

# ON THE BOUNDEDNESS OF SOLUTIONS OF A FORCED DISCONTINUOUS OSCILLATOR

TERE M-SEARA<sup>1,2,3</sup>, \*LUAN V. M. F. SILVA<sup>4</sup>, AND JORDI VILLANUEVA<sup>1</sup>

**ABSTRACT.** We study the global boundedness of the solutions of a non-smooth forced oscillator with a periodic and real analytic forcing. We show that the impact map associated with this discontinuous equation becomes a real analytic and exact symplectic map when written in suitable canonical coordinates. By an accurate study of the behaviour of the map for large amplitudes and by employing a parametrization KAM theorem, we show that the periodic solutions of the unperturbed oscillator persist as two-dimensional tori under conditions that depend on the Diophantine conditions of the frequency, but are independent on both the amplitude of the orbit and of the specific value of the frequency. This allows the construction of a sequence of nested invariant tori of increasing amplitude that confine the solutions within them, ensuring their boundedness. The same construction may be useful to address such persistence problem for a larger class of non-smooth forced oscillators.

## 1. INTRODUCTION

The study of the boundedness of solutions in Duffing-type equations has been a long-standing problem in dynamical systems, dating back to the seminal works of Littlewood in the 1960's [13, 14]. In the referred works, Littlewood proved the existence of unbounded solutions to the equation

$$(1) \quad \ddot{x} + g(x) = p(t),$$

with  $p(t)$  being bounded and periodic and for certain function  $g(x)$ , usually called saturation function. However, in contrast to this unbounded behavior, Littlewood [15] raised the question of whether all solutions of (1) would indeed be bounded, prompting further investigations into different conditions on  $p(t)$  and  $g(x)$  beyond the aforementioned ones. It was Morris [17] who first made a significant contribution by providing the initial example, solely under the assumption that  $p(t)$  is a periodic continuous function and  $g(x) = 2x^3$ . Some years later, Dieckerhoff and Zehnder [3] further expanded upon the Morris result by investigating the equation

$$\ddot{x} + x^{2n+1} + \sum_{j=0}^{2n} x^j p_j(t) = 0, \quad n \geq 1,$$

with  $p_j$  denoting periodic  $\mathcal{C}^\infty$ -functions.

As noted in [12], in 1998, during a presentation at the Academia Sinica, Ortega [21] introduced the idea of investigating the global stability of (1), assuming the presence of a bounded saturation function  $g(x)$ . He specifically suggested to take  $g(x) = \arctan(x)$  and aimed to identify the conditions on  $p(t)$  that would ensure this desired result. However, due to the bounded saturation function causing a subtle twist at infinity, applying the standard twist map theorem becomes challenging in obtaining global conclusions. In this circumstance, Li [12] was

---

\*Corresponding author: Luan V. M. F. Silva, luanmattos@ime.unicamp.br

2010 *Mathematics Subject Classification.* 37J40, 34A36, 70H08.

*Key words and phrases.* Non-smooth oscillators, boundedness of solutions, invariant tori, KAM theory, parametrization method.

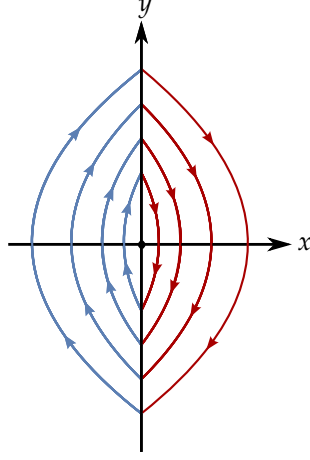


FIGURE 1. Phase portrait of (2) for the unperturbed case ( $\varepsilon = 0$ ).

the first to establish a result based on Ortega's proposal, assuming that  $p(t)$  is a  $C^\infty$  periodic function with zero average. This initial result was further enhanced by Wang in [28], where the regularity requirement on  $p(t)$  was relaxed to  $C^5$  while also considering a small condition on its average.

As a limit scenario within the framework proposed by Ortega [21], our investigation focuses on examining the boundedness of solutions for a non-smooth forced oscillator described by the equation

$$(2) \quad \ddot{x} + \text{sign}(x) = \varepsilon p(t),$$

with  $\text{sign}$  representing the standard sign function,  $\varepsilon \geq 0$  being a small real parameter, and the forcing term being a real analytic and  $2\pi$ -periodic function

$$(3) \quad p(t) = a_0 + \sum_{k \geq 1} (a_k \cos(kt) + b_k \sin(kt)),$$

so that it is an analytic function of  $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ .

In the unperturbed scenario  $\varepsilon = 0$ , all solutions of equation (2) are periodic (see Fig. 1). This implies that, in particular, all solutions remain bounded in the  $(x, \dot{x})$ -plane for  $\varepsilon = 0$ . A natural question that arises from this observation is whether the bounded nature of all solutions remains true when subjected to small perturbations.

Significant progress has been made regarding the analysis of equation (2) in the non-perturbative context  $\varepsilon = 1$ . For example, in [4], Enguiça and Ortega proved that equation (2) has infinitely many bounded solutions in the  $(x, \dot{x})$ -plane (of arbitrarily large amplitude) if the function  $p(t)$  is, not necessarily periodic, but bounded and satisfies that the limit

$$\bar{p} = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_t^{t+T} p(s) ds$$

exists uniformly in  $t \in \mathbb{R}$ , with  $|\bar{p}| < 1$ . Recently, Novaes and Silva [19] proved that all solutions of (2) remain bounded if  $p(t)$  is a Lebesgue integrable periodic function with vanishing average.

Further investigations related to the perturbative equation (2) can also be found in the research literature. For instance, Jacquemard and Teixeira [7] and Novaes and Silva [20] studied the persistence of periodic orbits, while Burra and Zanolin [1] discovered the existence of chaos near the origin under the assumption that  $p(t)$  is a piecewise constant and periodic function. The significance of equation (2) extends beyond theoretical interest, finding practical applications in the field of electronics. In this context, the presence of the sign function can be interpreted as a representation of a smooth oscillator operating in conjunction with a relay, as emphasized in [7].

The main result of this work is Theorem A, where we prove the boundedness of all solutions of (2) by assuming that  $p(t)$  is a real analytic and  $2\pi$ -periodic function, regardless of the value of the average of  $p(t)$ . The main idea behind the proof of Theorem A is to determine the existence, for any given  $\varepsilon > 0$  sufficiently small, of an infinite family of nested two-dimensional invariant tori in the extended phase space  $(t, x, \dot{x}) \in \mathbb{T} \times \mathbb{R}^2$ . The projection of these tori onto the  $(x, \dot{x})$ -plane surrounds the origin  $(0, 0)$  and their amplitude becomes unbounded as we move away from it. Therefore, each of these tori perpetually confines all solutions with initial conditions enclosed within it.

One standard approach to find these tori is to perform changes of variables to the discontinuous differential equations to overcome their lack of regularity, as described in [10, 11] and references therein. This approach leads to a finite differentiable perturbed system which allows the application of KAM theory to its stroboscopic Poincaré map. Subsequently, variants of Moser's twist map theorem [18] are employed to identify closed invariant curves for this map. By doing so, the existence of the tori of the original system can be determined. However, for equation (2), the application of such coordinate changes leads to a new system for which it is not apparent that we can establish a uniform upper bound  $\varepsilon^* > 0$  in a way that all solutions of (2) would be bounded if  $0 < \varepsilon < \varepsilon^*$ . Explicitly, achieving this result requires proving that there exist invariant tori of the system of arbitrarily large amplitude for any value of  $\varepsilon$  such that  $0 < \varepsilon < \varepsilon^*$ . So far, all efforts made to obtain this result by means of traditional techniques have not worked. The primary reason for this outcome is due to the fact that the twist condition of the system obtained after applying the different coordinate changes to (2) diminishes significantly as we progressively distance ourselves from the origin. Consequently, in order to be able to directly apply the classical procedure to (2), we must show that, after the transformation, the size of the resulting perturbation strongly decreases with this distance. Showing this strong decrease for this system does not appear to be an easy task, since we are then forced to follow closely the effect that these regularizing transformations have on the system, with the aggravating factor that this must be done in terms of a  $C^r$ -topology. So it is natural to think of a different approach. Indeed, our alternative approach allows, on the one hand, to generate an analytical impact map (which simplifies the study of its asymptotic properties with respect to a finite differentiable scenario) and, on the other hand, its construction arises naturally from the structure of the studied system, so it does not require discussing the effect that eventual regularizing transformations may have on it.

Our method to prove Theorem A involves analyzing the so called impact map associated with the equation (see Fig. 2 and equation (20)) in appropriate canonical time-energy coordinates (see equation (21)). These coordinates arise in a natural way after rewriting (2) as an autonomous Hamiltonian system (see equation (14)). In these coordinates, the impact map is a real analytic and exact symplectic map of the annulus (see Proposition 1), which turns out to be a perturbation of an integrable twist map. Although the impact map in these canonical coordinates possesses favorable properties regarding to the study of the persistence of the

invariant curves of the integrable approximation  $\varepsilon = 0$ , at first sight the same drawbacks outlined above regarding to the standard approximation persist for this analytic map. Explicitly, if we apply to it a KAM theorem for the persistence of invariant curves of analytic exact symplectic maps, it is easy to establish the persistence, up to a certain value of  $\varepsilon > 0$ , of any given unperturbed curve having a Diophantine rotation number. However, it is not clear that we can simultaneously establish the persistence of invariant curves of arbitrarily large amplitude.

**Remark 1.** *It is worth noting that the process we use to introduce these canonical time-energy coordinates can easily be extended, at least locally, to a very wide range of non-smooth forced oscillators. When these coordinates can be introduced, we have that the analyticity, the exact symplectic character and the perturbative form of the impact map follow at zero cost simply from the structure of the initial system. This means that if we are only interested in proving a local persistence result of invariant curves, such a result can usually be reached through a very moderate amount of work. The fact that proving Theorem A requires performing the elaborate process described below is because in our case we intend to prove a global persistence result.*

In order to prove Theorem A, it has been necessary to perform a precise analysis of the behaviour of the impact map for large amplitudes. We have identified its dominant terms and we have controlled the asymptotic size of the remaining ones. As a consequence of this analysis we have been able to carry out the following construction. We consider a specific invariant curve of the unperturbed impact map with sufficiently large amplitude and with frequency that satisfies the Diophantine estimates (26), for some pair  $(\gamma, \nu)$ . This curve is characterized by a large enough (in modulus) value of the action variable of the map, that we call  $E_0^*$ . Next, we localize the impact map around  $E_0^*$  and perform an appropriate conjugation (scaling) on this localized map, depending on  $E_0^*$ . In this way, we construct another analytic and exact symplectic map of the annulus. This map, which we can refer to as  $F_{E_0^*}$ , is different action by action and closely describes the dynamics around the selected curve. Regardless of the amplitude of the curve, for  $F_{E_0^*}$  we have both that the twist condition is always of order one and that the size of the perturbative terms of the corresponding integrable approximation can also be bounded uniformly for all the actions (see Proposition 3). Then, we study the persistence of invariant curves of the impact map of arbitrarily large amplitude by applying a parametrization KAM theorem for analytic and exact symplectic twist maps of the annulus (see Theorem B) to these maps  $F_{E_0^*}$ . Using this approach we do not have to worry about either the smallness of the twist condition or the control of the size of the non-integrable part as the amplitude increases. This approach of addressing the persistence of a family of tori of a specific system in terms of a parametric family of systems labelled by their frequency or frequency vector (here  $E_0^*$  determines the frequency) has been used previously in other contexts of KAM theory (see e.g. [9, 8]).

The final conclusion of this approach is that there exists  $\varepsilon^* > 0$  such that, for any  $0 < \varepsilon < \varepsilon^*$ , there are infinitely many invariant curves of the unperturbed impact map that survive the perturbation. The amplitude of these curves tends to infinite and so do their frequencies. But the Diophantine constant  $\gamma$  of all these frequencies can be bounded from below by the same value. Actually,  $\varepsilon^*$  only depends on this uniformly lower bound of  $\gamma$  and on the periodic function  $p(t)$ . By integrating by (2) these invariant curves of the impact map and expressed in terms of  $(t, x, \dot{x})$ , we construct the infinite collection of nested invariant tori  $\{\mathcal{S}_j^\varepsilon\}_{j=0}^\infty$  for the equation (2) in the extended phase space, thereby ensuring the boundedness of all its solutions, as long as  $\varepsilon$  remains within this specified range. It is worth mentioning that the existence of invariant tori using the KAM method with impact maps has also been investigated in other studies, as seen in [24, 25, 29, 30].

For the sake of completeness, in this paper we have not only included the statement of the parametrization KAM Theorem B, but we have also included its proof. Firstly, we have done this because Theorem B is perfectly suited for application to the local persistence of invariant curves of analytic and close to integrable exact symplectic twist maps of the annulus. Other results in the literature on parametrization KAM theory are sometimes developed for systems with a more general structure and, therefore, need to be interpreted appropriately in order to be applied to the present context. Furthermore, we also think that its rather schematic but comprehensive proof may be useful for the interested reader who is not familiar with this kind of KAM results.

Finally, although we have addressed Theorem A for a real analytic and periodic function  $p(t)$ , the same approach should be valid when  $p(t)$  is only finitely differentiable. The main modifications with respect to the analytic case should appear when performing the estimates of the impact map for large amplitudes, since the size of the derivatives involved should be controlled without using Cauchy estimates. Of course, we must also consider using an appropriate version of the twist theorem for finitely differentiable exact symplectic maps of the annulus.

The structure of this paper is outlined as follows. Section 2 provides the main result of the paper, Theorem A, as well as an overview of the different results (Propositions 1, 2, and 3 and Theorem B whose sequential application leads to Theorem A. In Section 3, we prove Theorem B, which is a KAM result (based in the parametrization method) for the persistence of quasi-periodic invariant curves of an analytic and exact symplectic twist map of the annulus. Section 4 examines the properties and bounds of the impact map associated with equation (2) and proves Proposition 2. Section 5 focuses on ensuring the exact symplectic properties and appropriate bounds of the localized and scaled impact map in suitable coordinates, as stated in Proposition 3. Finally, Section 6 offers a comprehensive result concerning the exact symplectic character for the impact map written in suitable variables (see Proposition 1).

## 2. MAIN RESULTS

The main result of this work concerns to the global stability of the solutions of (2), as follows.

**Theorem A.** *There exists  $\varepsilon^* > 0$ , depending on  $p$ , such that if  $0 \leq \varepsilon < \varepsilon^*$ , then all solutions of (2) are bounded.*

In this section we prove Theorem A by the recurrent application of the results and constructions presented below. The proof of each technical result necessary to establish the theorem is postponed to a later section of the paper, as is indicated in each case.

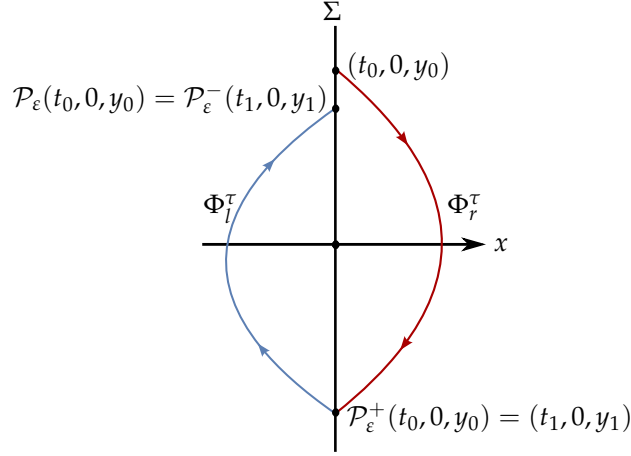
To prove Theorem A, we first check that a suitable impact map associated to (2), in good coordinates, is an analytic and exact-symplectic map which turns out to be a perturbation of an integrable twist map. Here, we briefly explain the formal construction of the impact map and we postpone the quantitative details to Section 4.

If we introduce  $y = \dot{x}$ , the differential equation (2) can be seen as the vector field

$$(4) \quad \begin{cases} \dot{t} = 1, \\ \dot{x} = y, \\ \dot{y} = -\text{sign}(x) + \varepsilon p(t), \end{cases}$$

in the extended phase space  $\mathbb{T} \times \mathbb{R}^2$ , where the presence of the function  $\text{sign}$  configures the plane  $\Sigma := \{(t, x, y) \in \mathbb{R}^3 : x = 0\}$  as a region of discontinuity of (4). The subsets of  $\Sigma$

$$(5) \quad \Sigma^+ := \{(t, 0, y) \in \Sigma : y > 0\} \quad \text{and} \quad \Sigma^- := \{(t, 0, y) \in \Sigma : y < 0\}$$

FIGURE 2. Impact map  $\mathcal{P}_\varepsilon = \mathcal{P}_\varepsilon^- \circ \mathcal{P}_\varepsilon^+$ .

play an important role in defining the domain of the half impact maps. We denote the solutions of (4) with initial condition  $(t_0, x_0, y_0)$  and  $(t_1, x_1, y_1)$ , if  $x_0 > 0$  and  $x_1 < 0$ , respectively, by

$$(6) \quad \Phi_r^\tau(t_0, x_0, y_0; \varepsilon) = (t_r^\tau(t_0, x_0, y_0; \varepsilon), x_r^\tau(t_0, x_0, y_0; \varepsilon), y_r^\tau(t_0, x_0, y_0; \varepsilon)),$$

and

$$(7) \quad \Phi_l^\tau(t_1, x_1, y_1; \varepsilon) = (t_l^\tau(t_1, x_1, y_1; \varepsilon), x_l^\tau(t_1, x_1, y_1; \varepsilon), y_l^\tau(t_1, x_1, y_1; \varepsilon)),$$

respectively, where  $t_r^\tau(t_0, x_0, y_0; \varepsilon) = t_0 + \tau$  and  $t_l^\tau(t_1, x_1, y_1; \varepsilon) = t_1 + \tau$ . For points in  $\Sigma^+$ , we will use  $\Phi_r^\tau$  as the flow in  $\Sigma^+$  “points to the right”. Analogously, for points in  $\Sigma^-$  we will use  $\Phi_l^\tau$  as the flow in  $\Sigma^-$  “points to the left”. By adopting the Filippov convention (see [5] for more detailed information) for the solutions of (4), we find out that such solutions are obtained by the concatenation of  $\Phi_r^\tau$  with  $\Phi_l^\tau$  along the crossing region  $\Sigma$ . This construction ensures the uniqueness and continuity of solutions. For  $(t_0, 0, y_0) \in \Sigma^+$ , we denote by  $\tau^+(t_0, y_0; \varepsilon)$  the smallest positive time such that

$$(8) \quad x_r^{\tau^+(t_0, y_0; \varepsilon)}(t_0, 0, y_0; \varepsilon) = 0.$$

We note that, as (4) depends periodically in time,  $\tau^+(t_0, y_0; \varepsilon)$  is  $2\pi$ -periodic in  $t_0$ . The half positive impact map is given by

$$(9) \quad \mathcal{P}_\varepsilon^+ : (t_0, 0, y_0) \in \Sigma^+ \mapsto (t_0 + \tau^+(t_0, y_0; \varepsilon), 0, y_r^{\tau^+(t_0, y_0; \varepsilon)}(t_0, 0, y_0; \varepsilon)) \in \Sigma^-.$$

The initial condition  $(t_1, 0, y_1) \in \Sigma^-$ , follows the left flow given by (7). We consider  $\tau^-(t_1, y_1; \varepsilon)$  as the smallest positive time such that

$$(10) \quad x_l^{\tau^-(t_1, y_1; \varepsilon)}(t_1, 0, y_1; \varepsilon) = 0,$$

with  $\tau^-(t_1, y_1; \varepsilon)$  being  $2\pi$  periodic in  $t_1$ . Then, the half negative impact map is given by

$$(11) \quad \mathcal{P}_\varepsilon^- : (t_1, 0, y_1) \in \Sigma^- \mapsto (t_1 + \tau^-(t_1, y_1; \varepsilon), 0, y_l^{\tau^-(t_1, y_1; \varepsilon)}(t_1, 0, y_1; \varepsilon)) \in \Sigma^+.$$

Thus, the complete impact map for equation (2) is given by the composition of the negative with the positive half impact map, respectively, i.e.,

$$\mathcal{P}_\varepsilon : (t_0, 0, y_0) \in \Sigma^+ \mapsto (\bar{t}_0, 0, \bar{y}_0) = (\mathcal{P}_\varepsilon^- \circ \mathcal{P}_\varepsilon^+)(t_0, 0, y_0) \in \Sigma^+,$$

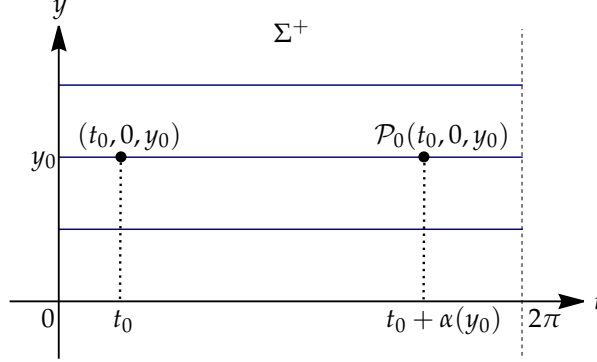


FIGURE 3. Intersections between the invariant tori in the unperturbed scenario and the subset  $\Sigma^+$ . In the curve  $\mathbb{T} \times \{0\} \times \{y_0\}$ , we can see the dynamics generated by the impact map  $\mathcal{P}_0$ .

with  $2\pi$ -periodic dependence in  $t_0$ , so that it can be read as a map of the annulus  $\mathbb{T} \times \{0\} \times \mathbb{R}^+$  (see Fig. 2).

When  $\varepsilon = 0$ , it is easy to check that

$$\mathcal{P}_0(t_0, 0, y_0) = (t_0 + \alpha(y_0), 0, y_0),$$

with  $\alpha(y_0) = 4y_0$  is the period of the periodic solution of the unperturbed system (2) with initial condition  $(t_0, 0, y_0)$ . Moreover, given  $y_0 > 0$  and for sufficiently small  $\varepsilon \geq 0$ , the trajectories of (2) starting in  $\Sigma^+$  cross  $\Sigma^+$  again, then  $\mathcal{P}_\varepsilon$  is well defined and also analytic. Since  $\alpha'(y_0) = 4 > 0$ , for all  $y_0$ , it follows that  $\mathcal{P}_0$  is an integrable twist map of the annulus and any circle of the form  $\mathbb{T} \times \{0\} \times \{y_0\}$ , for all  $y_0 > 0$ , is an invariant curve of  $\mathcal{P}_0$  with frequency  $\alpha(y_0)$  (see Fig. 3). For those  $y_0$  such that its associated rotation number  $\alpha(y_0)/2\pi$  is irrational, the motion on this curve is quasi-periodic. In Lemmas 4 and 5 we provide asymptotic expressions for the times of impact  $\tau^+$  and  $\tau^-$ , respectively, for large (enough) amplitudes and whenever  $0 \leq \varepsilon < \varepsilon_*^+$  and  $0 \leq \varepsilon < \varepsilon_*^-$ , respectively, with  $\varepsilon_*^+$  and  $\varepsilon_*^-$  depending only on the function  $p$ . These results allow us to control how far  $\mathcal{P}_\varepsilon$  is to be an integrable map.

**Remark 2.** As the map  $\mathcal{P}_\varepsilon$  is defined on  $\Sigma^+ \subset \{x = 0\}$ , until the contrary is said, we will omit the  $x$ -variable and treat  $\mathcal{P}_\varepsilon$  as a function of  $(t_0, y_0)$ , regarding it as a map on the annulus. The same convention will be applied to the maps  $\mathcal{P}_\varepsilon^+$  and  $\mathcal{P}_\varepsilon^-$ .

A fundamental concept for our approach is the exact-symplectic character for twist maps of the annulus. However, considering the cylinder  $\mathbb{T} \times \mathbb{R}^+$  endowed with the 2-form  $dt_0 \wedge dy_0$ , we cannot prove that the map  $\mathcal{P}_\varepsilon$  is symplectic. To overcome this obstacle, we re-write it in terms of a new couple of variables:  $t_0$  and the associated symplectic conjugate, i.e., the energy  $E_0$  according to the (piecewise) Hamiltonian structure of (4) (see (14)). The exact symplectic character of  $\mathcal{P}_\varepsilon$  in these variables follows at once from the application of Proposition 1 to this piecewise Hamiltonian. We postpone the proof of this proposition to Section 6.

**Proposition 1.** Let  $H(x, y, t)$  be a non-autonomous Hamiltonian, with respect to the 2-form  $dx \wedge dy$ , and with  $2\pi$ -periodic dependence in  $t$ . By adding  $E$  as a conjugate variable of  $t$ , we introduce the autonomous Hamiltonian  $\mathcal{H}(x, t, y, E) = H(x, y, t) + E$ , with respect to the 2-form  $dx \wedge dy + dt \wedge dE$ . We denote the solutions associated with  $\mathcal{H}$  by

$$(x^\tau(x_0, t_0, y_0, E_0), t^\tau(x_0, t_0, y_0, E_0), y^\tau(x_0, t_0, y_0, E_0), E^\tau(x_0, t_0, y_0, E_0)),$$

with  $t^\tau(x_0, t_0, y_0, E_0) = t_0 + \tau$ , where  $\tau$  represents the new time, and  $(x_0, t_0, y_0, E_0)$  denotes the initial conditions at time  $\tau = 0$ .

We consider the section  $\Xi = \{x = 0\} \cap \{\mathcal{H} = 0\}$  and we suppose that the equation  $\mathcal{H}(0, t, y, E) = 0$  allows writing, at least locally for  $(t, E) \in U \subset \mathbb{R}^2$ , the variable  $y$  as  $y = y(t, E)$ , being  $y(t, E)$  a smooth function with  $2\pi$ -periodic dependence in  $t$ . We also suppose that it is possible to define the impact time  $\tilde{\tau}(t_0, E_0)$  as the positive time for which we have

$$(12) \quad x^{\tilde{\tau}(t_0, E_0)}(0, t_0, y(t_0, E_0), E_0) = 0,$$

for any  $(t_0, E_0) \in U$ , being  $\tilde{\tau}(t_0, E_0)$  a smooth function with  $2\pi$ -periodic dependence in  $t_0$ . Then, associated to this section  $\Xi$ , we introduce the map:

$$F : U \longrightarrow \mathbb{R}^2, \\ (t_0, E_0) \longmapsto (t_1, E_1) = (t_0 + f_{t_0}(t_0, E_0), f_{E_0}(t_0, E_0)),$$

where:

$$f_{t_0}(t_0, E_0) = \tilde{\tau}(t_0, E_0),$$

$$f_{E_0}(t_0, E_0) = E^{\tilde{\tau}(t_0, E_0)}(0, t_0, y(t_0, E_0), E_0) = -H(0, y^{\tilde{\tau}(t_0, E_0)}(0, t_0, y(t_0, E_0), E_0), t_0 + \tilde{\tau}(t_0, E_0)),$$

meaning that if we consider initial conditions  $(0, t_0, y(t_0, E_0), E_0) \in \Xi$ , then the corresponding solution integrated up to time  $\tilde{\tau}(t_0, E_0)$ , given by  $(0, t_1, y_1, E_1)$ , with  $y_1 = y^{\tilde{\tau}(t_0, E_0)}(0, t_0, y(t_0, E_0), E_0)$ , also belongs to  $\Xi$ . Then, the map  $F$  is exact symplectic with respect to the 1-form  $E_0 dt_0$ , i.e., there is a function  $S(t_0, E_0)$ ,  $2\pi$ -periodic in  $t_0$ , such that  $F^*(E_1 dt_1) = E_0 dt_0 + dS(t_0, E_0)$ .

We notice that the solutions of (4) are solutions of the non-autonomous Hamiltonian system for  $(x, y)$ , with continuous piecewise analytic Hamilton function

$$(13) \quad \begin{cases} H_\varepsilon^r(x, y, t) = \frac{y^2}{2} + x(1 - \varepsilon p(t)) & \text{if } x > 0, \\ H_\varepsilon^l(x, y, t) = \frac{y^2}{2} - x(1 + \varepsilon p(t)) & \text{if } x < 0. \end{cases}$$

Adding the energy  $E$  as a conjugate variable to the time  $t$ , we transform  $H_\varepsilon^\sigma$  into the two-degrees of freedom autonomous Hamiltonian

$$(14) \quad \mathcal{H}_\varepsilon^\sigma(x, t, y, E) = H_\varepsilon^\sigma(x, y, t) + E, \quad \sigma = r, l,$$

for which we denote as  $\tau$  the new time variable. Notice that if we restrict to the zero energy level  $\mathcal{H}_\varepsilon^\sigma = 0$ , then we recover all the solutions of  $H_\varepsilon^\sigma$ , for  $\sigma = r, l$ . If the solutions  $t_\sigma^\tau$ ,  $x_\sigma^\tau$ , and  $y_\sigma^\tau$  are known, then the expressions for  $E_\sigma^\tau$  are explicit from the relation  $\mathcal{H}_\varepsilon^\sigma = 0$ . So, based on the construction of the half impact maps (9) and (11), it follows from Proposition 1 that both impact maps  $\mathcal{P}_\varepsilon^+$  and  $\mathcal{P}_\varepsilon^-$ , in canonical coordinates  $(t_0, E_0)$ , give rise to exact symplectic maps  $\bar{\mathcal{P}}_\varepsilon^+$  and  $\bar{\mathcal{P}}_\varepsilon^-$ , respectively, with respect to the 1-form  $E_0 dt_0$ . Therefore their composition, to be denoted as  $\bar{\mathcal{P}}_\varepsilon = \bar{\mathcal{P}}_\varepsilon^- \circ \bar{\mathcal{P}}_\varepsilon^+$ , is also exact symplectic.

From expressions (13) and (14), it follows that both equations  $\mathcal{H}_\varepsilon^\sigma = 0$ , for  $\sigma = r, l$ , give rise, for  $x = 0$ , to the relation  $y^2/2 + E = 0$ . So, once we have computed the map  $(\bar{t}_0, \bar{y}_0) = \mathcal{P}_\varepsilon(t_0, y_0)$ , then the exact-symplectic map  $(\bar{t}_0, \bar{E}_0) = \bar{\mathcal{P}}_\varepsilon(t_0, E_0)$  is obtained by replacing  $y_0 > 0$  by  $E_0 = -y_0^2/2 < 0$  and  $\bar{y}_0 > 0$  by  $\bar{E}_0 = -\bar{y}_0/2$ .

To apply KAM theory to  $\bar{\mathcal{P}}_\varepsilon$  (in fact, to the localized and scaled map introduced in Proposition 3) we need, in particular, to handle the regularity of the map. In this paper, we regard into the analytical context as our framework for applying KAM theory. Although the differential equation (2) is discontinuous, both  $\bar{\mathcal{P}}_\varepsilon^+$  and  $\bar{\mathcal{P}}_\varepsilon^-$  are analytic maps. This is because (2) is



piecewise analytic and both impact times  $\tau^+(t_0, y_0; \varepsilon)$  and  $\tau^-(t_1, y_1; \varepsilon)$  are analytic functions of the initial conditions. Next point is to compute an integrable approximation for the maps  $\mathcal{P}_\varepsilon(t_0, y_0)$  and  $\bar{\mathcal{P}}_\varepsilon(t_0, E_0)$  and to control how far are them from this integrable approximation as  $y_0 \rightarrow +\infty$  and  $E_0 \rightarrow -\infty$ , respectively. To achieve these purposes, we need to ensure that both  $\tau^+(t_0, y_0; \varepsilon)$  and  $\tau^-(t_1, y_1; \varepsilon)$  are well defined for any  $y_0 \gg 0$  and  $y_1 \ll 0$ , respectively, and to accurately analyze their behaviour for large values of  $y_0$  and  $y_1$ . Since we are interested in dealing with analytical dependence, computations concerning  $\tau^+$  and  $\tau^-$  must be carried out not only for real values of the initial conditions, but also for values of  $(t_0, y_0)$  and  $(t_1, y_1)$  in appropriate complex strips around the real domains selected for each of these four variables. The width of these analyticity domains can be selected independently of  $\varepsilon$ , provided that  $\varepsilon > 0$  is sufficiently small. We use the estimates on  $\tau^+(t_0, y_0; \varepsilon)$  and  $\tau^-(t_1, y_1; \varepsilon)$  in complex domains to control  $\mathcal{P}_\varepsilon^+$ ,  $\mathcal{P}_\varepsilon^-$ , and  $\mathcal{P}_\varepsilon$ . Since this is the most technical part of this work, we postpone the details to Section 4, and give here only the final result concerning the behaviour of the impact map  $\mathcal{P}_\varepsilon$  in Proposition 2 below.

We introduce some notations to be used throughout the paper. First, we denote by  $|z|$  the Euclidean norm of any complex number  $z \in \mathbb{C}$  and we extend the same notation for the sup-norm of any complex valued vector or matrix. We also introduce:

$$(15) \quad \Delta(\rho) = \{\theta \in \mathbb{C} : |\operatorname{Im}(\theta)| < \rho\}, \quad \|f\|_\rho = \sup_{\theta \in \Delta(\rho)} \{|f(\theta)|\}.$$

Here,  $f = f(\theta)$  is a function (real valued, vector valued or complex valued),  $2\pi$ -periodic in  $\theta$ , that can be analytically extended to the complex strip  $\Delta(\rho)$  of width  $\rho > 0$  and that it is bounded up to the boundary of this strip. Given positive quantities  $y^*$ ,  $\bar{\rho}$ , and  $\tilde{\rho}$ , we define the following sets of  $\mathbb{C}^2$ :

$$(16) \quad \mathcal{D}^+(y^*, \bar{\rho}, \tilde{\rho}) = \{(t_0, y_0) \in \Delta(\bar{\rho}) \times \Delta(\tilde{\rho}) : |y_0| > y^* \text{ and } \operatorname{Re}(y_0) > 0\},$$

$$(17) \quad \mathcal{D}^-(y^*, \bar{\rho}, \tilde{\rho}) = \{(t_1, y_1) \in \Delta(\bar{\rho}) \times \Delta(\tilde{\rho}) : |y_1| > y^* \text{ and } \operatorname{Re}(y_1) < 0\}.$$

We will check that both  $\mathcal{P}_\varepsilon^+$  and  $\mathcal{P}_\varepsilon$  are defined in sets of the form (16), while  $\mathcal{P}_\varepsilon^-$  is defined in a set of the form (17).

We introduce the  $2\pi$ -periodic real analytic functions  $\tilde{P}_j(t)$ , for  $j = -1, 0, 1, 2$ , defined as:

$$(18) \quad \tilde{P}_0(t) = p(t) - a_0, \quad a_0 = \langle p \rangle, \quad \tilde{P}_1'(t) = \tilde{P}_0(t), \quad \tilde{P}_2'(t) = \tilde{P}_1(t), \quad \tilde{P}_{-1}(t) = \tilde{P}_0'(t),$$

where  $p(t)$  is the real analytic  $2\pi$ -periodic function defined in (3) and  $\langle f \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) dt$  denotes the average of a  $2\pi$ -periodic function  $f(t)$ . The functions  $\tilde{P}_1(t)$  and  $\tilde{P}_2(t)$  are unique under the normalization  $\langle \tilde{P}_j \rangle = 0$ ,  $j = 1, 2$ . Due to the analyticity of  $p(t)$ , there are constants  $0 < \rho < 1$  and  $\tilde{\rho} > 0$  such that

$$(19) \quad \|\tilde{P}_i\|_\rho \leq \tilde{\rho}, \quad \text{for } j = -1, 0, 1, 2.$$

Both,  $\rho$  and  $\tilde{\rho}$ , are set fixed throughout the paper.

**Proposition 2.** *With notations, definitions and hypotheses above, there are positive constants  $\bar{\rho}$  and  $\tilde{\rho}$ , only depending on  $\rho$ , and positive numbers  $0 < \varepsilon_{\mathcal{P}}^* < \min\{1, 1/2|a_0|\}$  and  $\tilde{C}$ , only depending on  $a_0$ ,  $\rho$ , and  $\tilde{\rho}$ , such that the following holds. For any  $0 \leq \varepsilon < \varepsilon_{\mathcal{P}}^*$ , the impact map  $(\bar{t}_0, \bar{y}_0) = \mathcal{P}_\varepsilon(t_0, y_0)$  is well defined and analytic if  $(t_0, y_0) \in \mathcal{D}^+(4, \bar{\rho}, \tilde{\rho})$  (see (16)), and takes the form*

$$(20) \quad \mathcal{P}_\varepsilon : \begin{cases} \bar{t}_0 = t_0 + \alpha_\varepsilon(y_0) + \varepsilon f_{t_0}(t_0, y_0; \varepsilon), \\ \bar{y}_0 = y_0 + \varepsilon f_{y_0}(t_0, y_0; \varepsilon), \end{cases}$$

with  $\alpha_\varepsilon(y_0) = \frac{4y_0}{1-a_0^2\varepsilon^2}$ , being the functions  $f_{t_0}(t_0, y_0; \varepsilon)$  and  $f_{y_0}(t_0, y_0; \varepsilon)$   $2\pi$ -periodic in  $t_0$ . Furthermore, the following estimates hold:

$$\left| \partial_{t_0}^i \partial_{y_0}^j f_{t_0}(t_0, y_0; \varepsilon) \right| \leq \begin{cases} \tilde{C} & \text{if } j = 0, \\ \frac{\tilde{C}}{|y_0|} & \text{if } j \neq 0, \end{cases} \quad \text{and} \quad \left| \partial_{t_0}^i \partial_{y_0}^j f_{y_0}(t_0, y_0; \varepsilon) \right| \leq \tilde{C},$$

for any  $(t_0, y_0) \in \mathcal{D}^+(4, \bar{\rho}, \bar{\rho})$ ,  $0 \leq \varepsilon < \varepsilon_p^*$ , and  $0 \leq i + j \leq 2$ .

The impact map  $\mathcal{P}_\varepsilon(t_0, y_0)$  lacks the desired exact symplectic character with respect to the 2-form  $dt_0 \wedge dy_0$  that we would like to have in order to be able to apply KAM theory to it. However, as noted above, this exactness in terms of the 2-form  $dt_0 \wedge dE_0$  is achieved by replacing the variable  $y_0 > 0$  by  $E_0 = -y_0^2/2 < 0$ . In this way, we obtain an analytic exact symplectic map  $(\bar{t}_0, \bar{E}_0) = \bar{\mathcal{P}}_\varepsilon(t_0, E_0)$ , which takes the form:

$$(21) \quad \bar{\mathcal{P}}_\varepsilon : \begin{cases} \bar{t}_0 = t_0 + \bar{\alpha}_\varepsilon(E_0) + \varepsilon \bar{f}_{t_0}(t_0, E_0; \varepsilon), \\ \bar{E}_0 = E_0 + \varepsilon \bar{f}_{E_0}(t_0, E_0; \varepsilon), \end{cases}$$

with  $\bar{\alpha}_\varepsilon(E_0) = \frac{4\sqrt{-2E_0}}{1-a_0^2\varepsilon^2}$ , being the functions  $\bar{f}_{t_0}(t_0, E_0; \varepsilon)$  and  $\bar{f}_{E_0}(t_0, E_0; \varepsilon)$   $2\pi$ -periodic in  $t_0$ .

The complex domain of definition for  $E_0$  of  $\bar{\mathcal{P}}_\varepsilon$ , as well as the corresponding estimates for  $\bar{f}_{t_0}$  and  $\bar{f}_{E_0}$ , follow at once from the results of Proposition 2 on  $\mathcal{P}_\varepsilon$ , only taking into account the relations  $E_0 = -y_0^2/2 < 0$  and  $\bar{E}_0 = -\bar{y}_0^2/2$ . However, we are not going to state the analogue of Proposition 2 for  $\bar{\mathcal{P}}_\varepsilon$  since, as noted above, our parameterization KAM theorem is not directly applied to  $\bar{\mathcal{P}}_\varepsilon$  but to a map obtained through a process of localization and scaling of  $\bar{\mathcal{P}}_\varepsilon$  around a specific action  $E_0^* \ll 0$ . This map is referred to as the scaled impact map, and its main properties are detailed in Proposition 3.

The map  $\bar{\mathcal{P}}_\varepsilon$  takes the form of a close to integrable twist map of the annulus  $\mathbb{T} \times \mathbb{R}$ . Thus, the natural question arising from this fact is: which of the invariant curves for the unperturbed map (when  $\varepsilon = 0$ ) persist after small perturbations? This question could be answered by means of the celebrated KAM theory, like the Moser's twist mapping theorem [18] as well as other versions of this theorem (see e.g. [11, 16, 22, 23]). However, all these results provide conditions for such persistence which, in particular, depend explicitly on the region considered and on the fact that some estimates uniformly hold in the selected region. Those reasons make very difficult the task to obtain persistence of invariant curves in unbounded domains, under conditions independent of the amplitude of the curve that we want to persist. Since the twist condition  $\bar{\alpha}'_\varepsilon(E_0)$  for the impact map  $\bar{\mathcal{P}}_\varepsilon$  tends to zero as  $E_0 \rightarrow -\infty$ , we do not know any KAM result that can be globally applied in our context to show the persistence of curves of arbitrarily large amplitude.

In this work we address the persistence of these invariant curves in terms of a result based on the parametrization KAM theory introduced in [2] (see [6] for a wide overview of parametrization techniques in dynamical systems). Specifically, in Theorem B we provide a KAM theorem concerning the persistence of a specific quasi-periodic invariant curve for a close to integrable, analytic and exact symplectic twist map  $F$  of the form:

$$(22) \quad (\phi, I) \in \mathbb{T} \times \mathbb{R} \mapsto F(\phi, I) = (\phi + \alpha(I) + f_\phi(\phi, I), I + f_I(\phi, I)) \in \mathbb{T} \times \mathbb{R},$$

endowed by the 2-form  $d\phi \wedge dI = -d(Id\phi)$ , i.e., verifying  $F^*(Id\phi) = Id\phi + dV(\phi, I)$ , for some function  $V(\phi, I)$  depending  $2\pi$ -periodically in  $\phi$ . By a specific invariant curve we mean the curve that is a perturbation of  $\mathbb{T} \times \{I_0^*\}$ , for some  $I_0^*$ , but that has the same frequency  $\omega = \alpha(I_0^*)$ . Explicitly, given an analytic curve  $\mathcal{T} \subset \mathbb{T} \times \mathbb{R}$ , we say that it is  $F$ -invariant with

frequency  $\omega$  if there is an analytic parametrization  $\varphi : \mathbb{T} \rightarrow \mathbb{T} \times \mathbb{R}$  of the curve  $\mathcal{T}$  for which the following equation holds

$$(23) \quad F(\varphi(\theta)) = \varphi(\theta + \omega),$$

for all  $\theta \in \mathbb{R}$ . We refer to (23) as the invariance equation for  $\varphi$ . Then, if  $\varphi$  verifies (23), the curve  $\mathcal{T} = \varphi(\mathbb{T})$  is invariant by the map  $F$  and the pull-back by  $\varphi$  of the dynamics on  $\mathcal{T}$  becomes the rigid rotation on  $\mathbb{T}$  of frequency  $\omega$ , i.e.,  $T_\omega(\cdot) = \cdot + \omega$ . Hence, the rotation number of  $\mathcal{T}$  is  $\omega/2\pi$  and the dynamics on the curve is quasi-periodic if  $\omega/2\pi$  is an irrational number. Actually, to discuss the persistence of this curve, we are going to assume that  $\omega/2\pi$  is a Diophantine number of type  $(\gamma, \nu)$  (see (26)). Since we are interested in invariant curves which are isotopic to  $\mathbb{T} \times \{I_0^*\}$ , we are going to consider parametrizations of the form  $\varphi(\theta) = (\theta + \varphi_\phi(\theta), \varphi_I(\theta))$ , with  $\varphi_\phi$  and  $\varphi_I$  being  $2\pi$ -periodic in  $\theta$ .

The idea of the parametrization method is to solve for  $\varphi$  the equation (23) by means of a quasi-Newton method. This method iteratively modifies  $\varphi$ , but not the map  $F$  which remains unchanged along the process. This fact eases the discussion of which conditions on  $F$  are needed in order to ensure the persistence of the target curve with respect to the classical approaches to KAM theory which are based on the application of a sequence of canonical transformations to the map  $F$ . To measure the distance of  $\varphi$  from being a solution of (23), the invariance error associated to  $\varphi$  is defined as

$$(24) \quad e(\theta) = F(\varphi(\theta)) - \varphi(\theta + \omega), \quad \theta \in \mathbb{T}.$$

Hence, it is usually said that  $\mathcal{T} = \varphi(\mathbb{T})$  is a quasi-torus of  $F$  if the  $2\pi$ -periodic function  $e$  is sufficiently small, in the sense of the norm  $\|e\|_\rho$  for some  $\rho > 0$ .

**Theorem B.** Consider  $F = F(\phi, I)$  a real analytic and exact symplectic map of the annulus  $\mathbb{T} \times \mathbb{R}$  of the form (22), take a particular action  $I_0^* \in \mathbb{R}$  and denote  $\omega = \alpha(I_0^*)$ . We suppose:

(H1)  $F$  can be analytically extended to  $\mathcal{U} := \Delta(\rho_0) \times D(I_0^*, R_0)$ , for some  $\rho_0, R_0 \in (0, 1)$ , with  $\Delta(\rho_0)$  being the complex strip introduced in (15) and  $D(I_0^*, R_0)$  being the complex disc:

$$(25) \quad D(I_0^*, R_0) := \{I \in \mathbb{C} : |I - I_0^*| < R_0\}.$$

(H2)  $\omega/2\pi$  is a Diophantine number of type  $(\gamma, \nu)$ , for some  $0 < \gamma \leq 1$  and  $\nu \geq 2$ , i.e.,

$$(26) \quad \left| \frac{\omega}{2\pi} - \frac{p}{q} \right| \geq \frac{\gamma}{q^\nu}, \quad \forall p \in \mathbb{Z}, \forall q \in \mathbb{N}.$$

(H3) There are positive constants  $c_1, c_2, c_3$  and  $c$ , with  $c \leq 1$ , for which the functions  $\alpha, f_\phi$  and  $f_I$  in (22) satisfy:

$$c_1 \leq |\alpha'(I)| \leq c_2, \quad |\alpha''(I)| \leq c_3, \quad |\partial_\phi^i \partial_I^j f_\phi(\phi, I)| \leq c, \quad |\partial_\phi^i \partial_I^j f_I(\phi, I)| \leq c,$$

for every  $(\phi, I) \in \mathcal{U}$  and  $0 \leq i + j \leq 2$ . In particular,  $\alpha'(I_0^*) \neq 0$  and the (unperturbed) map  $(\phi, I) \rightarrow (I + \alpha(I), I)$  is a twist map if  $I \approx I_0^*$ .

Then, there is a constant  $c^* \geq 1$ , that only depends on  $1/c_1, c_2, c_3$  and  $\nu$ , and that can be defined in such a way it is an increasing function of all these quantities, for which the following holds. Assume that  $c$  is small enough so that it verifies

$$(27) \quad c^* c \leq \gamma^2 \delta_0^{2\nu} \min\{c_1 \delta_0/8, \delta_0/12, R_0/3\}, \quad c^* c \leq 2^{-4\nu-2} \gamma^4 \delta_0^{4\nu+1},$$

where  $\delta_0 = \rho_0/12$ . Then, there exists a real analytic function  $\varphi^* : \mathbb{T} \rightarrow \mathbb{T} \times \mathbb{R}$ , of the form  $\varphi^*(\theta) = (\theta + \varphi_\phi^*(\theta), \varphi_I^*(\theta))$ , with  $\varphi_\phi^*(\theta), \varphi_I^*(\theta)$  being  $2\pi$ -periodic in  $\theta$ , and with  $(\varphi^*)'(\theta) \neq 0$  for all  $\theta$ , which turns out to be a solution of the invariance equation (23). Consequently,  $\varphi^*$  is the parametrization of

an invariant curve  $\mathcal{T} = \varphi^*(\mathbb{T})$  of  $F$  with frequency  $\omega$ . Moreover,  $\varphi$  is defined for  $\theta \in \Delta(\rho_0/2)$  and satisfies:

$$(28) \quad \left\| \varphi_\phi^* \right\|_{\rho_0/2} \leq \frac{2c^*c}{\gamma^2\delta_0^{2\nu}}, \quad \left\| \varphi_I^* - I_0^* \right\|_{\rho_0/2} \leq \frac{2c^*c}{\gamma^2\delta_0^{2\nu}}, \quad \left\| (\varphi_s^*)' \right\|_{\rho_0/2} \leq \frac{2c^*c}{\gamma^2\delta_0^{2\nu+1}},$$

for  $s \in \{\phi, I\}$ .

The proof of Theorem B is postponed to Section 3 which in turn is presented in a very synthetic way, since it is an adaptation to the context of previous results on parametrization KAM theory. The interested reader is referred to the references [2, 6] quoted above for more details and for the geometric motivation of the constructions used to prove it.

**Remark 3.** In order to apply Theorem B, we need  $c$  (“the size of the perturbation”) to be smaller than an expression of  $\mathcal{O}(\gamma^4)$ , while classical KAM theory results usually only require  $c$  to be smaller than an expression of  $\mathcal{O}(\gamma^2)$ . This discrepancy in the  $\gamma$  exponent between classical and parametrization methods in KAM theory is due to the way in which the proof of the parametrization results is usually carried out. Getting a  $\mathcal{O}(\gamma^2)$  condition for  $c$  using KAM parametrization methods requires a more sophisticated approach than the one used here to prove Theorem B. Such approach (see [27] for the details) is not considered here because a  $\mathcal{O}(\gamma^4)$  condition for  $c$  is sufficient for our purposes.

Using Theorem B we are going to show that the persistence conditions of any unperturbed curve of the impact map  $\bar{\mathcal{P}}_\varepsilon$  (of large enough amplitude) do not depend on the amplitude of this curve, but only on the Diophantine constant  $\gamma$  of its frequency  $\omega$ . However, despite  $\bar{\mathcal{P}}_\varepsilon$  in (21) is an analytic and quasi-integrable exact symplectic twist map of the annulus, some conditions of Theorem B are not satisfied by it when  $E_0 \rightarrow -\infty$ . We mainly stress the fact that the twist condition becomes extremely small, since  $\bar{\alpha}'_\varepsilon(E_0) = -\frac{4}{1-a_0^2\varepsilon^2} \frac{1}{\sqrt{-2E_0}}$ . To address these shortcomings, we propose a construction in which Theorem B is not applied directly to the map  $\bar{\mathcal{P}}_\varepsilon$ , but to a collection of maps that follow from the localization of  $\bar{\mathcal{P}}_\varepsilon$  around a specific curve  $\mathbb{T} \times \{E_0^*\}$  of the unperturbed approximation, followed by an appropriate scaling depending on the selected action  $E_0^*$  (see Proposition 3). So, the specific map to which we apply Theorem B is different curve by curve, but the persistence conditions will be uniform in  $\varepsilon$  for all  $E_0^* \ll 0$  and, therefore, for all these maps.

We consider fixed values of  $\varepsilon > 0$  and of the energy variable  $E_0^* < 0$ . We suppose that, for this couple,  $\omega = \bar{\alpha}_\varepsilon(E_0^*)$  is well defined (later on we will ask  $\omega/2\pi$  to satisfy an appropriate Diophantine condition (26)). Then, we consider the change of variables (scaling) adapted to the selected value of  $E_0^*$ :

$$(29) \quad \bar{\psi} : (\phi, I) \mapsto (t_0, E_0/\sqrt{-E_0^*}).$$

In particular, the new action  $I$  corresponding to the selected energy variable  $E = E_0^*$  is  $I = I_0^* = -\sqrt{-E_0^*}$ . The map  $F(\phi, I)$  that we obtain after conjugation of  $\bar{\mathcal{P}}_\varepsilon$  by  $\bar{\psi}$  (we omit the explicit  $\varepsilon$  dependence on  $F$  since  $\varepsilon$  is set fixed) is a real analytic exact symplectic map, that takes the form (22), and that verifies the properties stated in Proposition 3 below. For a better understanding of the statement of the Proposition 3, we point out that the relation between the variable  $I < 0$  introduced in (29) and the variable  $y_0 > 0$  of Proposition 2 is

$$(30) \quad y_0 = y_0(I) = \sqrt{-\sqrt{2}y_0^*I},$$

where  $y_0^* = -\sqrt{2}I_0^* = \sqrt{-2E_0^*}$  is the value of  $y_0$  related to  $E_0 = E_0^*$  and  $I = I_0^*$ .

**Proposition 3.** *With the same notations and hypotheses of Proposition 2, we consider a fixed couple  $0 < \varepsilon < \varepsilon_{\mathcal{P}}^*$  and  $y_0^* > 5$ , and define  $E_0^* = -(y_0^*)^2/2$  and  $I_0^* = -y_0^*/\sqrt{2}$ . We then perform the scaling (29) to the map  $\bar{\mathcal{P}}_\varepsilon$  of (21) and we obtain a map  $F(\phi, I)$  of the form*

$$(31) \quad F(\phi, I) = (\phi + \alpha(I) + f_\phi(\phi, I), I + f_I(\phi, I)),$$

with  $\alpha(I) = \frac{4}{1-a_0^2\varepsilon^2}\sqrt{-\sqrt{2}y_0^*I}$  and  $f_\phi(\phi, I)$  and  $f_I(\phi, I)$  being  $2\pi$ -periodic in  $\phi$ . The map  $F$  is exact symplectic with respect to the 1-form  $\text{Id}\phi$  and it is real analytic if  $(\phi, I) \in \Delta(\bar{\rho}) \times D(I_0^*, \bar{\rho})$  (see (15) and (25)), where  $\bar{\rho}$  and  $\bar{\rho}$  are given in Proposition 2. Moreover, there is a constant  $\bar{C}$ , that only depends on  $a_0$ ,  $\rho$ , and  $\bar{p}$  (see (18) and (19)), such that

$$\frac{5\sqrt{2}}{3} < |\alpha'(I)| < \frac{10\sqrt{2}}{3}, \quad |\alpha''(I)| \leq \frac{25}{24}, \quad \left| \partial_\phi^i \partial_I^j f_\phi(\phi, I) \right| \leq \varepsilon \bar{C}, \quad \left| \partial_\phi^i \partial_I^j f_I(\phi, I) \right| \leq \varepsilon \bar{C},$$

for any  $(\phi, I) \in \Delta(\bar{\rho}) \times D(I_0^*, \bar{\rho})$ ,  $0 < \varepsilon < \varepsilon_{\mathcal{P}}^*$ , and  $0 \leq i + j \leq 2$ .

The proof of Proposition 3 is given in Section 5. It is worth noting that, if we vary the values of  $I^* \rightarrow -\infty$ , after the scaling we have that the size of the twist condition  $\alpha'(I)$  has lower an upper bounds independent on  $I^*$  in a disk of center  $I^*$  an radius  $\bar{\rho}$  independent on  $I^*$ .

To conclude the proof of Theorem A, we observe that the scaled impact map (31) satisfies the conditions of Theorem B. Explicitly, we select a fixed  $0 \leq \varepsilon < \varepsilon_{\mathcal{P}}^*$ , where  $\varepsilon_{\mathcal{P}}^*$  is provided in Proposition 2, and consider a value of  $y = y_0^* > 5$  for the impact map  $\mathcal{P}_\varepsilon$  of the initial system (2) (see Fig. 2). According to the selected scaling, to such value  $y = y_0^*$  it corresponds the action value  $I = I_0^* = -y_0^*/\sqrt{2}$ . Hence, for this given couple  $(\varepsilon, y_0^*)$ , the corresponding frequency of the unperturbed map  $(\phi, I) \rightarrow (\phi + \alpha(I), I)$  in (31) is given by  $\omega_{y_0^*}^\varepsilon = \alpha(I_0^*) = \frac{4}{1-a_0^2\varepsilon^2}y_0^*$ . It is clear that, for a fixed  $\varepsilon$ , the expression  $\omega_{y_0^*}^\varepsilon$  is a strictly increasing function of  $y_0^*$  that goes to  $+\infty$  as  $y_0^* \rightarrow +\infty$ . Let us assume that if  $\omega = \omega_{y_0^*}^\varepsilon$ , then  $\omega/2\pi$  is a Diophantine number of type  $(\gamma, \nu)$  as defined in (26). Then, we take  $\rho_0 = \bar{\rho}$ ,  $R_0 = \bar{\rho}$ ,  $c_1 = \frac{5\sqrt{2}}{3}$ ,  $c_2 = \frac{10\sqrt{2}}{3}$ ,  $c_3 = \frac{25}{24}$  and  $c = \varepsilon \bar{C}$ , and apply Theorem B to the map  $F = F(\phi, I)$  in (31) corresponding to the selected couple  $(\varepsilon, y_0^*)$  in terms of these parameters, at the action  $I = I_0^*$ . Consequently, all the conditions of Theorem B are met and we conclude that there exists  $0 < \varepsilon^* \leq \varepsilon_{\mathcal{P}}^*$  such that  $F$  possesses an analytic invariant curve of frequency  $\omega = \omega_{y_0^*}^\varepsilon$ , provided that the selected  $\varepsilon$  is smaller than  $\varepsilon^*$ . Moreover,  $\varepsilon^*$  does only depend on  $\bar{\rho}$ ,  $\bar{\rho}$ ,  $\bar{C}$ ,  $\gamma$ , and  $\nu$ , but not on the selected value of  $y_0^* > 5$  nor on the size of the frequency  $\omega_{y_0^*}^\varepsilon$ . This invariant curve of  $F$  is close to the unperturbed circle  $\mathbb{T} \times \{I_0^*\}$  and it admits an analytic parameterization  $\varphi^* = (\varphi_{y_0^*}^\varepsilon)^*$ , of the form  $\varphi^*(\theta) = (\theta + \varphi_\phi^*(\theta), \varphi_I^*(\theta))$ ,  $\theta \in \mathbb{T}$ , that solves the corresponding invariance equation (23). The parameterization  $\varphi^*$  also satisfies the estimates given in (28), for some constant  $c^*$  that only depends on  $\nu$ , where  $\delta_0 = \bar{\rho}/12$ . In particular, this means that if we define  $K^* = 2c^*\bar{C}\gamma^{-2}\delta_0^{-2\nu-1}$ , we have the bounds:

$$\left\| \varphi_\phi^* \right\|_{\bar{\rho}/2} \leq K^* \varepsilon, \quad \left\| \varphi_I^* - I_0^* \right\|_{\bar{\rho}/2} \leq K^* \varepsilon, \quad \left\| (\varphi_s^*)' \right\|_{\bar{\rho}/2} \leq K^* \varepsilon, \quad s \in \{\phi, I\}.$$

Going back to the original variables  $(t_0, x_0 = 0, y_0)$  of the impact map  $\mathcal{P}_\varepsilon$ , we observe that by considering the above discussed relations  $t_0 = \phi$  and  $y_0 = \sqrt{-\sqrt{2}y_0^*I}$  (see (29) and (30)), we obtain the following analytical parametrization for the corresponding  $\omega_{y_0^*}^\varepsilon$ -quasi-periodic solution of  $\mathcal{P}_\varepsilon$  in  $\Sigma^+$  (see (5)), that we denote as  $\eta^+ = (\eta_{y_0^*}^\varepsilon)^+$  (in the following we will often

omit the dependence on  $(\varepsilon, y_0^*)$  since in this discussion we are dealing with a fixed pair of these values):

$$\eta^+(\theta) = (\theta + \eta_{t_0}^+(\theta), 0, \eta_{y_0}^+(\theta)) = \left( \theta + \varphi_\phi^*(\theta), 0, \sqrt{-\sqrt{2} y_0^* \varphi_I^*(\theta)} \right), \quad \theta \in \mathbb{T},$$

on the understanding that here we are treating  $t_0$  as a variable defined modulus  $2\pi$ , so in fact we are treating the plane  $\Sigma^+$  as a cylinder, that we denote by  $\tilde{\Sigma}^+$ . This parametrization verifies  $\mathcal{P}_\varepsilon(\eta^+(\theta)) = \eta^+(\theta + \omega)$ , where  $\omega = \omega_{y_0^*}^\varepsilon$ . Using that  $y_0^* = -\sqrt{2} I_0^*$ , we observe that

$$\begin{aligned} \eta_{y_0}^+(\theta) - y_0^* &= \frac{-\sqrt{2} y_0^* (\varphi_I^*(\theta) - I_0^*)}{\sqrt{-\sqrt{2} y_0^* \varphi_I^*(\theta) + y_0^*}}, \\ (\eta_{y_0}^+)'(\theta) &= \frac{-\sqrt{2} y_0^* (\varphi_I^*)'(\theta)}{\sqrt{-\sqrt{2} y_0^* \varphi_I^*(\theta) + y_0^*}} + \frac{(y_0^*)^2 (\varphi_I^*(\theta) - I_0^*) (\varphi_I^*)'(\theta)}{(\sqrt{-\sqrt{2} y_0^* \varphi_I^*(\theta) + y_0^*})^2 \sqrt{-\sqrt{2} y_0^* \varphi_I^*(\theta)}}. \end{aligned}$$

These expression allows to obtain the following estimates (we note that although  $\eta^+$  is an analytic function, for stability purposes bounds for “real” values of  $\theta$  suffice):

$$|\eta_{t_0}^+(\theta)| \leq K^* \varepsilon, \quad |\eta_{y_0}^+(\theta) - y_0^*| \leq \sqrt{2} K^* \varepsilon, \quad |(\eta_{t_0}^+)'(\theta)| \leq K^* \varepsilon, \quad |(\eta_{y_0}^+)'(\theta)| \leq 2 \sqrt{2} K^* \varepsilon, \quad \theta \in \mathbb{T}.$$

To get the last bound of this list perhaps we have to make  $\varepsilon^*$  slightly smaller or to slightly increase  $K^*$ , but in any case the obtained value of  $\varepsilon^*$  or  $K^*$  depends on the same parameters as the former one. Consequently, the closed curve  $\mathcal{T}^+ = (\mathcal{T}_{y_0^*}^\varepsilon)^+ = \eta^+(\theta) \subset \tilde{\Sigma}^+$  is invariant by the action of the impact map  $\mathcal{P}_\varepsilon$  and it remains close to the circle  $\mathbb{T} \times \{y_0^*\}$ , to which it is homotopic. For further uses, we also introduce the corresponding invariant curve in the cylinder  $\tilde{\Sigma}^-$  defined by the action of  $\mathcal{P}_\varepsilon^+$  on  $\mathcal{T}^+$ , i.e.,  $\mathcal{T}^- = (\mathcal{T}_{y_0^*}^\varepsilon)^- = \mathcal{P}_\varepsilon^+(\mathcal{T}^+) \subset \tilde{\Sigma}^-$ . Of course, we have  $\mathcal{P}_\varepsilon^-(\mathcal{T}^-) = \mathcal{T}^+$ .

By integration of  $\mathcal{T}^+$  by the flow of (4), this curve gives rise to an invariant two-dimensional torus  $\mathcal{S} = \mathcal{S}_{y_0^*}^\varepsilon$  of (4), on the understanding that we are still treating  $t_0$  as a variable defined modulus  $2\pi$ . The torus  $\mathcal{S}$  is continuous, piecewise analytic and  $\varepsilon$ -close to the Cartesian product of the corresponding  $y_0^*$ -invariant curve of the unperturbed system  $\varepsilon = 0$  (as displayed in Fig. 1) by the variable  $t \in \mathbb{T}$ , being homotopic to  $(t, x, \dot{x}) \in \mathbb{T} \times \{0\} \times \{y_0\}$ . Indeed,  $\mathcal{S}$  is not smooth in both the upper curve  $\mathcal{T}^+$  and the lower curve  $\mathcal{T}^-$ , since integration by (4) means to follow the flow  $\Phi_r^\tau$  of (6) from  $\mathcal{T}^+$  to  $\mathcal{T}^-$  and the flow  $\Phi_l^\tau$  of (7) from  $\mathcal{T}^-$  to  $\mathcal{T}^+$ . Parametrizations of the right and left components of  $\mathcal{S}$  are given by  $\sigma^+ = (\sigma_{y_0^*}^\varepsilon)^+$  and  $\sigma^- = (\sigma_{y_0^*}^\varepsilon)^-$ , respectively, which can be written as:

$$\sigma^+(\theta, \tau) = \Phi_r^\tau(\eta^+(\theta)), \quad \theta \in \mathbb{T}, \quad 0 \leq \tau \leq \tau^+(\tilde{\eta}^+(\theta); \varepsilon),$$

and

$$\sigma^-(\theta, \tau) = \Phi_l^\tau(\eta^-(\theta)), \quad \theta \in \mathbb{T}, \quad 0 \leq \tau \leq \tau^-(\tilde{\eta}^-(\theta); \varepsilon),$$

where  $\tau^+$  and  $\tau^-$  are the impact times introduced in (8) and (10), and we are denoting  $\eta^- = \mathcal{P}_\varepsilon^+(\eta^+)$  and  $\tilde{\eta}^+$  and  $\tilde{\eta}^-$  as the functions defined by the first and third components  $(t, y)$  of the parameterizations  $\eta^+$  and  $\eta^-$ , respectively (without the zero of the second component).

Since (4) is  $2\pi$  periodic in  $t$ , to make the variable  $t_0$  a true dynamical variable of the system we should lift it from  $\mathbb{T}$  to  $\mathbb{R}$ . Then, the torus  $\mathcal{S} = \mathcal{S}_{y_0^*}^\varepsilon$  becomes an infinite cylinder of  $(t, x, \dot{x}) \in \mathbb{R}^3$ , which is perpetually  $\varepsilon$ -close to the Cartesian product of one of the invariant curves of Fig. 1 by  $t \in \mathbb{R}$ . This means that it perpetually confines the trajectories whose initial

conditions belongs inside this cylinder, so that they remain bounded for al time. Furthermore, the amplitude in the  $(x, \dot{x})$ -plane of these cylinders goes to infinite as  $y_0^* \rightarrow +\infty$ .

The final part of the proof of Theorem A is a mere observation. Let us consider  $\omega_0 \in \mathbb{R}$  any given value such that  $\omega_0/2\pi$  is a Diophantine number according to definition (26). E.g., we can select  $\omega_0$  as  $2\pi$  times the Golden Mean  $(\sqrt{5} - 1)/2$ . Associated to  $\omega_0$ , we have a particular couple  $(\gamma, \nu)$  in (26), that we set fixed up to the end of this section. We then introduce  $\omega_k = \omega + 2\pi k$ , where  $k \in \mathbb{N}$ . On the one side,  $\omega_k \rightarrow +\infty$  as  $k \rightarrow +\infty$ . Consequently, since we have that  $\omega_{y_0^*}^\varepsilon = \frac{4}{1-a_0^2\varepsilon^2}y_0^*$ , then, for any given  $0 < \varepsilon < 1/|a_0|$ , we can define an infinite sequence of values of  $y_0^* = y_0^*(\varepsilon, k) > 5$  for which we have  $\omega_{y_0^*(\varepsilon, k)}^\varepsilon = \omega_k$ , for any  $k \geq k_0$ . The value of  $k_0$  is independent of  $\varepsilon$  and  $y_0^*(\varepsilon, k) \rightarrow +\infty$  as  $k \rightarrow +\infty$ . On the other side,  $\omega_k/2\pi$  is also a Diophantine number of type  $(\gamma, \nu)$ , for any  $k \in \mathbb{N}$ . We then select as specific value of  $\varepsilon^*$  in the statement of Theorem A the one provided by construction above when applied to any given  $y_0^* > 5$  for which  $\omega = \omega_{y_0^*}^\varepsilon$  is such that  $\omega/2\pi$  is a Diophantine number of type  $(\gamma, \nu)$ . Since  $\varepsilon^*$  depends on  $(\gamma, \nu)$ , but not on the size of the frequency  $\omega_{y_0^*}^\varepsilon$ , we conclude that for each  $0 < \varepsilon < \varepsilon^*$  we can construct a sequence of two-dimensional tori  $\{S_{y_0^*(\varepsilon, k)}^\varepsilon\}_{k \geq k_0}$  whose projection onto the plane  $(x, \dot{x})$  confine the trajectories surrounded by each of these tori. Since the amplitude of these tori go to infinity with  $k$ , we get the desired perpetual stability for this specific value of  $\varepsilon$ .

### 3. A PARAMETERIZATION RESULT FOR INVARIANT CURVES OF A TWIST MAP OF THE ANNULUS: PROOF OF THEOREM 2

Firstly, we introduce some notations and basic results to be used throughout the proof of Theorem B. Let  $g : \mathbb{R} \rightarrow \mathbb{R}$  be a real analytic and  $2\pi$ -periodic function that we assume that can be analytically extended to  $\Delta(\rho)$  (see (15)), for some  $\rho > 0$ , and denote by  $\tilde{g} = g - \langle g \rangle$ . Then, for any  $0 < \delta \leq \rho$  we have

$$|\langle g \rangle| \leq \|g\|_\rho, \quad \|\tilde{g}\|_\rho \leq 2\|g\|_\rho, \quad \|g'\|_{\rho-\delta} \leq \frac{\|g\|_\rho}{\delta}.$$

where the last inequality is a consequence of the application of the classical Cauchy estimates to  $g$ . Since  $\omega \in \mathbb{R}$  is set to be a fixed number, we introduce the notation  $f_+(\cdot) = f(\cdot + \omega)$  and we define the operator  $\mathcal{L}_\omega(f) = f_+ - f$ . Let us denote as  $g(\theta) = \sum_{k \in \mathbb{Z}} \hat{g}_k e^{ik\theta}$  the Fourier expansion of the  $2\pi$ -periodic and analytic function  $g$  above. If  $\omega/2\pi$  is an irrational number and  $\langle g \rangle = \hat{g}_0 = 0$ , then there is a unique formal  $2\pi$ -periodic solution  $f = \mathcal{L}_\omega^{-1}(g)$  of the linear equation  $\mathcal{L}_\omega(f) = g$ , provided that the normalization condition  $\langle f \rangle = 0$  is assumed on  $f$ . This solution is explicitly given by the expansion:

$$\mathcal{L}_\omega^{-1}(g) = \sum_{k \in \mathbb{Z} \setminus \{0\}} \frac{\hat{g}_k}{ik\omega} e^{ik\theta}.$$

In fact, all the formal solutions of  $\mathcal{L}_\omega(f) = g$  are of the form  $f = c + \mathcal{L}_\omega^{-1}(g)$ , for any constant  $c$ . If  $\omega/2\pi$  is a Diophantine number of type  $(\gamma, \nu)$  (see (26)), then so-called Rüssmann's estimates (see e.g. [6]) show that  $\mathcal{L}_\omega^{-1}(g)$  is also an analytic function and that there is a constant  $C_R = C_R(\nu)$ , for which  $\|\mathcal{L}_\omega^{-1}(g)\|_{\rho-\delta} \leq C_R(\gamma\delta^\nu)^{-1}\|g\|_\rho$ . Finally, we recall that being  $F$  an exact symplectic map of  $\mathbb{T} \times \mathbb{R}$ , then it is also a symplectic map and consequently the following equalities are satisfied

$$(32) \quad F^*(d\phi \wedge dI) = d\phi \wedge dI \quad \text{and} \quad (DF(p))^\top J(DF(p)) = J,$$

where  $DF(p)$  denotes the differential matrix of  $F$  at the point  $p$  and  $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$  is the matrix representation of the two form  $d\phi \wedge dI$ .

The proof of Theorem B is performed by means of the iterative application of the lemma below. It is worthy mentioning that  $\varphi$  and the related objects depend on  $\theta$ , but this dependence is going to be omitted in most of the expressions.

**Iterative Lemma.** *With the same hypotheses and notations on the statement of Theorem B, let us consider a parametrization of a curve  $\varphi : \mathbb{T} \rightarrow \mathbb{T} \times \mathbb{R}$  of the form  $\varphi(\theta) = (\theta + \varphi_\phi(\theta), \varphi_I(\theta))$ , for which the following holds. We suppose that  $\varphi_\phi$  and  $\varphi_I$  are  $2\pi$ -periodic functions that can be analytically extended to  $\Delta(\rho)$  (see (15)), for some  $0 < \rho \leq \rho_0 < 1$ , and verify:*

$$(33) \quad \|\varphi_\phi\|_\rho \leq \rho_0 - \rho, \quad \|\varphi_I - I_0^*\|_\rho \leq R_0, \quad \|\varphi'_\phi\|_\rho \leq 1/4 \quad \text{and} \quad \|\varphi'_I\|_\rho \leq 1/4.$$

We denote by  $e$  the invariance error (24) associated to  $\varphi$  and  $F$  and we define the scalar functions

$$(34) \quad \Omega = (\varphi')^\top \varphi', \quad A = \Omega^{-1} \Omega_+^{-1} (\varphi'_+)^{\top} DF(\varphi) J \varphi'.$$

We assume that  $|\langle A \rangle| \geq c_1/2$  and that there is a constant  $0 < \mu \leq 1$  for which  $\|e\|_\rho \leq \mu$  and  $\|e'\|_\rho \leq \mu$ . Then, there is a constant  $c^* \geq 1$ , that only depends on  $1/c_1, c_2, c_3$  and  $v$ , and that can be defined in such a way it is an increasing function of all these quantities, for which the following holds. We take  $0 < \delta < \rho/3$  for which

$$(35) \quad \|\varphi_I - I_0^*\|_\rho \leq R_0 - \frac{c^* \mu}{\gamma^2 \delta^{2v}}, \quad \|\varphi'_j\|_\rho \leq \frac{1}{4} - \frac{c^* \mu}{\gamma^2 \delta^{2v+1}},$$

for  $j \in \{\phi, I\}$ . Then, we can define a new parametrization  $\varphi^{(1)} = \varphi + \Delta\varphi$ , where  $\Delta\varphi : \mathbb{T} \rightarrow \mathbb{T} \times \mathbb{R}$  is  $2\pi$  periodic and analytic in  $\Delta(\rho^{(1)})$ , with  $\rho^{(1)} = \rho - 3\delta$ , which verifies:

$$\|\Delta\varphi\|_{\rho^{(1)}} \leq \frac{c^* \mu}{\gamma^2 \delta^{2v}}, \quad \text{and} \quad \|(\Delta\varphi)'\|_{\rho^{(1)}} \leq \frac{c^* \mu}{\gamma^2 \delta^{2v+1}}.$$

In addition, if we define the invariance error  $e^{(1)}$  and the function  $A^{(1)}$  analogously as we have defined  $e$  and  $A$  for  $\varphi$ , we also have:

$$\|A^{(1)} - A\|_{\rho^{(1)}} \leq \frac{c^* \mu}{\gamma^2 \delta^{2v+1}}, \quad \|e^{(1)}\|_{\rho^{(1)}} \leq \mu^{(1)}, \quad \text{and} \quad \|(e^{(1)})'\|_{\rho^{(1)}} \leq \mu^{(1)},$$

where  $\mu^{(1)} = \frac{c^* \mu^2}{\gamma^4 \delta^{4v+1}}$ .

Last formula means that the size of the new invariance error  $e^{(1)}$  is (almost) quadratic with respect to the size of  $e$ . Hence, we may expect (under suitable assumptions) that the size of the error goes iteratively to zero very fast with the step.

*Proof.* Using Newton's method as a benchmark tool, a natural way to define the correction  $\Delta\varphi$  for the initial parametrization  $\varphi$  is to try to obtain a new error  $e^{(1)}$  of quadratic size with respect to  $e$ . Explicitly:

$$e^{(1)} = F(\varphi^{(1)}) - \varphi_+^{(1)} = E^{(1)} + \mathcal{R}(\varphi, \Delta\varphi),$$

where

$$E^{(1)} = e + DF(\varphi)\Delta\varphi - \Delta\varphi_+, \quad \mathcal{R}(\varphi, \Delta\varphi) = F(\varphi + \Delta\varphi) - F(\varphi) - DF(\varphi)\Delta\varphi.$$

Hence,  $E^{(1)}$  stands for the linear approximation of  $e^{(1)}$  around  $\varphi$  and  $\mathcal{R}(\varphi, \Delta\varphi)$  is the corresponding quadratic reminder. In the aim of Newton's method, our ideal target is to try to set  $E^{(1)} = 0$ . The functional structure of the problem makes this objective unrealistic, so we



consider the less ambitious goal of obtaining an expression for  $E^{(1)}$  of quadratic size in terms of the size of  $e$ . This is enough for our purposes since it allows to avoid the nocive effect that the small divisors associated to the problem have on the convergence of the iterative application of this lemma. A construction like the one used to define  $\Delta\varphi$  is usually called a quasi-Newton method. Since hypotheses on the statement of the lemma ensure that  $\{\varphi', J\varphi'\}$  define a basis of  $\mathbb{R}^2$  at any point of  $\theta$ , we look for a correction  $\Delta\varphi$  of the form:

$$(36) \quad \Delta\varphi = a \varphi' + b \Omega^{-1} J\varphi',$$

where  $a$  and  $b$  are small scalar and  $2\pi$ -periodic functions which we select as the solutions of following (small-divisors) difference equations

$$(37) \quad \mathcal{L}_\omega(b) = \langle (\varphi'_+)^{\top} J e \rangle - (\varphi'_+)^{\top} J e \quad \text{and} \quad \mathcal{L}_\omega(a) = Ab + \Omega_+^{-1} (\varphi'_+)^{\top} e,$$

with the normalization condition  $\langle a \rangle = 0$ , since selecting a difference value for  $\langle a \rangle = 0$  only means a “change of origin” on the curve. To avoid increasing the length of the proof, we omit details on the analytic and geometric motivations of selecting such specific expressions for  $a$  and  $b$ , that are discussed in the references on parametric KAM Theory quoted in the paper. The right hand side of the equation for  $b$  has zero average, so  $b$  takes the form  $b = \langle b \rangle + \tilde{b}$ , where

$$\tilde{b} = \mathcal{L}_\omega^{-1} (\langle (\varphi'_+)^{\top} J e \rangle - (\varphi'_+)^{\top} J e),$$

and  $\langle b \rangle$  is selected in such a way the right hand side of the equation for  $a$  has zero average. Using that by hypothesis  $\langle A \rangle \neq 0$ , this means that

$$\langle b \rangle = -\langle A \rangle^{-1} \left( \langle A \tilde{b} \rangle + \langle \Omega_+^{-1} (\varphi'_+)^{\top} e \rangle \right).$$

Finally, we should select

$$a = \mathcal{L}_\omega^{-1} (Ab + \Omega_+^{-1} (\varphi'_+)^{\top} e).$$

In order to establish that taking  $\Delta\varphi$  defined in such a way means a quadratic size for  $E^{(1)}$  in terms of  $e$ , key computations are to show that the vectors  $DF(\varphi)\varphi'$  and  $DF(\varphi)\Omega^{-1}J\varphi'$  can be written in the following way

$$(38) \quad DF(\varphi)\varphi' = \varphi'_+ + e', \quad DF(\varphi)\Omega^{-1}J\varphi' = A \varphi'_+ + (1 + B) \Omega_+^{-1} J\varphi'_+,$$

where  $A$  is defined in the statement and  $B$  is the scalar function  $B = -\Omega^{-1}(DF(\varphi)^{-1}e')^{\top}\varphi'$ . Both expressions are straightforward. The first one follows by taking the derivative on the definition of the invariance error  $e$  in (24) and the second one follows by multiplying it by  $(\varphi')^{\top}$  and  $(\varphi')^{\top}J$ , and using that  $DF$  is a symplectic matrix (see (32)) at any point (we also recall that  $J^{\top} = J^{-1} = -J$  and that  $u^{\top}Ju = 0$  for any vector  $u$ ). The fact that  $e$  is an analytic periodic function means that the size of  $e'$  (and so the size of  $B$ ) is “comparable” to the size of  $e$  modulus the application of appropriate Cauchy estimates. In the literature, expressions in (38) are usually discussed in terms of the concept of quasi-reducibility of the the linearized system around a quasi-torus of a symplectic map. We refer the interested reader to the references quoted in the introduction for more details. Using expression in (38) and equation (36) it follows that:

$$E^{(1)} = a e' + B b \Omega_+^{-1} J\varphi'_+ + \zeta^{(1)},$$

where

$$\zeta^{(1)} = e + (a - a_+ + Ab)\varphi'_+ + (b - b_+)\Omega_+^{-1}J\varphi'_+ = e + (Ab - \mathcal{L}_\omega(a))\varphi'_+ - \mathcal{L}_\omega(b)\Omega_+^{-1}J\varphi'_+.$$

By using difference equations (37) for  $a$  and  $b$  and the fact that in the basis  $\{\varphi'_+, \Omega_+^{-1}J\varphi'_+\}$  the invariance error  $e$  is written as

$$e = (\Omega_+^{-1}(\varphi'_+)^{\top} e) \varphi'_+ - ((\varphi'_+)^{\top} J e) \Omega_+^{-1} J \varphi'_+,$$

we get the expression  $\zeta^{(1)} = -\langle (\varphi'_+)^{\top} J e \rangle \Omega_+^{-1} J \varphi'_+$ . Since we easily show that the size of  $e'$ ,  $B$ ,  $a$ , and  $b$  are all comparable to the size of  $e$ , the quadratic behaviour in  $e$  of  $E^{(1)}$  is equivalent to the quadratic behaviour in  $e$  of  $\zeta^{(1)}$  above. The fact that  $\zeta^{(1)}$  has quadratic size is not obvious, but it is a consequence of the exact symplectic properties of the map  $F$ . This is the only point of the proof in which the exactness of  $F$  is considered. We remind that being  $F$  exact symplectic with respect to the 1-form  $\text{Id}\phi$  means that

$$F^*(\text{Id}\phi) = \text{Id}\phi + d(V(\phi, I)),$$

for some function  $V : \mathbb{T} \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $2\pi$ -periodic in  $\phi$ . By performing the pull-back by  $\varphi : \mathbb{T} \rightarrow \mathbb{T} \times \mathbb{R}$ , we obtain:

$$(F \circ \varphi)^*(\text{Id}\phi) = \varphi^*(\text{Id}\phi) + d(V \circ \varphi).$$

By writting  $\varphi = \varphi(\theta)$  and  $F = (F_{\phi}, F_I)$ , and then considering the coordinate representation of this one-form in terms of the basis  $d\theta$ , we obtain the equality:

$$(F_I(\varphi))(F_{\phi}(\varphi))' = (1 + \varphi'_{\phi})\varphi_I + (V(\varphi))'.$$

Then, writting  $e = (e_{\phi}, e_I)$  and using that  $F_{\phi}(\varphi) = \theta + \omega + \varphi_{\phi,+} + e_{\phi}$ ,  $F_I(\varphi) = \varphi_{I,+} + e_I$  and that  $(\varphi'_+)^{\top} J e = -e_{\phi} \varphi'_{I,+} + e_I(1 + \varphi_{\phi,+})$ , it follows:

$$(\varphi'_+)^{\top} J e = -\mathcal{L}_{\omega}((1 + \varphi'_{\phi})\varphi_I) + (V(\varphi))' - e'_{\phi} e_I - (e_{\phi} \varphi_{I,+})'.$$

Since  $(1 + \varphi'_{\phi})\varphi_I$ ,  $e'_{\phi} e_I$ , and  $e_{\phi} \varphi_{I,+}$  are  $2\pi$ -periodic in  $\theta$ , we conclude that  $\langle (\varphi'_+)^{\top} J e \rangle = -\langle e'_{\phi} e_I \rangle$ . Consequently:

$$E^{(1)} = a e' + B b \Omega_+^{-1} J \varphi'_+ + \langle e'_{\phi} e_I \rangle \Omega_+^{-1} J \varphi'_+.$$

This last computation ends the formal part of the proof of the lemma. Now, we have to perform the rigorous bounds of all the objects involved in above computations. In particular, bounds below show that the previous formal computations are well-defined if the hypotheses of the lemma hold. Our main purpose is to construct a constant  $c^*$  for which the result holds. To achieve this aim, during the next sequence of bounds the value of  $c^*$  is redefined recursively to meet a finite number of conditions. At the end of this process, the last value of  $c^*$  obtained is that of the statement of the lemma.

From bounds on  $\varphi'_{\phi}$  and  $\varphi'_I$  in (33), we notice that  $\Omega = (1 + \varphi'_{\phi})^2 + (\varphi'_I)^2$  verifies  $\|\Omega - 1\|_{\rho} \leq 5/8$ . Then,  $3/8 \leq |\Omega(\theta)| \leq 13/8$ , for any  $\theta \in \Delta(\rho)$  and so  $\|\Omega\|_{\rho} \leq 2$  and  $\|\Omega^{-1}\|_{\rho} \leq 3$ . Since hypotheses in (33) ensure that  $DF(\varphi(\theta))$  is well-defined for all  $\theta \in \Delta(\rho)$ , we conclude that  $A$  is well defined in (34) and verifies an estimate of the form  $\|A\|_{\rho} \leq c^*$  for some  $c^*$  that, up to now, only depends on  $c_2$  (since we are assuming  $c < 1$ ).

We then use Rüssmann's estimates to control the size of the solutions for  $a$  and  $b$  of equations in (37). Firstly we have:

$$(39) \quad \|\tilde{b}\|_{\rho-\delta} \leq c_R \frac{\| \langle (\varphi'_+)^{\top} J e \rangle - (\varphi'_+)^{\top} J e \|_{\rho}}{\gamma \delta^{\nu}} \leq \frac{c^* \mu}{\gamma \delta^{\nu}},$$

where  $c_R = c_R(\nu)$  is the constant provided by Rüssmann's result,  $\delta$  is the one in the statement of the lemma and  $c^*$  is re-defined to meet both  $\|A\|_{\rho} \leq c^*$  and (39), and in such a way it turns out to be an increasing function of the quantities in which  $c^*$  depends. To achieve this last property, we observe that each time we have to re-define  $c^*$  in order to meet a new bound,

the new expression of  $c^*$  turns out to be a polynomial formula involving  $1/c_1, c_2, c_3, \nu$  and the preceding expression of  $c^*$ . Therefore, the new expression of  $c^*$  is obtained by taking the maximum between the former expression of  $c^*$  and this new polynomial expression. By proceeding in analogous way as we have done to get these bounds, we obtain (by recursively re-defining  $c^*$ ):

$$\begin{aligned} |\langle b \rangle| &\leq \frac{c^* \mu}{\gamma \delta^\nu}, \quad \|b\|_{\rho-\delta} \leq \frac{c^* \mu}{\gamma \delta^\nu}, \quad \|a\|_{\rho-2\delta} \leq \frac{c^* \mu}{\gamma^2 \delta^{2\nu}}, \quad \|\Delta \varphi\|_{\rho-2\delta} \leq \frac{c^* \mu}{\gamma^2 \delta^{2\nu}}, \\ \|(\Delta \varphi)'\|_{\rho-3\delta} &\leq \frac{c^* \mu}{\gamma^2 \delta^{2\nu+1}}, \quad \|B\|_\rho \leq c^* \mu, \quad \|E^{(1)}\|_{\rho-2\delta} \leq \frac{c^* \mu^2}{\gamma^2 \delta^{2\nu}}, \end{aligned}$$

where we are using standard Cauchy estimates to bound each  $\theta$ -derivatives and, to bound  $B$ , that the symplectic character of  $F$  implies that  $(DF)^{-1} = -J(DF)^\top J$ .

To control the remaining aspects of the proof, we need to ensure that the composition of  $F$  and its derivatives with the new parametrization  $\varphi^{(1)} = \varphi + \Delta \varphi$  are well defined. By writing  $\varphi^{(1)} = (\theta + \varphi_\phi^{(1)}, \varphi_I^{(1)})$  and  $\Delta \varphi = (\Delta \varphi_\phi, \Delta \varphi_I)$ , and using hypotheses in (33) and (35), we conclude that for any  $\theta \in \Delta(\rho - 2\delta)$  we have:

$$\begin{aligned} |\operatorname{Im}(\varphi_\phi^{(1)})| &\leq |\operatorname{Im}(\theta)| + \|\varphi_\phi\|_\rho + \|\Delta \varphi_\phi\|_{\rho-2\delta} \leq \rho_0 - \delta < \rho_0, \\ |\varphi_I^{(1)} - I_0^*| &\leq \|\varphi_I - I_0^*\|_\rho + \|\Delta \varphi_I\|_{\rho-2\delta} \leq R_0. \end{aligned}$$

We note that although hypotheses in (35) are formulated in terms of the final value of  $c^*$ , they are also valid in terms of the value of  $c^*$  defined so far. In particular, bounds above allow us to define the new error  $e^{(1)}$ . To obtain an estimate for the size of  $e^{(1)}$ , it only remains to bound the remainder of the Taylor expansion  $\mathcal{R}(\varphi, \Delta \varphi)$ , that we express as  $\mathcal{R} = (\mathcal{R}_\phi, \mathcal{R}_I)$ . Then, to estimate  $\mathcal{R}_j(\varphi, \Delta \varphi)(\theta)$ , for any fixed  $\theta \in \Delta(\rho - 2\delta)$  and with  $j \in \{\phi, I\}$ , we introduce the assistant function  $g_j(s) = F_j(\varphi + s\Delta \varphi)$ . Since  $g_j'(s) = DF_j(\varphi + s\Delta \varphi) \Delta \varphi$ , integrating  $g_j$  by parts we observe that

$$\mathcal{R}_j(\varphi, \Delta \varphi) = g_j(1) - g_j(0) - g_j'(0) = - \int_0^1 (s-1) g_j''(s) ds.$$

Thus,  $\|\mathcal{R}_j\|_{\rho-2\delta} \leq \sup_{s \in [0,1]} \{\|g_j''(s)\|_{\rho-2\delta}\}$ . But, since

$$g_j''(s) = \partial_\phi^2 F_j(\varphi + s\Delta \varphi) (\Delta \varphi_\phi)^2 + 2\partial_I \partial_\phi F_j(\varphi + s\Delta \varphi) (\Delta \varphi_\phi) (\Delta \varphi_I) + \partial_I^2 F_j(\varphi + s\Delta \varphi) (\Delta \varphi_I)^2,$$

it follows that

$$\|\mathcal{R}\|_{\rho-2\delta} \leq \frac{c^* \mu^2}{\gamma^4 \delta^{4\nu}}, \quad \|e^{(1)}\|_{\rho-2\delta} \leq \frac{c^* \mu^2}{\gamma^4 \delta^{4\nu}}, \quad \|e^{(1)}\|_{\rho-3\delta} \leq \frac{c^* \mu^2}{\gamma^4 \delta^{4\nu+1}},$$

and therefore, the bounds for  $e^{(1)}$  and  $(e^{(1)})'$  are met in terms of the expression for  $\rho^{(1)}$  and  $\mu^{(1)}$  defined on the statement. To get estimates above on  $g_j''(s)$ , a crucial aspect is that  $\varphi + s\Delta \varphi \in \Delta(\rho_0)$  for any  $s \in [0,1]$ , so that the composition of  $\varphi + s\Delta \varphi$  with  $F$  and its derivatives is well defined for any  $s \in [0,1]$ . It only remains to bound  $A^{(1)} - A$ , where

$$A^{(1)} = (\Omega^{(1)})^{-1} (\Omega_+^{(1)})^{-1} ((\varphi_+^{(1)})')^\top DF(\varphi^{(1)}) J(\varphi^{(1)})'.$$

We notice that hypotheses in (35) on  $\varphi'_\phi$  and  $\varphi'_I$  and bound above on  $(\Delta \varphi)'$  guarantee that  $\Omega^{(1)}$  is well defined in  $\Delta(\rho - 3\delta)$  and verifies the same bounds as  $\Omega$ . To get the bound for  $A^{(1)} - A$ ,

we consider the telescopic sum

$$\begin{aligned} A^{(1)} - A = & ((\Omega^{(1)})^{-1} - \Omega^{-1})(\Omega_+^{(1)})^{-1}((\varphi_+^{(1)})')^\top DF(\varphi^{(1)})J(\varphi^{(1)})' \\ & + \Omega^{-1}((\Omega_+^{(1)})^{-1} - \Omega_+^{-1})(\varphi_+^{(1)})'^\top DF(\varphi^{(1)})J(\varphi^{(1)})' \\ & + \Omega^{-1}\Omega_+^{-1}((\varphi_+^{(1)})' - \varphi_+'^\top)^\top DF(\varphi^{(1)})J(\varphi^{(1)})' \\ & + \Omega^{-1}\Omega_+^{-1}(\varphi_+'^\top)^\top (DF(\varphi^{(1)}) - DF(\varphi))J(\varphi^{(1)})' \\ & + \Omega^{-1}\Omega_+^{-1}(\varphi_+'^\top)^\top DF(\varphi)J(\varphi^{(1)} - \varphi)'. \end{aligned}$$

Then, we estimate every difference in the equality above leading to a bound of the form

$$\|A^{(1)} - A\|_{\rho-3\delta} \leq \frac{c^*\mu}{\gamma^2\delta^{2\nu+1}}.$$

The only expressions involved in this telescopic sum for which the corresponding bounds are not completely straightforward are  $(\Omega^{(1)})^{-1} - \Omega^{-1}$  and  $DF(\varphi^{(1)}) - DF(\varphi)$ . For the first one we write:

$$(\Omega^{(1)})^{-1} - \Omega^{-1} = \Omega^{-1}(\Omega^{(1)})^{-1}(\Omega - \Omega^{(1)}),$$

and we estimate the size of  $\Omega^{(1)} - \Omega$  in terms of the size of  $(\varphi^{(1)})' - \varphi' = (\Delta\varphi)'$  by also expressing  $\Omega^{(1)} - \Omega$  in terms of a telescopic sum analogous to the one above. Indeed, we get the estimate  $\|(\Omega^{(1)})^{-1} - \Omega^{-1}\|_{\rho-3\delta} \leq \frac{c^*\mu}{\gamma^2\delta^{2\nu+1}}$ . To bound  $DF(\varphi^{(1)}) - DF(\varphi)$ , we proceed analogously as we have done to bound  $\mathcal{R}(\varphi, \Delta\varphi)$ . For any fixed  $\theta \in \Delta(\rho - 2\delta)$  and for any couple  $j, k \in \{\phi, I\}$ , we consider the auxiliary function  $h_{j,k}(s) = \partial_j F_k(\varphi + s\Delta\varphi)$ , for  $s \in [0, 1]$ . By the Fundamental Theorem of Calculus, we see that

$$\partial_j F_k(\varphi^{(1)}) - \partial_j F_k(\varphi) = \int_0^1 \nabla(\partial_j F_k)(\varphi + s\Delta\varphi) ds \cdot \Delta\varphi,$$

where  $\nabla$  refers to the gradient operator with respect to  $(\phi, I)$ . Since all the involved compositions are well-defined for any  $\theta \in \Delta(\rho - 2\delta)$  and  $s \in [0, 1]$ , we get an estimate of the form  $\|\partial_j F_k(\varphi^{(1)}) - \partial_j F_k(\varphi)\|_{\rho-2\delta} \leq \frac{c^*\mu}{\gamma^2\delta^{2\nu}}$  for any entry of  $DF(\varphi^{(1)}) - DF(\varphi)$ .  $\square$

*Proof of Theorem B.* Our goal is to prove this result by the iterative application of Lemma 3. Hence, we start by considering  $c^* = c^*(1/c_1, c_2, c_3, \nu) \geq 1$  the value provided by this iterative lemma, that we set fixed from now on along the proof of the theorem.

**First step of the iterative scheme.** To perform the first application (for  $n = 0$ ) of Lemma 3, we select the parametrization  $\varphi(\theta) = \varphi^{(0)}(\theta) = (\theta, I_0^*)$ , so that  $\varphi_\phi(\theta) = 0$  and  $\varphi_I(\theta) = I_0^*$ . We also consider the constant values  $\rho = \rho^{(0)} = \rho_0$  and  $\delta = \delta^{(0)} = \rho_0/12$ . In particular,  $\rho^{(1)} = \rho^{(0)} - 3\delta^{(0)} > 0$ . Clearly, the parametrization  $\varphi$  is analytic in  $\Delta(\rho) = \Delta(\rho_0)$ , and we can easily check that:

$$e^{(0)}(\theta) = (f_\phi(\theta, I_0^*), f_I(\theta, I_0^*)), \quad A^{(0)}(\theta) = -\alpha'(I_0^*) - \partial_I f_\phi(\theta, I_0^*).$$

Hence, we can set  $\mu = \mu^{(0)} = c$  and we have:

$$\begin{aligned} \|\varphi_\phi^{(0)}\|_{\rho^{(0)}} &= 0 = \rho_0 - \rho^{(0)}, \quad \|\varphi_I^{(0)} - I_0^*\|_{\rho^{(0)}} = 0 \leq R_0 - \frac{c^*\mu^{(0)}}{\gamma^2(\delta^{(0)})^{2\nu}}, \\ \|(\varphi_s^{(0)})'\|_{\rho^{(0)}} &= 0 \leq \frac{1}{4} - \frac{c^*\mu^{(0)}}{\gamma^2(\delta^{(0)})^{2\nu+1}}, \quad |\langle A^{(0)} \rangle| \geq c_1 - c \geq \frac{c_1}{2}, \end{aligned}$$

for  $s \in \{\phi, I\}$ , provided that  $\frac{c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2\nu}} \leq \min\{R_0, \delta^{(0)}/4\}$  and  $c \leq c_1/2$ , conditions that are guaranteed by hypotheses (27) on the statement. Thus, we can apply Lemma 3 to  $\varphi^{(0)}$  and we can define a new parametrization of quasi-torus  $\varphi^{(1)}$  that verifies bounds on the statement of the lemma.

**A close expression for the iterative error.** To iteratively apply Lemma 3, we are going to select, at each step  $n \geq 0$  of the iterative scheme, the constant values  $\delta = \delta^{(n)} = \delta^{(0)}/2^n$  and  $\rho = \rho^{(n)} = \rho^{(0)} - 3 \sum_{j=0}^{n-1} \delta^{(j)}$ , for which we clearly have  $\rho^{(n)} \geq \rho_0/2$ , for any  $n \geq 0$ . Let us suppose for a moment that we can freely iterate for any  $n \geq 0$  by dealing with these selected values. Then, we have  $\mu^{(0)} = c$  and we get the following iterative expression for the error  $\mu = \mu^{(n)}$  after  $n$  iterations:

$$\mu^{(n+1)} = \frac{c^* (\mu^{(n)})^2}{\gamma^4(\delta^{(n)})^{4\nu+1}} = \frac{c^*}{\gamma^4(\delta^{(0)})^{4\nu+1}} 2^{(4\nu+1)n} (\mu^{(n)})^2.$$

Simple computations show that we obtain

$$\begin{aligned} \mu^{(n)} &= \left( \frac{c^*}{\gamma^4(\delta^{(0)})^{4\nu+1}} \right)^{2^n-1} 2^{(4\nu+1)(2^n-(n+1))} (\mu^{(0)})^{2^n} \\ (40) \quad &= \left( \frac{c^*}{\gamma^4(\delta^{(0)})^{4\nu+1}} \right)^{-1} 2^{-(4\nu+1)(n+1)} \left( \frac{2^{4\nu+1} c^* \mu^{(0)}}{\gamma^4(\delta^{(0)})^{4\nu+1}} \right)^{2^n}, \end{aligned}$$

expression that can easily be verified by induction. In particular, since hypotheses on  $\mu^{(0)} = c$  mean that  $\frac{2^{4\nu+1} c^* \mu^{(0)}}{\gamma^4(\delta^{(0)})^{4\nu+1}} \leq \frac{1}{2}$ , we conclude that this expression verifies  $\mu^{(n)} \rightarrow 0$  as  $n \rightarrow \infty$ . In addition, we also have that the following estimates hold for every  $n \geq 0$ :

$$\frac{c^* \mu^{(n)}}{\gamma^2(\delta^{(n)})^{2\nu}} \leq \frac{c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2\nu}} \left( \frac{1}{2} \right)^n, \quad \frac{c^* \mu^{(n)}}{\gamma^2(\delta^{(n)})^{2\nu+1}} \leq \frac{c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2\nu+1}} \left( \frac{1}{2} \right)^n.$$

Indeed, using that  $2^n \geq n+1$  and that  $\frac{2^{4\nu+1} c^* \mu^{(0)}}{\gamma^4(\delta^{(0)})^{4\nu+1}} \leq \frac{1}{2}$ , we have:

$$\begin{aligned} \frac{c^* \mu^{(n)}}{\gamma^2(\delta^{(n)})^{2\nu}} &= \frac{\gamma^2(\delta^{(0)})^{2\nu+1}}{2^{2\nu}} 2^{-(2\nu+1)(n+1)} \left( \frac{2^{4\nu+1} c^* \mu^{(0)}}{\gamma^4(\delta^{(0)})^{4\nu+1}} \right)^{n+1} \\ &= \frac{2^{2\nu+1} c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2\nu}} 2^{-(2\nu+1)(n+1)} \left( \frac{1}{2} \right)^n \\ &\leq \frac{c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2\nu}} \left( \frac{1}{2} \right)^n, \end{aligned}$$

and analogously for the other case.

**General step.** We iteratively apply Lemma 3 starting with the 0-data selected above and suppose that we have been able to iterate  $n$  times, for some  $n \geq 1$ , so we define the family of corrected parameterizations  $\{\varphi^{(j)}\}_{j=0}^n$ , for which the estimates provided by Lemma 3 hold in the corresponding domain  $\Delta(\rho^{(j)})$ , for each  $j = 1, \dots, n$ . In particular, the size of the errors  $\|e^{(j)}\|_{\rho^{(j)}}$  and  $\|(e^{(j)})'\|_{\rho^{(j)}}$  is controlled by  $\mu^{(j)}$  given by expression in (40). We want to show that we can iterate again by showing that  $\varphi^{(n)}$  fit into the conditions needed to apply the Iterative

Lemma. Specifically, we need to check, for any  $n \geq 1$ :

$$\|\varphi_\phi^{(n)}\|_{\rho^{(n)}} \leq \rho_0 - \rho^{(n)}, \quad \|\varphi_I^{(n)} - I_0^*\|_{\rho^{(n)}} \leq R_0 - \frac{c^* \mu^{(n)}}{\gamma^2(\delta^{(n)})^{2v}}, \quad \|(\varphi_s^{(n)})'\|_{\rho^{(n)}} \leq \frac{1}{4} - \frac{c^* \mu^{(n)}}{\gamma^2(\delta^{(n)})^{2v+1}},$$

for  $s \in \{\phi, I\}$ , and  $|\langle A^{(n)} \rangle| \geq c_1/2$ . Notice that

$$\|\varphi_\phi^{(n)}\|_{\rho^{(n)}} \leq \sum_{j=0}^{n-1} \|\Delta \varphi_\phi^{(j)}\|_{\rho^{(j+1)}} \leq \sum_{j=0}^{n-1} \frac{c^* \mu^{(j)}}{\gamma^2(\delta^{(j)})^{2v}} \leq \frac{c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2v}} \sum_{j=0}^{n-1} \left(\frac{1}{2}\right)^j \leq \frac{2c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2v}},$$

and the same bound holds for  $\|\varphi_I^{(n)} - I_0^*\|_{\rho^{(n)}}$ . We proceed analogously to get the bounds

$$\|(\varphi_s^{(n)})'\|_{\rho^{(n)}} \leq \frac{2c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2v+1}}, \text{ for } s \in \{\phi, I\}, \text{ by means of the sum of the bounds on } \|(\Delta \varphi_s^{(n)})'\|_{\rho^{(j+1)}}$$

for  $j = 0, \dots, n-1$ . Finally, to control the average  $\langle A^{(n)} \rangle$  we start by bounding

$$\|A^{(n)} - A^{(0)}\|_{\rho^{(n)}} \leq \sum_{j=0}^{n-1} \|A^{(j+1)} - A^{(j)}\|_{\rho^{(j+1)}} \leq \frac{2c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2v+1}}.$$

Then, bounding the size of the average  $\langle A^{(n)} - A^{(0)} \rangle$  by the norm of the function and using the estimate above  $|\langle A^{(0)} \rangle| \geq c_1 - c$ , we obtain:

$$|\langle A^{(n)} \rangle| \geq |\langle A^{(0)} \rangle| - |\langle A^{(n)} - A^{(0)} \rangle| \geq |\langle A^{(0)} \rangle| - \|A^{(n)} - A^{(0)}\|_{\rho^{(n)}} \geq c_1 - c - \frac{2c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2v+1}}.$$

Using that  $\rho_0 - \rho^{(n)} \geq \rho_0 - \rho^{(1)} = 3\delta^{(0)} = \rho_0/4$ , for any  $n \geq 1$ , the bound  $\|\varphi_\phi^{(n)}\|_{\rho^{(n)}} \leq \rho_0 - \rho^{(n)}$  is achieved provided that  $\frac{2c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2v}} \leq \rho_0/4$ . Next, using that  $\frac{c^* \mu^{(n)}}{\gamma^2(\delta^{(n)})^{2v}} \leq \frac{c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2v}}$  and  $\frac{c^* \mu^{(n)}}{\gamma^2(\delta^{(n)})^{2v+1}} \leq \frac{c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2v+1}}$ , the desired bounds on  $\|\varphi_I^{(n)} - I_0^*\|_{\rho^{(n)}}$ ,  $\|(\varphi_s^{(n)})'\|_{\rho^{(n)}}$  and  $|\langle A^{(n)} \rangle|$ , are fulfilled provided that  $\frac{3c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2v}} \leq R_0$ ,  $\frac{3c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2v+1}} \leq 1/4$  and  $c + \frac{2c^* \mu^{(0)}}{\gamma^2(\delta^{(0)})^{2v+1}} \leq c_1/2$ , respectively. All these bounds are guaranteed by the conditions on  $c$  on the statement of the theorem.

**Convergence of the iterative scheme.** Once we have verified that hypotheses on the statement allow us to compute the full sequence of iterates  $\{\varphi^{(n)}\}_{n=0}^\infty$ , the convergence of  $\varphi^* = \lim_{n \rightarrow \infty} \varphi^{(n)}$  in  $\Delta(\rho^*)$ , with  $\rho^* = \lim_{n \rightarrow \infty} \rho^{(n)} = \rho_0/2$ , is straightforward as the series  $\varphi^* = \varphi^{(0)} + \sum_{n=0}^\infty \Delta \varphi^{(n)}$  is absolutely convergent in terms of the norm  $\|\cdot\|_{\rho_0/2}$ . Estimates on  $\varphi^*$  on the statement are also straightforward from computations above.  $\square$

#### 4. THE IMPACT MAP: PROOF OF PROPOSITION 2

From the  $2\pi$ -periodic function  $p$  of the statement, we define the functions

$$P_1(\tau, t_0) := \int_0^\tau p(s + t_0) ds \quad \text{and} \quad P_2(\tau, t_0) := \int_0^\tau P_1(s, t_0) ds.$$

By considering the function  $\tilde{P}_1$ ,  $\tilde{P}_2$  and  $a_0$  introduced in (18), we can rewrite  $P_1$  and  $P_2$  as

$$P_1(\tau, t_0) = a_0 \tau + \tilde{P}_1(t_0 + \tau) - \tilde{P}_1(t_0) \quad \text{and} \quad P_2(\tau, t_0) = a_0 \frac{\tau^2}{2} - \tau \tilde{P}_1(t_0) + \tilde{P}_2(t_0 + \tau) - \tilde{P}_2(t_0).$$

With the functions  $P_1$  and  $P_2$  defined in (41) we can provide the explicit solutions of the system originating from (4). For  $x_0 > 0$ , we have

$$(42) \quad \begin{cases} t_r^\tau(t_0, x_0, y_0; \varepsilon) = \tau + t_0, \\ x_r^\tau(t_0, x_0, y_0; \varepsilon) = x_0 + \tau y_0 - \frac{\tau^2}{2} + \varepsilon P_2(\tau, t_0), \\ y_r^\tau(t_0, x_0, y_0; \varepsilon) = y_0 - \tau + \varepsilon P_1(\tau, t_0). \end{cases}$$

If  $x_1 < 0$ , we have

$$(43) \quad \begin{cases} t_l^\tau(t_1, x_1, y_1; \varepsilon) = \tau + t_1, \\ x_l^\tau(t_1, x_1, y_1; \varepsilon) = x_1 + \tau y_1 + \frac{\tau^2}{2} + \varepsilon P_2(\tau, t_1), \\ y_l^\tau(t_1, x_1, y_1; \varepsilon) = y_1 + \tau + \varepsilon P_1(\tau, t_1). \end{cases}$$

By expressions in (42) and (43), we notice that the times of impact  $\tau^+(t_0, y_0; \varepsilon)$  and  $\tau^-(t_1, y_1; \varepsilon)$ , introduced in (8) and (10), respectively, cannot be explicitly obtained, so we cannot deal with explicit expressions for the impact maps  $\mathcal{P}_\varepsilon^+$ ,  $\mathcal{P}_\varepsilon^-$ , and  $\mathcal{P}_\varepsilon$ . For this reason, we propose the following approach to deal with them.

**Asymptotic behaviour of the first impact time.** Since we are interested in discussing the persistence of quasi-periodic invariant curves of  $\mathcal{P}_\varepsilon$  of arbitrarily large amplitude  $y_0$  and  $0 < \varepsilon \ll 1$ , we start by discussing an asymptotic expression for  $\tau^+(t_0, y_0; \varepsilon)$  and  $\tau^-(t_1, y_1; \varepsilon)$  when  $y_0 \rightarrow +\infty$  and  $y_1 \rightarrow -\infty$ , respectively. Using expressions (41) for  $P_2$  and (42) for  $x_r^\tau$ , equation (8) for the first impact time  $\tau^+ = \tau^+(t_0, y_0; \varepsilon)$  reads as:

$$(44) \quad \tau^+ y_0 - \frac{1 - a_0 \varepsilon}{2} (\tau^+)^2 - \varepsilon \tau^+ \tilde{P}_1(t_0) + \varepsilon \tilde{P}_2(t_0 + \tau^+) - \varepsilon \tilde{P}_2(t_0) = 0.$$

Hence, for  $y_0 \gg 0$  sufficiently large, one expects the dominant part of the expression of  $\tau^+$  to be:

$$(45) \quad \tau_0^+(t_0, y_0; \varepsilon) = \frac{2(y_0 - \varepsilon \tilde{P}_1(t_0))}{1 - a_0 \varepsilon},$$

so it is natural to write

$$(46) \quad \tau^+(t_0, y_0; \varepsilon) = \tau_0^+(t_0, y_0; \varepsilon) + \varepsilon \tau_*^+(t_0, y_0; \varepsilon),$$

where  $\tau_*^+$  is a small correction (as  $y_0 \rightarrow +\infty$ ) satisfying properties to be given in the next lemma.

**Lemma 4.** *Let  $\tilde{P}_1$  and  $\tilde{P}_2$  be the real analytic and  $2\pi$ -periodic functions introduced in (18) which verify  $\tilde{P}_2' = \tilde{P}_1$  and  $\|\tilde{P}_j\|_\rho \leq \tilde{p}$ ,  $j = 1, 2$ , for  $0 < \rho < 1$  and  $\tilde{p} > 0$ . Given  $y_0^* \geq 1$ ,  $0 < \bar{\rho}_0 \leq \rho$ ,  $\bar{\rho}_0^+ > 0$ , and  $\bar{\rho}_0^+ > 0$ , such that  $\bar{\rho}_0^+ + 4\bar{\rho}_0^+ < \bar{\rho}_0$ , we consider the set  $\mathcal{D}_0^+ = \mathcal{D}^+(y_0^*, \bar{\rho}_0^+, \bar{\rho}_0^+)$  (see (16)) and define  $\varepsilon_*^+$  as*

$$\varepsilon_*^+ = \sup \left\{ \varepsilon > 0 : |a_0 \varepsilon| \leq \frac{1}{2} \quad \text{and} \quad \bar{\rho}_0^+ + 4\bar{\rho}_0^+ + 288\varepsilon \tilde{p} < \bar{\rho}_0 \right\}.$$

*Then, for every  $0 < \varepsilon \leq \varepsilon_*^+$ , the first impact time  $\tau^+$  can be written as in (45) and (46), with  $\tau_0^+(t_0, y_0; \varepsilon)$  and  $\tau_*^+(t_0, y_0; \varepsilon)$  being real analytic functions for  $(t_0, y_0) \in \mathcal{D}_0^+$ ,  $2\pi$ -periodic in  $t_0$ , and satisfying:*

$$(47) \quad \frac{|y_0|}{2} \leq |\tau_0^+(t_0, y_0; \varepsilon)| \leq 6|y_0|, \quad |\tau_*^+(t_0, y_0; \varepsilon)| \leq \frac{32\tilde{p}}{|y_0|}.$$

Moreover, given any  $0 < \hat{\rho}_0^+ < \min\{\bar{\rho}_0^+, \tilde{\rho}_0^+\}/2$ , there exists  $C_*^+ = C_*^+(\tilde{\rho}, \hat{\rho}_0^+) > 0$  such that

$$(48) \quad \left| \partial_{t_0}^i \partial_{y_0}^j \tau_*^+(t_0, y_0; \varepsilon) \right| \leq \frac{C_*^+}{|y_0|},$$

for every  $(t_0, y_0) \in \mathcal{D}^+(y_0^* + 2\hat{\rho}_0^+, \bar{\rho}_0^+ - 2\hat{\rho}_0^+, \bar{\rho}_0^+ - 2\hat{\rho}_0^+)$ ,  $0 < \varepsilon \leq \varepsilon_*^+$ , and  $0 \leq i + j \leq 2$ .

*Proof.* We start by observing that for any  $(t_0, y_0) \in \mathcal{D}_0^+$  and  $0 < \varepsilon \leq \varepsilon_*^+$ , the function  $\tau_0^+$  of (45) satisfies the estimates in (47). This is straightforward by observing that  $|y_0| \geq y_0^* \geq 1$ ,  $|a_0\varepsilon| \leq \frac{1}{2}$ , and that  $\varepsilon\tilde{p} < \frac{1}{2}$ . Indeed, using that  $1/2 \leq |1 - a_0\varepsilon| \leq 3/2$ , we have:

$$\frac{1}{2} \leq \frac{2}{3} \leq \frac{2}{|1 - a_0\varepsilon|} \left( 1 - \frac{\varepsilon|\tilde{P}_1(t_0)|}{|y_0|} \right) \leq \left| \frac{\tau_0^+(t_0, y_0; \varepsilon)}{y_0} \right| \leq \frac{2}{|1 - a_0\varepsilon|} \left( 1 + \frac{\varepsilon|\tilde{P}_1(t_0)|}{|y_0|} \right) \leq 6.$$

If we write  $\tau^+$  as in (46), and we impose that it satisfies equation (44), we have that  $\tau_*^+ = \tau_*^+(t_0, y_0; \varepsilon)$  satisfies the equation  $\tau_*^+ = \mathcal{F}^+(\tau_*^+)$ , where the non-linear functional  $\mathcal{F}^+ = \mathcal{F}_\varepsilon^+$  is defined as:

$$\mathcal{F}^+(\tau)(t_0, y_0) = \frac{2(\tilde{P}_2(t_0 + \tau_0^+(t_0, y_0; \varepsilon) + \varepsilon\tau(t_0, y_0; \varepsilon)) - \tilde{P}_2(t_0))}{(1 - a_0\varepsilon)(\tau_0^+ + \varepsilon\tau(t_0, y_0; \varepsilon))}.$$

We introduce the Banach space  $(\mathcal{X}, \|\cdot\|_{\mathcal{D}^+})$  defined by:

$$\mathcal{X} = \{\tau : \overline{\mathcal{D}_0^+} \rightarrow \mathbb{C}, f \text{ real analytic in } \mathcal{D}_0^+ \text{ and bounded up to the boundary, } \|\tau\|_{\mathcal{D}_0^+} < +\infty\},$$

where  $\|\tau\|_{\mathcal{D}_0^+} = \sup_{(t_0, y_0) \in \overline{\mathcal{D}_0^+}} \{|y_0| \tau(t_0, y_0)|\}$ . We denote as  $\mathcal{B}_{32\tilde{p}} \subset \mathcal{X}$  the closed ball of center

zero and radius  $32\tilde{p}$  in terms of the norm  $\|\cdot\|_{\mathcal{D}_0^+}$ . Since  $\mathcal{B}_{32\tilde{p}}$  is a closed set of  $(\mathcal{X}, \|\cdot\|_{\mathcal{D}_0^+})$ , it constitutes a complete metric space in terms of the distance induced by the considered norm. We are going to prove that if  $\varepsilon > 0$  is as in the statement, then  $\mathcal{F}^+ : \mathcal{B}_{32\tilde{p}} \rightarrow \mathcal{B}_{32\tilde{p}}$  is a contraction. In this way, using the fixed point theorem, we establish the well defined character of  $\tau_*^+$  in  $\mathcal{D}_0^+$  as well as the bound  $|\tau_*^+(t_0, y_0; \varepsilon)| \leq \frac{32\tilde{p}}{|y_0|}$ . To achieve this purpose, first we show that that  $\mathcal{F}^+(\tau)$  is well defined and belongs to  $\mathcal{X}$ , for every  $\tau \in \mathcal{B}_{32\tilde{p}}$ . First of all, we notice that if  $\tau \in \mathcal{B}_{32\tilde{p}}$ ,  $(t_0, y_0) \in \mathcal{D}_0^+$  and  $0 < \varepsilon \leq \varepsilon_*^+$ , then (we omit the dependence of  $\tau$  and  $\tau_0^+$  on  $t_0, y_0$  and  $\varepsilon$  unless necessary)

$$\begin{aligned} |\operatorname{Im}(t_0 + \tau_0^+ + \varepsilon\tau)| &= \left| \operatorname{Im}(t_0) + \frac{2}{1 - a_0\varepsilon} (\operatorname{Im}(y_0) + \varepsilon\operatorname{Im}(\tilde{P}_1(t_0))) + \varepsilon\operatorname{Im}(\tau) \right| \\ &\leq \bar{p}_0^+ + 4\tilde{p}_0^+ + 4\varepsilon\tilde{p} + 32\varepsilon\tilde{p} \\ &\leq \bar{\rho}_0, \end{aligned}$$

where we are bounding the size of the imaginary parts of  $\tilde{P}_1$  and  $\tau$  by its norms, and using that since  $y_0^* \geq 1$  then  $|\tau(t_0, y_0)| \leq 32\tilde{p}$ . This means that the composition  $\tilde{P}_2(t_0 + \tau_0^+ + \varepsilon\tau)$  is well defined and also analytic. Moreover, we also have that:

$$|\tau_0^+ + \varepsilon\tau| \geq |\tau_0^+| - \varepsilon|\tau| \geq \frac{|y_0|}{2} - \varepsilon \frac{|y_0\tau|}{|y_0|} \geq \frac{|y_0|}{2} \left( 1 - \frac{64\varepsilon\tilde{p}}{y_0^*} \right) \geq \frac{|y_0|}{4} \geq \frac{1}{4}.$$

Computations above guarantee the well defined character of  $\mathcal{F}^+$ . Next, we compute

$$\mathcal{F}^+(0) = \frac{2(\tilde{P}_2(t_0 + \tau_0^+) - \tilde{P}_2(t_0))}{(1 - a_0\varepsilon)\tau_0^+}.$$



Using bounds (47) on  $\tau_0^+$ , we have that

$$|\mathcal{F}^+(0)(t_0, y_0; \varepsilon)| \leq 16 \frac{\tilde{p}}{|y_0|}.$$

Therefore,  $\|\mathcal{F}^+(0)\|_{\mathcal{D}_0^+} < 16\tilde{p}$  and, consequently,  $\mathcal{F}^+(0) \in \mathcal{B}_{32\tilde{p}}$ . Now, we take  $\tau, \bar{\tau} \in \mathcal{B}_{32\tilde{p}}$ . Then:

$$\begin{aligned} |\mathcal{F}^+(\tau) - \mathcal{F}^+(\bar{\tau})| &= \left| \frac{2(\tilde{P}_2(t_0 + \tau_0^+ + \varepsilon\tau) - \tilde{P}_2(t_0))}{(1 - a_0\varepsilon)(\tau_0^+ + \varepsilon\tau)} - \frac{2(\tilde{P}_2(t_0 + \tau_0^+ + \varepsilon\bar{\tau}) - \tilde{P}_2(t_0))}{(1 - a_0\varepsilon)(\tau_0^+ + \varepsilon\bar{\tau})} \right| \\ &\leq 4 \left| \frac{\tilde{P}_2(t_0 + \tau_0^+ + \varepsilon\tau) - \tilde{P}_2(t_0 + \tau_0^+ + \varepsilon\bar{\tau})}{\tau_0^+ + \varepsilon\tau} \right| \\ &\quad + 4\varepsilon \left| \frac{(\tilde{P}_2(t_0 + \tau_0^+ + \varepsilon\bar{\tau}) - \tilde{P}_2(t_0))(\tau - \bar{\tau})}{(\tau_0^+ + \varepsilon\tau)(\tau_0^+ + \varepsilon\bar{\tau})} \right|. \end{aligned}$$

Thus, it follows from the Mean Value Theorem that

$$\left| \frac{\tilde{P}_2(t_0 + \tau_0^+ + \varepsilon\tau) - \tilde{P}_2(t_0 + \tau_0^+ + \varepsilon\bar{\tau})}{\tau_0^+ + \varepsilon\tau} \right| \leq 4\varepsilon\tilde{p}|\tau - \bar{\tau}|.$$

On the other hand

$$\left| \frac{(\tilde{P}_2(t_0 + \tau_0^+ + \varepsilon\bar{\tau}) - \tilde{P}_2(t_0))(\tau - \bar{\tau})}{(\tau_0^+ + \varepsilon\tau)(\tau_0^+ + \varepsilon\bar{\tau})} \right| \leq 32\tilde{p}|\tau - \bar{\tau}|.$$

Then, taking into account both bounds and knowing that  $288\varepsilon\tilde{p} \leq 1$ , it follows that

$$\|\mathcal{F}^+(\tau) - \mathcal{F}^+(\bar{\tau})\|_{\mathcal{D}_0^+} \leq 144\varepsilon\tilde{p}\|\tau - \bar{\tau}\|_{\mathcal{D}_0^+} \leq \frac{1}{2}\|\tau - \bar{\tau}\|_{\mathcal{D}_0^+},$$

which means that  $\mathcal{F}^+ : \mathcal{B}_{32\tilde{p}} \rightarrow \mathcal{X}$  is a contraction. Finally, if  $\tau \in \mathcal{B}_{32\tilde{p}}$ , then

$$\begin{aligned} \|\mathcal{F}^+(\tau)\|_{\mathcal{D}_0^+} &\leq \|\mathcal{F}^+(\tau) - \mathcal{F}^+(0) + \mathcal{F}^+(0)\|_{\mathcal{D}_0^+} \\ &\leq \|\mathcal{F}^+(\tau) - \mathcal{F}^+(0)\|_{\mathcal{D}_0^+} + \|\mathcal{F}^+(0)\|_{\mathcal{D}_0^+} \\ &\leq \frac{1}{2}\|\tau\|_{\mathcal{D}_0^+} + 16\tilde{p} \\ &\leq 32\tilde{p}, \end{aligned}$$

which means that  $\mathcal{F}^+ : \mathcal{B}_{32\tilde{p}} \rightarrow \mathcal{B}_{32\tilde{p}}$  is well defined and it is also a contraction as wanted. Consequently,  $\mathcal{F}^+$  has a unique fixed point  $\tau_*^+ \in \mathcal{B}_{32\tilde{p}}$  and, therefore, it satisfies bounds in (48).

In order to obtain the bounds for the first and second order partial derivatives of  $\tau_*^+$ , with respect  $t_0$  and  $y_0$ , we apply standard Cauchy estimates to the function  $\eta^+(t_0, y_0; \varepsilon) = y_0\tau_*^+(t_0, y_0; \varepsilon)$ . We use that  $|\eta^+| \leq 32\tilde{p}$  in  $\mathcal{D}_0^+ = \mathcal{D}^+(y_0^*, \bar{\rho}_0^+, \bar{\rho}_0^+)$  and we shrink the complex width of the domain by  $\hat{\rho}_0^+$  when performing each derivative. In this way, it is not difficult to see that there is a constant  $C_*^+ = C_*^+(\tilde{p}, \hat{\rho}_0^+) > 0$  for which equation (48) holds.  $\square$

We can analogously discuss the asymptotic behaviour of the time for impact of the negative solutions  $\tau^- = \tau^-(t_1, y_1; \varepsilon)$ . From expressions in (43) for the integration of (13) for  $x < 0$ , it verifies the equation:

$$\tau^- y_1 + \frac{1 + a_0\varepsilon}{2}(\tau^-)^2 - \varepsilon\tau^- \tilde{P}_1(t_1) + \varepsilon\tilde{P}_2(t_1 + \tau^-) - \varepsilon\tilde{P}_2(t_1) = 0,$$

which motivates to write, if  $y_1 \ll 0$ ,

$$(49) \quad \tau^-(t_1, y_1; \varepsilon) = \tau_0^-(t_1, y_1; \varepsilon) + \varepsilon \tau_*^-(t_1, y_1; \varepsilon), \quad \tau_0^-(t_1, y_1; \varepsilon) = -\frac{2(y_1 - \varepsilon \tilde{P}_1(t_1))}{1 + a_0 \varepsilon}.$$

Estimates concerning  $\tau_0^-(t_1, y_1; \varepsilon)$  and  $\tau_*^-(t_1, y_1; \varepsilon)$  are stated in Lemma 5, whose proof is completely analogous to that of lemma 4.

**Lemma 5.** *Let  $\tilde{P}_1$  and  $\tilde{P}_2$  be as in lemma 4. Given  $y_1^* \geq 1$ ,  $0 < \bar{\rho}_1 \leq \rho$ ,  $\bar{\rho}_1^- > 0$ , and  $\bar{\rho}_1^- > 0$ , such that  $\bar{\rho}_1^- + 4\bar{\rho}_1^- < \bar{\rho}_1$ , we consider the set  $\mathcal{D}_1^- = \mathcal{D}^+(y_1^*, \bar{\rho}_1^-, \bar{\rho}_1^-)$  (see (17)) and define  $\varepsilon_*^-$  as*

$$\varepsilon_*^- = \sup \left\{ \varepsilon > 0 : |a_0 \varepsilon| \leq \frac{1}{2} \quad \text{and} \quad \bar{\rho}_1^- + 4\bar{\rho}_1^- + 288\varepsilon \bar{p} < \bar{\rho}_1 \right\}.$$

*Then, for every  $0 < \varepsilon \leq \varepsilon_*^-$ , the first impact time  $\tau^-$  can be written as in (49), with  $\tau_0^-(t_1, y_1; \varepsilon)$  and  $\tau_*^-(t_1, y_1; \varepsilon)$  being real analytic functions for  $(t_1, y_1) \in \mathcal{D}_1^-$ ,  $2\pi$ -periodic in  $t_1$ , and satisfying:*

$$\frac{|y_1|}{2} \leq |\tau_1^-(t_1, y_1; \varepsilon)| \leq 6|y_1|, \quad |\tau_*^-(t_1, y_1; \varepsilon)| \leq \frac{32\bar{p}}{|y_1|}$$

*Besides that, given any  $0 < \hat{\rho}_1^- < \min\{\bar{\rho}_1^-, \tilde{\rho}_1^-\}/2$ , there exists  $C_*^- = C_*(\tilde{p}, \hat{\rho}_1^-) > 0$  such that*

$$\left| \partial_{t_1}^i \partial_{y_1}^j \tau_*^-(t_1, y_1; \varepsilon) \right| \leq \frac{C_*^-}{|y_1|},$$

*for every  $(t_1, y_1) \in \mathcal{D}^-(y_1^* + 2\hat{\rho}_1^-, \bar{\rho}_1^- - 2\hat{\rho}_1^-, \bar{\rho}_1^- - 2\hat{\rho}_1^-)$ ,  $0 < \varepsilon \leq \varepsilon_*^-$ , and  $0 \leq i + j \leq 2$ .*

**Expressions of Impact Map  $\mathcal{P}_\varepsilon$ .** We use the asymptotic expressions for the impact times  $\tau^+$  and  $\tau^-$  obtained in lemmas 4 and 5 to provide formulas for the half positive impact map  $\mathcal{P}_\varepsilon^+(t_0, y_0)$  and the half negative impact map  $\mathcal{P}_\varepsilon^-(t_1, y_1)$  which make apparent the dominant terms of both maps when  $y_0 \gg 0$  and  $y_1 \ll 0$ , respectively. Then, by composition of both maps, we obtain the corresponding expression  $\mathcal{P}_\varepsilon = \mathcal{P}_\varepsilon^- \circ \mathcal{P}_\varepsilon^+(t_0, y_0)$ .

Let us denote  $(t_1(t_0, y_0; \varepsilon), y_1(t_0, y_0; \varepsilon)) = \mathcal{P}_\varepsilon^+(t_0, y_0)$ . Using formulas in (9), (42), (45), and (46), we obtain:

$$(50) \quad t_1(t_0, y_0; \varepsilon) = t_0 + \tau^+(t_0, y_0; \varepsilon) = t_0 + \frac{2(y_0 - \varepsilon \tilde{P}_1(t_0))}{1 - a_0 \varepsilon} + \varepsilon \tau_*^+(t_0, y_0; \varepsilon),$$

and also using (41):

$$(51) \quad \begin{aligned} y_1(t_0, y_0; \varepsilon) &= y_0 - \tau^+(t_0, y_0; \varepsilon) + \varepsilon P_1(\tau^+(t_0, y_0; \varepsilon), t_0) \\ &= -y_0 + \varepsilon (\tilde{P}_1(t_1(t_0, y_0; \varepsilon)) + \tilde{P}_1(t_0) - (1 - a_0 \varepsilon) \tau_*^+(t_0, y_0; \varepsilon)). \end{aligned}$$

Let us denote  $(\bar{t}_0(t_0, y_0; \varepsilon), \bar{y}_0(t_0, y_0; \varepsilon)) = \mathcal{P}_\varepsilon(t_0, y_0)$ . By using definitions above, we have that  $(\bar{t}_0(t_0, y_0; \varepsilon), \bar{y}_0(t_0, y_0; \varepsilon)) = \mathcal{P}_\varepsilon^-(t_1(t_0, y_0; \varepsilon), y_1(t_0, y_0; \varepsilon))$ . We combine computations and formulas above with those in (11), (43), and (49) to obtain expressions below for  $\bar{t}_0(t_0, y_0; \varepsilon)$  and  $\bar{y}_0(t_0, y_0; \varepsilon)$ . Although in these computations  $t_1 = t_1(t_0, y_0; \varepsilon)$  and  $y_1 = y_1(t_0, y_0; \varepsilon)$  through formulas in (50) and (51), we omit their dependence on  $t_0$ ,  $y_0$ , and  $\varepsilon$ . We have:

$$(52) \quad \begin{aligned} \bar{t}_0(t_0, y_0; \varepsilon) &= t_1 + \tau^-(t_1, y_1; \varepsilon) \\ &= t_0 + \frac{2(y_0 - \varepsilon \tilde{P}_1(t_0))}{1 - a_0 \varepsilon} + \varepsilon \tau_*^+(t_0, y_0; \varepsilon) - \frac{2(y_1 - \varepsilon \tilde{P}_1(t_1))}{1 + a_0 \varepsilon} + \varepsilon \tau_*^-(t_1, y_1; \varepsilon) \\ &= t_0 + \frac{4y_0}{1 - a_0^2 \varepsilon^2} + \varepsilon \left( -\frac{4\tilde{P}_1(t_0)}{1 - a_0^2 \varepsilon^2} + \frac{3 - a_0 \varepsilon}{1 + a_0 \varepsilon} \tau_*^+(t_0, y_0; \varepsilon) + \tau_*^-(t_1, y_1; \varepsilon) \right), \end{aligned}$$

and (we also omit the dependence of  $\bar{t}_0 = \bar{t}_0(t_0, y_0; \varepsilon)$  on  $t_0, y_0$ , and  $\varepsilon$ ) using again (41) and (43), and formula (49) for  $\tau^-$ ,

$$\begin{aligned}\bar{y}_0(t_0, y_0; \varepsilon) &= y_1 + \tau^-(t_1, y_1; \varepsilon) + \varepsilon P_1(\tau^-(t_1, y_1; \varepsilon), t_1) \\ &= y_1 + (1 + a_0 \varepsilon) \tau^-(t_1, y_1; \varepsilon) + \varepsilon (\tilde{P}_1(\bar{t}_0) - \tilde{P}_1(t_1)) \\ &= -y_1 + \varepsilon (\tilde{P}_1(\bar{t}_0) + \tilde{P}_1(t_1) + (1 + a_0 \varepsilon) \tau_*^-(t_1, y_1; \varepsilon)) \\ &= y_0 + \varepsilon (\tilde{P}_1(\bar{t}_0) - \tilde{P}_1(t_0) + (1 - a_0 \varepsilon) \tau_*^+(t_0, y_0; \varepsilon) + (1 + a_0 \varepsilon) \tau_*^-(t_1, y_1; \varepsilon)).\end{aligned}$$

Consequently, the impact map in coordinates  $(t_0, y_0)$  takes the form:

$$\mathcal{P}_\varepsilon : \begin{cases} \bar{t}_0 = t_0 + \alpha_\varepsilon(y_0) + \varepsilon f_{t_0}(t_0, y_0; \varepsilon), \\ \bar{y}_0 = y_0 + \varepsilon f_{y_0}(t_0, y_0; \varepsilon), \end{cases}$$

with

$$(53) \quad \alpha_\varepsilon(y_0) = \frac{4y_0}{1 - a_0^2 \varepsilon^2},$$

$$(54) \quad f_{t_0}(t_0, y_0; \varepsilon) = -\frac{4\tilde{P}_1(t_0)}{1 - a_0^2 \varepsilon^2} + \frac{3 - a_0 \varepsilon}{1 + a_0 \varepsilon} \tau_*^+(t_0, y_0; \varepsilon) + \tau_*^-(t_1(t_0, y_0; \varepsilon), y_1(t_0, y_0; \varepsilon); \varepsilon),$$

and

$$(55) \quad \begin{aligned} f_{y_0}(t_0, y_0; \varepsilon) &= \tilde{P}_1(\bar{t}_0(t_0, y_0; \varepsilon)) - \tilde{P}_1(t_0) + (1 - a_0 \varepsilon) \tau_*^+(t_0, y_0, \varepsilon) \\ &\quad + (1 + a_0 \varepsilon) \tau_*^-(t_1(t_0, y_0, \varepsilon), y_1(t_0, y_0, \varepsilon); \varepsilon), \end{aligned}$$

where  $t_1(t_0, y_0; \varepsilon)$ ,  $y_1(t_0, y_0; \varepsilon)$  are given by (50), (51), respectively, and  $\bar{t}_0 = \bar{t}_0(t_0, y_0; \varepsilon)$  is given by the first component of (20).

**Estimates on the Impact Map  $\mathcal{P}_\varepsilon$  (Proof of Proposition 2).** The most important aspect to be addressed in order to provide quantitative estimates for  $\mathcal{P}_\varepsilon$  is to specify a complex domain for  $(t_0, y_0)$  in which we can ensure that the map is well defined and in which we can use the estimates on the impact times  $\tau^+$  and  $\tau^-$  provided by Lemmas 4 and 5, respectively. This is stated precisely in Proposition 2 and below we give the details of the proof.

With the same notations and hypotheses of Lemma 4, we select quantities below to apply Lemmas 4 and 5, respectively.

- We set  $y_0^* = 3$ ,  $\bar{\rho}_0 = \rho$ ,  $\bar{\rho}_0^+ = \rho/8$ ,  $\bar{\rho}_0^- = \rho/48$  and  $\hat{\rho}_0^+ = \rho/192$  in Lemma 4. Then, the value  $\varepsilon_*^+ > 0$  on the statement of the lemma is defined by the conditions  $|a_0 \varepsilon| \leq 1/2$  and  $288\varepsilon\bar{p} < 19\rho/24$ . Consequently, if  $C_*^+ = C_*^+(\bar{p}, \hat{\rho}_0^+) > 0$  is the constant provided by Lemma 4, the following estimates hold if  $(t_0, y_0) \in \mathcal{D}_0^+ := \mathcal{D}^+(4, 11\rho/96, \rho/96)$ ,  $0 < \varepsilon \leq \varepsilon_*^+$ , and  $0 \leq i + j \leq 2$ :

$$\frac{|y_0|}{2} \leq |\tau_0^+(t_0, y_0; \varepsilon)| \leq 6|y_0|, \quad |\tau_*^+(t_0, y_0; \varepsilon)| \leq \frac{32\bar{p}}{|y_0|}, \quad \left| \partial_{t_0}^i \partial_{y_0}^j \tau_*^+(t_0, y_0; \varepsilon) \right| \leq \frac{C_*^+}{|y_0|}.$$

We observe that in this case  $\mathcal{D}_0^+ \subset \mathcal{D}^+(y_0^* + 2\hat{\rho}_0^+, \bar{\rho}_0^+ - 2\hat{\rho}_0^+, \bar{\rho}_0^+ - 2\hat{\rho}_0^+)$ .

- We set  $y_1^* = 1$ ,  $\bar{\rho}_1 = \rho$ ,  $\bar{\rho}_1^- = \rho/3$ ,  $\bar{\rho}_1^+ = \rho/12$  and  $\hat{\rho}_1^- = \rho/48$  in Lemma 5. Then, the value  $\varepsilon_*^- > 0$  on the statement of the lemma is defined by the conditions  $|a_0 \varepsilon| \leq 1/2$  and  $288\varepsilon\bar{p} < \rho/3$ . Consequently, if  $C_*^- = C_*^-(\bar{p}, \hat{\rho}_1^-) > 0$  is the constant provided by Lemma 5, the following estimates hold if  $(t_1, y_1) \in \mathcal{D}_0^- := \mathcal{D}^-(2, 7\rho/24, \rho/24)$ ,  $0 < \varepsilon \leq \varepsilon_*^-$ , and  $0 \leq i + j \leq 2$ :

$$\frac{|y_1|}{2} \leq |\tau_0^-(t_1, y_1; \varepsilon)| \leq 6|y_1|, \quad |\tau_*^-(t_1, y_1; \varepsilon)| \leq \frac{32\bar{p}}{|y_1|}, \quad \left| \partial_{t_1}^i \partial_{y_1}^j \tau_*^-(t_1, y_1; \varepsilon) \right| \leq \frac{C_*^-}{|y_1|}.$$

Now, what we observe is that  $\mathcal{D}_0^- \subset \mathcal{D}^-(y_1^* + 2\hat{\rho}_1^-, \bar{\rho}_1^- - 2\hat{\rho}_1^-, \tilde{\rho}_1^- - 2\hat{\rho}_1^-)$ .

We set  $\bar{\rho} = 11\rho/96$  and  $\tilde{\rho} = \rho/96$  as the values in the statement of Proposition 2, so that we have  $\mathcal{D}^+(4, \bar{\rho}, \tilde{\rho}) = \mathcal{D}_0^+$ . The value  $\varepsilon = \varepsilon_{\mathcal{P}}^*$  of the proposition is defined by the conditions  $|a_0\varepsilon| \leq 1/2$  and  $864\varepsilon\tilde{\rho} < \rho$  which, in particular, guarantee that  $\varepsilon_{\mathcal{P}}^* < \min\{\varepsilon_*^+, \varepsilon_*^-\}$ . For any  $(t_0, y_0) \in \mathcal{D}_0^+$  and  $0 < \varepsilon < \varepsilon_{\mathcal{P}}^*$ , we have (we are using (50), (51), estimates above provided by Lemmas 4 and 5 and the standard trick of bounding the size of the imaginary part of a complex number by its modulus when needed):

$$\begin{aligned} |\operatorname{Im}(t_1(t_0, y_0; \varepsilon))| &\leq |\operatorname{Im}(t_0)| + 4(|\operatorname{Im}(y_0)| + \varepsilon\|\tilde{P}_1\|_\rho) + \frac{32\tilde{\rho}\varepsilon}{|y_0|} < \frac{15}{96}\rho + 12\tilde{\rho}\varepsilon \leq \frac{7}{24}\rho, \\ |\operatorname{Im}(y_1(t_0, y_0; \varepsilon))| &\leq |\operatorname{Im}(y_0)| + \varepsilon\left(2\|\tilde{P}_1\|_\rho + \frac{3}{2}\frac{32\tilde{\rho}\varepsilon}{|y_0|}\right) < \frac{\rho}{96} + 14\tilde{\rho}\varepsilon \leq \frac{\rho}{24}, \end{aligned}$$

where we are bounding  $|\tilde{P}_1(t_1(t_0, y_0; \varepsilon))| \leq \|\tilde{P}_1\|_\rho$ , since the first computation guarantees that  $|\operatorname{Im}(t_1(t_0, y_0; \varepsilon))| < \rho$ . Moreover:

$$|y_1(t_0, y_0; \varepsilon)| \geq |y_0| - \varepsilon\left(2\|\tilde{P}_1\|_\rho + \frac{3}{2}\frac{32\tilde{\rho}\varepsilon}{|y_0|}\right) > |y_0|\left(1 - \frac{7}{2}\tilde{\rho}\varepsilon\right) \geq \frac{|y_0|}{2}$$

and, in particular,  $|y_1(t_0, y_0; \varepsilon)| \geq 2$ . These estimates guarantee that  $(t_1(t_0, y_0; \varepsilon), y_1(t_0, y_0; \varepsilon)) \in \mathcal{D}_0^-$ . In addition, we also consider (52) (recall that we are denoting  $t_1 = t_1(t_0, y_0; \varepsilon)$  and  $y_1 = y_1(t_0, y_0; \varepsilon)$ ):

$$|\operatorname{Im}(\bar{t}_0(t_0, y_0; \varepsilon))| \leq |\operatorname{Im}(t_0)| + \frac{16}{3}|\operatorname{Im}(y_0)| + \varepsilon\left(\frac{16}{3}\tilde{\rho} + 7\frac{32\tilde{\rho}}{|y_0|} + \frac{32\tilde{\rho}}{|y_1|}\right) < \frac{49}{288}\rho + \frac{232}{3}\tilde{\rho}\varepsilon \leq \rho.$$

Consequently, all the expressions and compositions of functions involved in formulas (54) and (55) are well defined within the selected ranges for  $t_0, y_0$  and  $\varepsilon$ . So, the map  $\mathcal{P}_\varepsilon(t_0, y_0; \varepsilon)$  is analytic in  $\mathcal{D}_0^+ = \mathcal{D}^+(4, \bar{\rho}, \tilde{\rho})$ , for any  $0 < \varepsilon < \varepsilon_{\mathcal{P}}^*$ .

To finish the proof of Proposition 2, it only remains to establish the existence of a constant  $\tilde{C} = \tilde{C}(a_0, \rho, \tilde{\rho})$  for which the bounds on the statement are fulfilled. To do this, we proceed analogously as we have done to define  $c_1^*$  in the proof of the Iterative Lemma 3 of the KAM part, i.e., we are going to redefine recursively an increasing value for  $\tilde{C}$  in order to meet a finite list of bounds. The value of  $\tilde{C}$  in the statement of the proposition is the final value for  $\tilde{C}$  thus obtained. We have checked that all the compositions involved in the definition of  $\mathcal{P}_\varepsilon$  by means of formulas from (50) to (55) are well defined. Therefore, all the bounds we have previously established throughout the proof of Proposition 2 apply to them. Consequently, we can bound all the objects involved by computing their derivatives by the chain rule and then applying such bounds to them. As an example, by computing the derivatives of  $t_1(t_0, y_0; \varepsilon)$  from (50), we obtain the following expressions (we omit most of the dependencies in  $t_0, y_0$  and  $\varepsilon$ ):

$$\begin{aligned} \partial_{t_0} t_1 &= 1 - \frac{\varepsilon}{1 - a_0\varepsilon} \tilde{P}_1'(t_0) + \varepsilon \partial_{t_0} \tau_*^+, \quad \partial_{y_0} t_1 = \frac{2}{1 - a_0\varepsilon} + \varepsilon \partial_{y_0} \tau_*^+, \\ \partial_{t_0}^2 t_1 &= -\frac{\varepsilon}{1 - a_0\varepsilon} \tilde{P}_1''(t_0) + \varepsilon \partial_{t_0}^2 \tau_*^+, \quad \partial_{t_0} \partial_{y_0} t_1 = \varepsilon \partial_{t_0} \partial_{y_0} \tau_*^+ \quad \partial_{y_0}^2 t_1 = \varepsilon \partial_{y_0}^2 \tau_*^+. \end{aligned}$$

Hence, we can easily establish the following estimates for them if  $(t_0, y_0) \in \mathcal{D}_0^+$  and  $0 < \varepsilon < \varepsilon_{\mathcal{P}}^*$  (for certain constant  $\tilde{C}$  defined recursively):

$$|\partial_{t_0} t_1| \leq \tilde{C}, \quad |\partial_{y_0} t_1| \leq \tilde{C}, \quad |\partial_{t_0}^2 t_1| \leq \tilde{C}\varepsilon, \quad |\partial_{t_0} \partial_{y_0} t_1| \leq \frac{\tilde{C}\varepsilon}{|y_0|}, \quad |\partial_{y_0}^2 t_1| \leq \frac{\tilde{C}\varepsilon}{|y_0|}.$$

By proceeding similarly with formula (51), we obtain analogous bounds for the partial derivatives of  $y_1 = y_1(t_0, y_0; \varepsilon)$ , just by replacing above  $t_1$  by  $y_1$ .

Finally, we deal with the first and second order derivatives with respect to  $t_0$  and  $y_0$  of  $f_{t_0}$  (see (54)),  $\bar{t}_0$  (see (20) and (53)), and  $f_{y_0}$  (see (55)). We express them in terms of the derivatives of  $t_1$ ,  $y_1$ ,  $\tau_*^+$ , and  $\tau_*^-$ , and bound them using all the estimates obtained previously throughout the proof of the proposition (in particular, we recall that  $|y_1| \geq |y_0|/2$ ). Then, we obtain the following estimates in the considered domains for  $t_0$ ,  $y_0$  and  $\varepsilon$ :

$$\begin{aligned} |f_{t_0}| &\leq \tilde{C}, \quad |\partial_{t_0} f_{t_0}| \leq \tilde{C}, \quad |\partial_{y_0} f_{t_0}| \leq \frac{\tilde{C}}{|y_0|}, \quad |\partial_{t_0}^2 f_{t_0}| \leq \tilde{C}, \quad |\partial_{t_0} \partial_{y_0} f_{t_0}| \leq \frac{\tilde{C}}{|y_0|}, \quad |\partial_{y_0}^2 f_{t_0}| \leq \frac{\tilde{C}}{|y_0|}, \\ |\partial_{t_0} \bar{t}_0| &\leq \tilde{C}, \quad |\partial_{y_0} \bar{t}_0| \leq \tilde{C}, \quad |\partial_{t_0}^2 \bar{t}_0| \leq \tilde{C}\varepsilon, \quad |\partial_{t_0} \partial_{y_0} \bar{t}_0| \leq \frac{\tilde{C}\varepsilon}{|y_0|}, \quad |\partial_{y_0}^2 \bar{t}_0| \leq \frac{\tilde{C}\varepsilon}{|y_0|}, \\ |f_{y_0}| &\leq \tilde{C}, \quad |\partial_{t_0} f_{y_0}| \leq \tilde{C}, \quad |\partial_{y_0} f_{y_0}| \leq \tilde{C}, \quad |\partial_{t_0}^2 f_{y_0}| \leq \tilde{C}, \quad |\partial_{t_0} \partial_{y_0} f_{y_0}| \leq \tilde{C}, \quad |\partial_{y_0}^2 f_{y_0}| \leq \tilde{C}. \end{aligned}$$

We omit further details because they are not difficult, but cumbersome.

### 5. THE LOCALIZED IMPACT MAP: PROOF OF PROPOSITION 3

As noted above, many dependencies on  $\varepsilon$  are omitted as it is set fixed throughout the proof. **Exactness of  $F$  with respect to  $\text{Id}\phi$ .** We first show that the map  $F(\phi, I) = (\bar{\psi}^{-1} \circ \bar{\mathcal{P}}_\varepsilon \circ \bar{\psi})(\phi, I)$  generated by applying the conjugation  $\bar{\psi}$  of (29) to the exact symplectic map  $\bar{\mathcal{P}}_\varepsilon$  of (21) is still exact symplectic. We know, by construction, that there is a function  $\bar{V}(t_0, E_0)$ , depending  $2\pi$ -periodically in  $t_0$ , such that  $\bar{\mathcal{P}}_\varepsilon^*(E_0 dt_0) = E_0 dt_0 + d\bar{V}(t_0, E_0)$ . Consequently, if we define  $V(\phi, I) = \frac{1}{\sqrt{-E_0^*}} \bar{V}(\phi, \sqrt{-E_0^*} I)$ , then we have:

$$\begin{aligned} F^*(\text{Id}\phi) &= \bar{\psi}^* \bar{\mathcal{P}}_\varepsilon^* (\bar{\psi}^{-1})^* (\text{Id}\phi) = \bar{\psi}^* \bar{\mathcal{P}}_\varepsilon^* \left( \frac{1}{\sqrt{-E_0^*}} E_0 dt_0 \right) \\ &= \frac{1}{\sqrt{-E_0^*}} \bar{\psi}^* (E_0 dt_0 + d\bar{V}(t_0, E_0)) = \text{Id}\phi + \bar{V}(\phi, I). \end{aligned}$$

**Explicit expression of  $F$ .** Since Proposition 2 gives us a close control over the properties of the map  $\mathcal{P}_\varepsilon(t_0, y_0)$  of (20), it is natural to express the map  $F$  of (31) in terms of it. Given a couple  $(\phi, I)$ , we denote  $(\bar{t}_0, \bar{y}_0) = \mathcal{P}_\varepsilon(t_0, y_0)$ , and  $(\bar{\phi}, \bar{I}) = F(\phi, I)$ . Then, we want to express  $(\bar{\phi}, \bar{I})$  in terms of  $\phi$ ,  $I$ ,  $\alpha_\varepsilon$ ,  $f_{t_0}$ , and  $f_{y_0}$ . Here, we are only concerned with the formal aspects of these calculations, so we do not discuss in which domain the considered expressions are well defined. Using the relations  $t_0 = \phi$ ,  $y_0 = \sqrt{-\sqrt{2} y_0^* I}$  (see (30)),  $\bar{\phi} = \bar{t}_0$ , and  $\bar{I} = -\frac{(\bar{y}_0)^2}{\sqrt{2} y_0^*}$ , we have:

$$\bar{\phi} = \bar{t}_0 = t_0 + \alpha_\varepsilon(y_0) + \varepsilon f_{t_0}(t_0, y_0; \varepsilon) = \phi + \alpha_\varepsilon \left( \sqrt{-\sqrt{2} y_0^* I} \right) + \varepsilon f_{t_0} \left( \phi, \sqrt{-\sqrt{2} y_0^* I}; \varepsilon \right).$$

Consequently,  $\alpha(I)$  is as given in the statement of Proposition 3 and

$$(56) \quad f_\phi(\phi, I) = \varepsilon f_{t_0} \left( \phi, \sqrt{-\sqrt{2} y_0^* I}; \varepsilon \right).$$

Furthermore,

$$\begin{aligned} \bar{I} &= -\frac{(\bar{y}_0)^2}{\sqrt{2} y_0^*} = -\frac{1}{\sqrt{2} y_0^*} (y_0 + \varepsilon f_{y_0}(t_0, y_0; \varepsilon))^2 \\ &= -\frac{1}{\sqrt{2} y_0^*} \left( y_0^2 + 2\varepsilon y_0 f_{y_0}(t_0, y_0; \varepsilon) + \varepsilon^2 (f_{y_0}(t_0, y_0; \varepsilon))^2 \right) \end{aligned}$$

$$= I - \frac{\varepsilon}{\sqrt{2}y_0^*} \left( 2y_0 f_{y_0}(t_0, y_0; \varepsilon) + \varepsilon (f_{y_0}(t_0, y_0; \varepsilon))^2 \right).$$

Consequently,

$$(57) \quad f_I(\phi, I) = -\frac{\varepsilon}{\sqrt{2}y_0^*} \left( 2\sqrt{-\sqrt{2}y_0^* I} f_{y_0}(\phi, \sqrt{-\sqrt{2}y_0^* I}; \varepsilon) + \varepsilon \left( f_{y_0}(\phi, \sqrt{-\sqrt{2}y_0^* I}; \varepsilon) \right)^2 \right).$$

**Bounding the action of  $\bar{\psi}$ .** As noted above, we can express the conjugation  $\bar{\psi}$  of (29) as  $t_0 = \phi$  and  $y_0 = \sqrt{-\sqrt{2}y_0^* I}$ . Now, we want to show that if  $(\phi, I) \in \Delta(\bar{\rho}) \times D(I_0^*, \bar{\rho})$  (see (15) and (25)), then  $(t_0, y_0) \in \mathcal{D}^+(4, \bar{\rho}, \bar{\rho})$  (see (16)) by establishing some quantitative bounds on the obtained value for  $y_0$ . Clearly, for  $t_0$  this assertion is obvious. Now let us denote  $I = I_0^* + \Delta I$ , with  $y_0^* = -\sqrt{2}I^*$  and  $|\Delta I| < \bar{\rho}$ , which means that  $I \in D(I_0^*, \bar{\rho})$ . Then, we have:

$$y_0 = y_0^* \sqrt{1 - \frac{\sqrt{2}\Delta I}{y_0^*}} = y_0^* \sum_{j=0}^{\infty} (-1)^j \binom{1/2}{j} \left( \frac{\sqrt{2}\Delta I}{y_0^*} \right)^j.$$

Since  $y_0^* > 5$  and  $\bar{\rho} < 1$ , we clearly have that  $\frac{\sqrt{2}|\Delta I|}{y_0^*} < \frac{1}{2}$ , meaning that the series above is well defined. Then, it follows :

$$\begin{aligned} |y_0 - y_0^*| &\leq y_0^* \sum_{j=1}^{\infty} \left| \binom{1/2}{j} \right| \left( \frac{\sqrt{2}|\Delta I|}{y_0^*} \right)^j = y_0^* \left( 1 - \sum_{j=0}^{\infty} \left| \binom{1/2}{j} \right| \left( \frac{\sqrt{2}|\Delta I|}{y_0^*} \right)^j \right) \\ &= y_0^* \left( 1 - \sqrt{1 - \frac{\sqrt{2}|\Delta I|}{y_0^*}} \right) \leq |\Delta I| < \bar{\rho}, \end{aligned}$$

where we have used that  $\left| \binom{1/2}{j} \right| = (-1)^{j-1} \binom{1/2}{j}$  and that by the Mean Value Theorem we can bound  $|1 - \sqrt{1-x}| \leq \frac{\sqrt{2}|x|}{2}$  if  $|x| \leq 1/2$ . In particular, using that  $y_0^*$  is a real number and bounding the size of the imaginary part of  $y_0 - y_0^*$  by its modulus, we conclude that

$$|\operatorname{Im}(y_0)| = |\operatorname{Im}(y_0 - y_0^*)| \leq |y_0 - y_0^*| < \bar{\rho},$$

which means that  $\bar{\psi}(\Delta(\bar{\rho}) \times D(I_0^*, \bar{\rho})) \subset \mathcal{D}^+(4, \bar{\rho}, \bar{\rho})$ , as wanted. Moreover, we also observe that

$$|y_0| \geq y_0^* - |y_0 - y_0^*| > y_0^* - \bar{\rho} > 4, \quad |y_0| \leq y_0^* + \bar{\rho}.$$

Using these bounds, recalling the explicit relation between  $y_0$  and  $I$ , given by  $y_0 = y_0(I) = \sqrt{-\sqrt{2}y_0^* I}$ , and computing the derivatives  $\partial_I y_0 = -\frac{\sqrt{2}}{2} \frac{y_0^*}{y_0}$  and  $\partial_I^2 y_0 = -\frac{(y_0^*)^2}{2y_0^3}$ , one easily obtains the estimates:

$$(58) \quad \frac{\sqrt{2}}{2} \frac{y_0^*}{y_0^* + \bar{\rho}} \leq |\partial_I y_0| \leq \frac{\sqrt{2}}{2} \frac{y_0^*}{y_0^* - \bar{\rho}}, \quad |\partial_I^2 y_0| \leq \frac{(y_0^*)^2}{2(y_0^* - \bar{\rho})^3}, \quad \forall I \in D(I_0^*, \bar{\rho}).$$

**Bounding  $\alpha'(I)$ ,  $\alpha''(I)$ ,  $f_\phi$  and  $f_I$ .** The first consequence of (58) are the bounds on  $\alpha'(I)$  and  $\alpha''(I)$  in the statement. Actually, we use that  $\alpha(I) = 4(1 - a_0^2 \varepsilon^2)^{-1} y_0(I)$ . Then, the bounds  $\frac{5\sqrt{2}}{3} < |\alpha'(I)| < \frac{10\sqrt{2}}{3}$  and  $|\alpha''(I)| \leq \frac{25}{24}$ ,  $\forall I \in D(I_0^*, \bar{\rho})$ , are straightforward from (58), by observing that since  $|a_0 \varepsilon| < 1/2$  (see Proposition 2) then  $1 < (1 - a_0^2 \varepsilon^2)^{-1} < 4/3$ , and recalling that  $y_0^* > 5$  and  $\bar{\rho} \in (0, 1)$ .

As we have shown that  $\bar{\psi}(\Delta(\bar{\rho}) \times D(I_0^*, \bar{\rho})) \subset \mathcal{D}^+(4, \bar{\rho}, \bar{\rho})$ , then Proposition 2 allows us to conclude that expressions for  $f_\phi$  and  $f_I$  in (56) and (57) are well defined and analytic if

$(\phi, I) \in \Delta(\bar{\rho}) \times D(I_0^*, \bar{\rho})$ , provided that  $0 < \varepsilon < \varepsilon_{\mathcal{P}}^*$  and  $y_0^* > 5$ . Expressions in (56) and (57) also allow us to obtain explicit formulas for the first and second order derivatives of  $f_\phi$  and  $f_I$ , with respect to  $\phi$  and  $I$ , in terms of the derivatives of  $f_{t_0}$ ,  $f_{y_0}$  and  $y_0(I) = \sqrt{-\sqrt{2} y_0^* I}$ . Then, by considering bounds above on  $y_0(I)$  and its derivatives as well those on  $f_{t_0}$ ,  $f_{y_0}$  and its derivatives in  $\mathcal{D}^+(4, \bar{\rho}, \bar{\rho})$  provided by Proposition 2, it is not difficult to conclude the existence of a constant  $\bar{C}$  for which hold all the bounds on the statement on  $f_\phi$ ,  $f_I$  and its derivatives.

## 6. EXACT SYMPLECTIC STRUCTURE: PROOF OF PROPOSITION 1

It is well known that that the time- $T$ -flow of a Hamiltonian system, the Poincaré map originating from an autonomous Hamiltonian system (at a fixed energy level) and the stroboscopic map associated to a periodic in time Hamiltonian system are exact symplectic (see e.g. [26]). Thus, we dedicate this section to the proof of Proposition 1, which basically shows that the impact map of a smooth periodic Hamiltonian system has the exact symplectic character if it is expressed using time-energy variables.

*Proof.* Let  $\mathcal{H}(x, y)$  be an autonomous Hamiltonian with  $n$  degrees of freedom with respect to the 2-form  $dx \wedge dy$ . We denote by  $\varphi^T(x_0, y_0) = (x^\tau(x_0, y_0), y^\tau(x_0, y_0))$  the flow associated to the system originating from  $\mathcal{H}$ , where  $\tau$  represents the time variable. It is a known result in the literature (see e.g. [26]) that the time- $T$ -map, for every  $T \in \mathbb{R}$ , is exact symplectic. Specifically,

$$x^T dy^T = (\varphi^T)^*(x_0 dy_0) = x_0 dy_0 + d\mathcal{S}^T(x_0, y_0),$$

with  $\varphi^T$  the time- $T$ -flow of  $\mathcal{H}$  and with  $\mathcal{S}^T$  being given by

$$\mathcal{S}^T(x_0, y_0) = \int_0^T [\langle x^s(x_0, y_0), \partial_\tau y^s(x_0, y_0) \rangle + \mathcal{H}(\varphi^s(x_0, y_0))] ds.$$

If set  $T = T(x_0, y_0)$  to depend on the initial conditions, we easily get the relation

$$(59) \quad x^{T(x_0, y_0)} d(y^{T(x_0, y_0)}) = x_0 dy_0 + d(\mathcal{S}^{T(x_0, y_0)}(x_0, y_0)) - \mathcal{H}(\varphi^{T(x_0, y_0)}(x_0, y_0)) d(T(x_0, y_0)).$$

We then apply the construction above to the 2-degrees of freedom autonomized Hamiltonian  $\mathcal{H}(x, t, y, E) = H(x, y, t) + E$  of the statement of the proposition, being  $x = (x, t)$  and  $y = (y, E)$ , and we select  $T(x_0, t_0, y_0, E_0) > 0$  as the impact time defined as  $\tilde{\tau}(t_0, E_0)$  in (12), but now for each initial condition  $(x_0, t_0, y_0, E_0)$  close to the point  $(0, t_0, y(t_0, E_0), E_0)$  of the section  $\Xi$ . Hence, it verifies:

$$x^{T(x_0, t_0, y_0, E_0)}(x_0, t_0, y_0, E_0) = 0.$$

We define the map

$$\begin{aligned} \Phi : U &\longrightarrow \Xi \\ (t_0, E_0) &\longmapsto (0, t_0, y(t_0, E_0), E_0), \end{aligned}$$

and perform the pull-back  $\Phi^*$  on equation (59) written in the context at hand. Using that  $\tilde{\tau}(t_0, E_0) = (T \circ \Phi)(t_0, E_0)$ , the definition of the impact map  $(t_1, E_1) = F(t_0, E_0)$ , that  $\mathcal{H} = 0$  along the solutions starting at  $\Xi$  and that the coordinate  $x$  vanishes at booth the points  $\Phi(t_0, E_0)$  and the point obtained after integrating time  $\tilde{\tau}(t_0, E_0)$  the point  $\Phi(t_0, E_0)$ , we conclude that

$$F^*(t_1 dE_1) = t_0 dE_0 + d\bar{\mathcal{S}}(t_0, E_0),$$

where we are introducing

$$\bar{\mathcal{S}}(t_0, E_0) = \mathcal{S}^{\tilde{\tau}(t_0, E_0)}(0, t_0, y(t_0, E_0), y_0).$$

By using the relation  $d(t_1 \cdot E_1) = t_1 dE_1 + E_1 dt_1$ , we have:

$$\begin{aligned} F^*(E_1 dt_1) &= d(F^*(t_1 \cdot E_1)) - F^*(t_1 dE_1) = d(F^*(t_1 \cdot E_1)) - t_0 dE_0 - d\tilde{S}(t_0, E_0) \\ &= E_0 dt_0 + dS(t_0, E_0), \end{aligned}$$

where

$$S(t_0, E_0) = F^*(t_1 \cdot E_1) - t_0 \cdot E_0 - \tilde{S}(t_0, E_0).$$

We observe that using that  $t^s(\Phi(t_0, E_0)) = t_0 + s$  but that  $x^s(\Phi(t_0, E_0))$ ,  $y^s(\Phi(t_0, E_0))$ , and  $E^s(\Phi(t_0, E_0))$  depend on  $t_0$  in a  $2\pi$ -periodic way, we can write

$$\tilde{S}(t_0, E_0) = \int_0^{\tilde{\tau}(t_0, E_0)} t_0 \cdot \partial_\tau E^s(\Phi(t_0, E_0)) ds + \tilde{S}(t_0, E_0) = t_0 \cdot (f_{E_0}(t_0, E_0) - E_0) + \tilde{S}(t_0, E_0),$$

where  $\tilde{S}(t_0, E_0)$  is  $2\pi$ -periodic in  $t_0$ . Then, using that  $F^*(t_1 \cdot E_1) = (t_0 + f_{t_0}(t_0, E_0)) \cdot f_{E_0}(t_0, E_0)$ , we conclude that

$$S(t_0, E_0) = f_{t_0}(t_0, E_0) \cdot f_{E_0}(t_0, E_0) - \tilde{S}(t_0, E_0),$$

is  $2\pi$ -periodic in  $t_0$ , as claimed in the statement.  $\square$

#### COMPETING INTERESTS

To the best of our knowledge, no conflict of interest, financial or other, exists.

#### AUTHORS' CONTRIBUTIONS

All persons who meet authorship criteria are listed as authors, and all authors certify that they have participated sufficiently in the work to take public responsibility for the content, including participation in the conceptualization, methodology, formal analysis, investigation, writing-original draft preparation and writing-review & editing.

#### ACKNOWLEDGMENTS

T. M-Seara and J. Villanueva are supported by the grant AGRUPS-2023: Grup de Sistemes Dinàmics de la UPC. T. M-Seara is supported by the grant PID-2021-122954NB-I00 funded by MCIN/AEI/10.13039/501100011033 and "ERDF A way of making Europe", the Catalan Institution for Research and Advanced Studies (via an ICREA Academia Prize 2019) and the Spanish State Research Agency, through the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M). LVMFS is partially supported by São Paulo Research Foundation (FAPESP) grants 2018/22398-0 and 2021/11515-8.

#### REFERENCES

- [1] L. Burra and F. Zanolin. Chaos in a periodically perturbed second-order equation with signum nonlinearity. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 30(2):2050031, 9, 2020.
- [2] R. de la Llave, A. González, A. Jorba, and J. Villanueva. KAM theory without action-angle variables. *Nonlinearity*, 18(2):855–895, 2005.
- [3] R. Dieckerhoff and E. Zehnder. Boundedness of solutions via the twist-theorem. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 14(1):79–95, 1987.
- [4] R. Enguiça and R. Ortega. Functions with average and bounded motions of a forced discontinuous oscillator. *J. Dynam. Differential Equations*, 31(3):1185–1198, 2019.
- [5] A. F. Filippov. *Differential equations with discontinuous righthand sides*, volume 18 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1988. Translated from the Russian.
- [6] A. Haro, M. Canadell, J.-L. Figueras, A. Luque, and J.-M. Mondelo. *The parameterization method for invariant manifolds*, volume 195 of *Applied Mathematical Sciences*. Springer, [Cham], 2016. From rigorous results to effective computations.



- [7] A. Jacquemard and M. Teixeira. Periodic solutions of a class of non-autonomous second order differential equations with discontinuous right-hand side. *Physica D: Nonlinear Phenomena*, 241(22):2003–2009, 2012. Dynamics and Bifurcations of Nonsmooth Systems.
- [8] A. Jorba and J. Villanueva. On the normal behaviour of partially elliptic lower-dimensional tori of Hamiltonian systems. *Nonlinearity*, 10(4):783–822, 1997.
- [9] A. Jorba and J. Villanueva. On the persistence of lower-dimensional invariant tori under quasi-periodic perturbations. *J. Nonlinear Sci.*, 7(5):427–473, 1997.
- [10] M. Kunze, T. Küpper, and J. You. On the application of KAM theory to discontinuous dynamical systems. *J. Differential Equations*, 139(1):1–21, 1997.
- [11] M. Levi. Quasiperiodic motions in superquadratic time-periodic potentials. *Comm. Math. Phys.*, 143(1):43–83, 1991.
- [12] X. Li. Boundedness of solutions for Duffing’s equations with semilinear potentials. *J. Differential Equations*, 176(1):248–268, 2001.
- [13] J. E. Littlewood. Unbounded solutions of an equation  $\ddot{y} + g(y) = p(t)$ , with  $p(t)$  periodic and bounded, and  $g(y)/y \rightarrow \infty$  as  $y \rightarrow \pm\infty$ . *J. London Math. Soc.*, 41:497–507, 1966.
- [14] J. E. Littlewood. Unbounded solutions of  $\ddot{y} + g(y) = p(t)$ . *J. London Math. Soc.*, 41:491–496, 1966.
- [15] J. E. Littlewood. *Some problems in real and complex analysis*. D. C. Heath and Company Raytheon Education Company, Lexington, Mass., 1968.
- [16] B. Liu and J. J. Song. Invariant curves of reversible mappings with small twist. *Acta Math. Sin. (Engl. Ser.)*, 20(1):15–24, 2004.
- [17] G. R. Morris. A case of boundedness in Littlewood’s problem on oscillatory differential equations. *Bull. Austral. Math. Soc.*, 14(1):71–93, 1976.
- [18] J. Moser. On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Gottingen Math.-Phys. Kl. II*, 1962:1–20, 1962.
- [19] D. D. Novaes and L. V. M. F. Silva. Invariant tori and boundedness of solutions of non-smooth oscillators with Lebesgue-integrable forcing term. *Zeitschrift für angewandte Mathematik und Physik*, 75(1):10, 2024.
- [20] D. D. Novaes and L. V. M. F. Silva. A Melnikov analysis on a family of second order discontinuous differential equations. *São Paulo Journal of Mathematical Sciences*, pages 1–19, 2024.
- [21] R. Ortega. A report on the boundedness for semilinear Duffing’s equation. Institute of Mathematics, Academia Sinica, 1998.
- [22] R. Ortega. Boundedness in a piecewise linear oscillator and a variant of the small twist theorem. *Proc. London Math. Soc.* (3), 79(2):381–413, 1999.
- [23] R. Ortega. Twist mappings, invariant curves and periodic differential equations. In *Nonlinear analysis and its applications to differential equations (Lisbon, 1998)*, volume 43 of *Progr. Nonlinear Differential Equations Appl.*, pages 85–112. Birkhäuser Boston, Boston, MA, 2001.
- [24] R. Ortega. Dynamics of a forced oscillator having an obstacle. In *Variational and topological methods in the study of nonlinear phenomena (Pisa, 2000)*, volume 49 of *Progr. Nonlinear Differential Equations Appl.*, pages 75–87. Birkhäuser Boston, Boston, MA, 2002.
- [25] D. Qian and X. Sun. Invariant tori for asymptotically linear impact oscillators. *Sci. China Ser. A*, 49(5):669–687, 2006.
- [26] D. Treschev and O. Zubelevich. *Introduction to the perturbation theory of Hamiltonian systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2010.
- [27] J. Villanueva. A parameterization method for Lagrangian tori of exact symplectic maps of  $\mathbb{R}^{2r}$ . *SIAM J. Appl. Dyn. Syst.*, 17(3):2289–2331, 2018.
- [28] Y. Wang. Boundedness of solutions in a class of Duffing equations with a bounded restore force. *Discrete Contin. Dyn. Syst.*, 14(4):783–800, 2006.
- [29] V. Zharnitsky. Quasiperiodic motion in the hamiltonian systems of the billiard type. *Phys. Rev. Lett.*, 81:4839–4842, Nov 1998.
- [30] V. Zharnitsky. Invariant tori in Hamiltonian systems with impacts. *Comm. Math. Phys.*, 211(2):289–302, 2000.

<sup>1</sup> DEPARTAMENT DE MATEMÀTIQUES, UNIVERSITAT POLITÈCNICA DE CATALUNYA, DIAGONAL 647, 08028 BARCELONA, SPAIN

<sup>2</sup> INSTITUT DE MATEMÀTIQUES DE LA UPC - BARCELONA TECH (IMTECH), PAU GARGALLO 14, 08028 BARCELONA, SPAIN.

<sup>3</sup> CENTRE DE RECERCA MATEMÀTICA, EDIFICI C, CAMPUS BELLATERRA, 08193 BELLATERRA, SPAIN.

<sup>4</sup> DEPARTAMENTO DE MATEMÁTICA, UNIVERSIDADE ESTADUAL DE CAMPINAS, RUA SÉRGIO BARUQUE DE HOLANDA, 651, CIDADE UNIVERSITÁRIA ZEFERINO VAZ, 13083-859, CAMPINAS, SP, BRAZIL