Exponentially small splitting of separatrices under fast quasiperiodic forcing

Amadeu Delshams[1], Vassili Gelfreich[2], Àngel Jorba[1], Tere M. Seara[1]

Revised version, January 1997

[1] Departament de Matemàtica Aplicada I
Universitat Politècnica de Catalunya
Diagonal 647, 08028 Barcelona, Spain
amadeu@ma1.upc.es, angel@tere.upc.es, tere@ma1.upc.es

[2] Departament de Matemàtica Aplicada i Anàlisi
Universitat de Barcelona
Gran via 585, 08007 Barcelona, Spain
and
Chair of Applied Mathematics
St.Petersburg Academy of Aerospace Instrumentation
Bolshaya Morskaya 67, 190000, St. Petersburg, Russia
gelf@maia.ub.es, gelf@misha.usr.saai.ru

Abstract

We consider fast quasiperiodic perturbations with two frequencies $(1/\varepsilon, \gamma/\varepsilon)$ of a pendulum, where γ is the golden mean number. The complete system has a two-dimensional invariant torus in a neighbourhood of the saddle point. We study the splitting of the three-dimensional invariant manifolds associated to this torus. Provided that the perturbation amplitude is small enough with respect to ε, and some of its Fourier coefficients (the ones associated to Fibonacci numbers), are separated from zero, it is proved that the invariant manifolds split and that the value of the splitting, which turns out to be exponentially small with respect to ε, is correctly predicted by the Melnikov function.

AMS classification scheme numbers: 34C37, 58F27, 58F36, 11J25

1 Introduction

At the end of the last century, H. Poincaré [Poi99] discovered the phenomenon of the splitting of separatrices, which seems to be the main cause of the stochastic behaviour in Hamiltonian systems. He formulated the *general problem of dynamics* as a perturbation of an integrable Hamiltonian

$$H(I,\varphi,\varepsilon) = H_0(I) + \varepsilon H_1(I,\varphi),$$

where ε is a small parameter, $I = (I_1, I_2, \dots, I_n)$, $\varphi = (\varphi_1, \varphi_2, \dots, \varphi_n)$. The values of the actions I, such that the unperturbed frequencies $\omega_k(I) = \partial H_0/\partial I_k$ are rationally dependent, are called *resonances.*

As a model for the motion near a resonance, Poincaré studied the pendulum with a high-frequency perturbation, which can be described by the Hamiltonian

$$\frac{y^2}{2} + \cos x + \mu \sin x \cos \frac{t}{\varepsilon}.$$

His calculations of the splitting, originally validated only for $|\mu|$ exponentially small with respect to ε, predicted correctly the splitting up to $|\mu| \leq \varepsilon^p$ for any positive parameter p [Gel97, Tre94]. The main problem in studying such kind of systems is that the splitting is exponentially small with respect to ε. Namely, Neishtadt's theorem [Nei84] implies that in a Hamiltonian of the form

$$H\left(x,y,t/\varepsilon\right) = H_0(x,y) + H_1\left(x,y,t/\varepsilon\right),$$

where the Hamiltonian system of H_0 has a saddle and an associated homoclinic orbit, and the perturbation H_1 is a periodic function of time with zero mean value, the splitting can be bounded from above by $O(e^{-\mathrm{const}/\varepsilon})$. For this estimate to be valid all the functions have to be real analytic in x and y, but C^1 dependence on time is sufficient. Lately, the constant in the exponent was related to the position of complex time singularities of the unperturbed homoclinic orbit [HMS88, Fon93, Fon95].

The above-mentioned systems provide a realistic model for the motion near a resonance only in the case of two degrees of freedom. If one considers simple resonances of systems with more than two degrees of freedom, one can choose all the angles except one to be fast variables.

The simplest case is a quasiperiodic perturbation of a planar Hamiltonian system. Neishtadt's averaging theorem was generalized to this case by C. Simó [Sim94], but the upper bounds provided for the splitting depend in an essential way on the frequency vector of the perturbation. For a perturbation of the pendulum depending on two frequencies, C. Simó [Sim94] checked numerically that a proper modification of the Melnikov method gives the correct prediction for the splitting.

Autonomous models with perturbations that depend on time in a quasiperiodic way appear in several problems of Celestial Mechanics. For instance, the motion of a spacecraft in the Earth–Moon system can be modeled assuming that Earth and Moon revolve in circles around their common centre of masses (this gives an autonomous model), and the main perturbations (difference between the circular and the real motion of the Moon, effect of the Sun, etc.) are modeled as a time-dependent quasiperiodic function. For more details, see [DJS91] or [GJMS91].

In the present paper we consider a quasiperiodic high-frequency perturbation of the pendulum, described by the Hamiltonian function

$$\frac{\omega \cdot I}{\varepsilon} + h(x, y, \theta, \varepsilon), \tag{1.1}$$

where

$$\omega \cdot I = \omega_1 I_1 + \omega_2 I_2, \qquad h(x, y, \theta, \varepsilon) = \frac{y^2}{2} + \cos x + \varepsilon^p m(\theta_1, \theta_2) \cos x,$$

with symplectic form $\mathrm{d}x \wedge \mathrm{d}y + \mathrm{d}\theta_1 \wedge \mathrm{d}I_1 + \mathrm{d}\theta_2 \wedge \mathrm{d}I_2$. We assume that ε is a small positive parameter and p is a positive parameter. Mainly due to a technical limitation imposed by the Extension Theorem (Theorem 3), we will restrict ourselves to the case $p > 3$. We also assume that the frequency is of the form ω/ε for

$$\omega = (1, \gamma), \qquad \gamma = \frac{\sqrt{5} + 1}{2}. \tag{1.2}$$

The number γ is the famous golden mean number, which is the "most irrational" number [Khi63, Lan91]. The equations of motion associated with Hamiltonian (1.1) are

$$\begin{aligned}
\dot{x} &= y, & \dot{y} &= (1 + \varepsilon^p m(\theta_1, \theta_2)) \sin x, \\
\dot{\theta}_1 &= \frac{1}{\varepsilon}, & \dot{I}_1 &= -\varepsilon^p \cos x \frac{\partial m}{\partial \theta_1}(\theta_1, \theta_2), \\
\dot{\theta}_2 &= \frac{\gamma}{\varepsilon}, & \dot{I}_2 &= -\varepsilon^p \cos x \frac{\partial m}{\partial \theta_2}(\theta_1, \theta_2).
\end{aligned} \tag{1.3}$$

Actions I_1 and I_2 have only been introduced to put the Hamiltonian in autonomous form, but are not relevant from a dynamical point of view (note that they do not appear in the right hand sides of the equations of motion).

The function m is assumed to be a 2π-periodic function of two variables θ_1 and θ_2. Thus, it can be represented as a Fourier series:

$$m(\theta_1, \theta_2) = \sum_{k_1, k_2} m_{k_1 k_2} e^{i(k_1 \theta_1 + k_2 \theta_2)}. \tag{1.4}$$

We assume that, for some positive numbers r_1 and r_2,

$$\sup_{k_1, k_2} \left| m_{k_1 k_2} e^{r_1 |k_1| + r_2 |k_2|} \right| < \infty, \tag{1.5}$$

and that there are positive numbers a and k_0, such that

$$|m_{k_1 k_2}| \geq a e^{-r_1 |k_1| - r_2 |k_2|}, \tag{1.6}$$

for all $|k_1|/|k_2|$, which are continuous fraction convergents of γ with $|k_2| \geq k_0$. In fact, k_1 and k_2 are consecutive Fibonacci numbers: $k_1 = \pm F_{n+1}$ and $k_2 = \mp F_n$. The Fibonacci numbers are defined by the recurrence: $F_0 = 1$, $F_1 = 1$, $F_{n+1} = F_n + F_{n-1}$ for $n \geq 1$. We call the corresponding terms in the perturbation to be *resonant* or *Fibonacci terms*.

For example, the function

$$m(\theta_1, \theta_2) = \frac{\cos\theta_1 \cos\theta_2}{(\cosh r_1 - \cos\theta_1)(\cosh r_2 - \cos\theta_2)}$$

satisfies these conditions.

The upper bound (1.5) implies that the function m is analytic on the strip $\{|\operatorname{Im}\theta_1| < r_1\} \times \{|\operatorname{Im}\theta_2| < r_2\}$. Equation (1.6) implies that this function can not be prolonged analytically onto a larger strip. Let us select $\alpha \in (0,1]$. Estimate (1.5) implies that

$$|m(\theta_1, \theta_2)| \leq K\varepsilon^{-2\alpha} \tag{1.7}$$

on the strip

$$|\operatorname{Im}\theta_1| \leq r_1 - \varepsilon^\alpha, \quad |\operatorname{Im}\theta_2| \leq r_2 - \varepsilon^\alpha. \tag{1.8}$$

Formula (1.6) implies that the upper bound (1.7) can not be improved. The value of the splitting depends essentially on the width of these strips. The function m under consideration has a singularity "of second order", in the sense that the upper bound (1.7) for the maximum of the modulus is quadratic with respect to the inverse of the distance to the boundary of the strip. In a similar way the case of a singularity of any "order" q can be considered. In this case $m_{k_1 k_2}$ should be replaced by $m_{k_1 k_2}/|k|^{q-2}$ in (1.5) and (1.6).

An example from [DGJS97b] shows that the Melnikov function and the splitting of separatrices can be of the order of some power of ε if the function m is not analytic, but has only a finite number of continuous derivatives. This makes a first qualitative difference between periodic and quasiperiodic perturbations. Indeed, in the periodic case, only the C^1 dependence with respect to θ of the perturbed Hamiltonian is needed to prove that the splitting is $O(e^{-\mathrm{const}/\varepsilon})$. (In both cases, the analyticity with respect to x, y is essential.)

The Hamiltonian (1.1) can be considered as a singular perturbation of the pendulum

$$h_0 = \frac{y^2}{2} + \cos x. \tag{1.9}$$

The unperturbed system has a saddle point $(0,0)$ and a homoclinic trajectory given by

$$x_0(t) = 4\arctan(e^t), \qquad y_0(t) = \dot{x}_0(t). \tag{1.10}$$

The complete system (1.3) has a whiskered torus $\mathcal{T}$: $(0, 0, \theta_1, \theta_2)$. The whiskers are $3D$-hypersurfaces in the $4D$-extended phase space $(x, y, \theta_1, \theta_2)$. These invariant manifolds are close to the unperturbed pendulum separatrix.

Our *main result* is that for $p > 3$ and small $\varepsilon > 0$ the invariant manifolds split, and that the value of the splitting is correctly predicted by the Melnikov function

$$M(\theta_1, \theta_2; \varepsilon) = \int_{-\infty}^{\infty} \{h_0, h\}(x_0(t), y_0(t), \theta_1 + t/\varepsilon, \theta_2 + \gamma t/\varepsilon, \varepsilon)\, dt. \tag{1.11}$$

To give a more precise statement, we need to introduce a $2\log\gamma$-periodic function c defined by

$$c(\delta) = C_0 \cosh\left(\frac{\delta - \delta_0}{2}\right) \quad \text{for} \quad \delta \in [\delta_0 - \log\gamma, \delta_0 + \log\gamma], \tag{1.12}$$

where

$$C_0 = \sqrt{\frac{2\pi(\gamma r_1 + r_2)}{\gamma + \gamma^{-1}}}, \qquad \delta_0 = \log\varepsilon^*, \quad \varepsilon^* = \frac{\pi(\gamma + \gamma^{-1})}{2\gamma^2(r_1\gamma + r_2)},$$

and continued by $2\log\gamma$-periodicity onto the whole real axis. The function c is piecewise-analytic and continuous.

Theorem 1 (Main Theorem) *Given positive constants $T_1 < T_2$, there exists a canonical coordinate system $(H, T, \theta_1, \theta_2)$, such that for $p > 3$, $T_1 \le T \le T_2$ and real θ_1 and θ_2, the stable manifold has the equation $H = 0$, and the unstable manifold can be represented as the graph of a function $H = H^u(T, \theta_1, \theta_2; \varepsilon)$, where the function H^u depends 2π-periodically on θ_1 and θ_2 and is close to the Melnikov function:*

$$|H^u(T, \theta_1, \theta_2; \varepsilon) - M(\theta_1 - T/\varepsilon, \theta_2 - \gamma T/\varepsilon; \varepsilon)| \le \mathrm{const}\, \varepsilon^{2p-4} \exp\left(-\frac{c(\log\varepsilon)}{\sqrt{\varepsilon}}\right), \tag{1.13}$$

with $c(\delta)$ as defined in (1.12). If condition (1.6) *is fulfilled, then there exists $\varepsilon_0 > 0$ such that, for $0 < \varepsilon < \varepsilon_0$, the maximum of the modulus of the Melnikov function is larger than the right hand side of (1.13).*

This result, which was already announced in [DGJS97a], is not trivial since the Melnikov function is exponentially small with respect to ε. As we will see, for a fixed small ε the resonant terms with k_1, $k_2 \sim \mathrm{const}/\sqrt{\varepsilon}$ are the ones that give the largest contribution to the Melnikov function. Condition (1.6) is not needed to get an upper bound for the Melnikov function of the form

$$\mathrm{const}\, \varepsilon^{p-1} \exp\left(-\frac{c(\log\varepsilon)}{\sqrt{\varepsilon}}\right), \tag{1.14}$$

as well as upper bound (1.13), which provide together an upper estimate for the splitting of separatrices.

We need condition (1.6) to ensure a *lower* bound for the maximum of the Melnikov function of the same form (1.14), which dominates the error bound (1.13) and gives rise to a lower estimate for the splitting of separatrices. If condition (1.6) is not satisfied, we only get an upper bound for the splitting. In particular, this happens if m is a trigonometric polynomial or an entire function. Nevertheless, if the inequality (1.6) is satisfied for a sufficiently large (but finite) number of consecutive resonant terms, we can still find small positive numbers $\varepsilon_0', \varepsilon_0 > 0$, such that the Melnikov function is greater than the error for small finite $\varepsilon \in (\varepsilon_0', \varepsilon_0)$. Then, from a practical point of view, the Melnikov theory gives a good approximation, but not an asymptotic formula.

It is remarkable that the exponent of the asymptotic expression (1.14) for the Melnikov function is different from the case of a periodic perturbation. There appears not only a different power of ε, but a periodic function $c(\log \varepsilon)$ instead of a constant.

In the case of an entire function m, we think that the method used in the present paper can be modified in order to improve the estimate of the error and to prove that the Melnikov function gives an actual asymptotic at least when the resonant terms decrease not faster than $1/k!$.

We note that the Melnikov function is not invariant with respect to canonical changes of variables. After a change, e.g. after a step of the classical averaging procedure, a lot of non-zero harmonics, which were not present in the original system, can appear. If in the original system the Fibonacci terms were not big enough, these new harmonics may give larger contribution to the splitting. This idea was used in [Sim94] to detect the splitting for a system with only 4 terms initially present.

The assumption (1.2) that ω in the frequency vector is just $(1, \gamma)$ can be relaxed. The generalization of the present result to the case when γ is a quadratic number is straightforward, with a similar expression (1.14) for the size of the Melnikov function. The case in which $\omega = (\omega_1, \omega_2)$, with the ratio ω_1/ω_2 being of constant type (the continued fraction expansion has bounded coefficients), but not quadratic, can be similarly analyzed, but in this case $c(\delta)$ is no longer a periodic function. In some sense one can say, properly speaking, that there are no asymptotics. But it seems that there still exist upper and lower bounds, with the factor $\sqrt{\varepsilon}$ in the denominator of the exponential term. The case of two frequencies whose ratio ω_1/ω_2 is not of constant type, as well as the case of more than two perturbing frequencies, is more complicated.

Our model is based on the paper [Sim94] by C. Simó, where a lot of semi-numerical computations were presented. It can be thought as an intermediate step between a Hamiltonian with one and a half degrees of freedom and a Hamiltonian with n degrees of freedom like the following generalization of Arnold's example

$$H(x, \phi, y, I, \varepsilon, \mu) = \frac{1}{2}y^2 + \frac{1}{2}I^2 + \varepsilon(\cos x - 1) + \mu F(x, \phi), \tag{1.15}$$

where $x \in \mathbb{T}$, $\phi \in \mathbb{T}^{n-1}$ are the coordinates, $y \in \mathbb{R}$ $I \in \mathbb{R}^{n-1}$ are the momenta, which was introduced by P. Lochak. It is remarkable that in his paper [Loc92, V§2], P. Lochak was already putting emphasis on perturbations F with arbitrarily high harmonics, in contrast with the original Arnold's example [Arn64], in order to get realistic estimates for the splitting of separatrices. A similar Hamiltonian (the fast rotator-pendulum model) was studied by L. Chierchia and G. Gallavotti [CG94], and by G. Gallavotti [Gal94], working to all orders of perturbation theory, and expressing the coefficients of the μ-th order contribution to the splitting of separatrices as improper time integrals from

$t = -\infty$ to $t = +\infty$. However, by using a perturbation F with finite harmonics, they only were able to get upper estimates for the splitting.

A related result is the Jeans–Landau–Teller approximation for adiabatic invariants, where the change of actions is given in first order by a sort of Melnikov function. Exponentially small upper bounds for the change of actions were obtained by G. Benettin, A. Carati and G. Gallavotti [BCG95], proving cancellations through tree–like diagrams. A numerical study performed by G. Benettin, A. Carati and F. Fassò [BCF96] for the case of a large asymptotic frequency $\lambda(1, \gamma)$ ($\lambda = 1/\varepsilon$ in our notation) with $\gamma = \sqrt{2}$, a quadratic number, shows a good agreement between the numerical values of the Melnikov function and the change of actions.

While we were revising this paper, we became aware of a new version of a remarkable preprint [RW97] by M. Rudnev and S. Wiggins, devoted also to the Hamiltonian (1.15). Assuming similar conditions to (1.5), (1.6) for an even perturbation F (in particular, F possesses arbitrarily high harmonics), they give exponentially small upper and lower bounds for the splitting of separatrices for $\mu = O(\varepsilon^p)$. It is important to notice that their results apply to $n \geq 3$ degrees of freedom. It is interesting to remark here that since we restrict ourselves to a more concrete model, we obtain more information about the limit behaviour of the Melnikov function and the splitting of separatrices, for a lower value of the exponent p.

The rest of the paper is devoted to the proof of the Main Theorem. In contrast with the above-mentioned papers, the method used in the present paper is based on the geometrical ideas proposed by Lazutkin [Laz84] for the study of the separatrix splitting for the standard map, and adapted to differential equations by the authors [Gel90, DS92, Gel93]. In section 2, the Melnikov function is carefully analyzed, to provide its asymptotic behaviour. In section 3, like in [DS92, DS96], and as a first step to give a description of the dynamics near the $2D$-dimensional hyperbolic invariant torus $\mathcal{T}$, it is established the existence of a convergent normal form. This result on the normal form theorem is similar to Moser's theorem [Mos56] on the normal form near a periodic hyperbolic orbit, and to [CG94] and [RW95] in more general situations. However, our proof (see section 8) is based on a quadratically convergent scheme, which allows us to show that the loss of domain in the phases θ is bounded by $\sqrt{\varepsilon}$, as required to obtain an asymptotic formula for the Melnikov function.

Besides, the Normal Form Theorem ensures that the local unstable manifold is $\mathrm{O}\left(\varepsilon^{p-1}\right)$-close to the unperturbed separatrix. In section 4, the Extension Theorem, proved in section 9, extends this local approximation for solutions of system (1.3) to a global one, on a suitably chosen complex domain. Since the unperturbed homoclinic orbit comes back to the domain of the normal form, the same happens to the unstable manifold, which can be compared with the local stable manifold. This comparison is performed in section 5, where Theorem 4 is proved, implying immediately the Main Theorem. Finally, section 6 is devoted to the arithmetic properties of the golden mean number γ, and in section 7 some analytic properties of the quasiperiodic functions are studied.

2 Melnikov Function

As is well-known, the Melnikov function (1.11) gives a first order approximation of the difference between the values of the unperturbed pendulum energy h_0 on the stable and unstable manifolds. Next lemma describes its main features.

Lemma 1 (Properties of the Melnikov function) *The Melnikov function defined by* (1.11) *is a* 2π*-periodic function of* θ_1 *and* θ_2*, such that*

1) $M(\theta_1 - T/\varepsilon, \theta_2 - \gamma T/\varepsilon; \varepsilon)$ *is analytic on the product of strips*

$$\{|\operatorname{Im}\theta_1| < r_1\} \times \{|\operatorname{Im}\theta_2| < r_2\} \times \{|\operatorname{Im}T| < \pi/2\};$$

2) the maximum of the modulus of the Melnikov function, $\max_{(\theta_1,\theta_2)\in\mathbb{T}^2}|M(\theta_1,\theta_2;\varepsilon)|$*, taken on the real arguments, can be bounded from above and from below by terms of the form*

$$\text{const}\,\varepsilon^{p-1}\exp\left(-\frac{c(\log\varepsilon)}{\sqrt{\varepsilon}}\right)$$

with different ε*-independent constants, where the function* c *in the exponent is defined by* (1.12)*;*

3) for a fixed small ε *only 4 terms dominate in the Fourier series for the Melnikov function and the rest can be estimated from above by* $O(e^{-C_1/\sqrt{\varepsilon}})$*, where the constant* $C_1 > \max c(\delta) = C_0\cosh(\log\sqrt{\gamma})$.

Remark 1 The number of leading terms depends on ε. In fact, the largest terms correspond to $(k_1,k_2) = \pm\left(F_{n(\varepsilon)+1}, -F_{n(\varepsilon)}\right)$, where $F_{n(\varepsilon)}$ is the Fibonacci number closest to $F^*(\varepsilon) = \sqrt{\phi_0/\varepsilon}$, ϕ_0 being a constant to be defined later in this section. Except for a small neighbourhood of $\varepsilon = \varepsilon^*\gamma^{-n}$, there is only one Fibonacci number closest to $F^*(\varepsilon)$, and then only two corresponding terms dominate in Fourier series.

Proof of the lemma. Taking into account the explicit formula (1.10) for $x_0(t)$ and $y_0(t)$ we obtain easily that

$$\begin{aligned} M(\theta_1,\theta_2;\varepsilon) &= \varepsilon^p\int_{-\infty}^{\infty} y_0(t)\sin(x_0(t))m(\theta_1 + t/\varepsilon, \theta_2 + \gamma t/\varepsilon)\,dt \\ &= -\varepsilon^p\int_{-\infty}^{\infty}\frac{4\sinh t}{\cosh^3 t}m(\theta_1 + t/\varepsilon, \theta_2 + \gamma t/\varepsilon)\,dt. \end{aligned}$$

To prove the assertion 1) of the lemma we note that

$$M(\theta_1 - T/\varepsilon, \theta_2 - \gamma T/\varepsilon; \varepsilon) = -\varepsilon^p\int_{-\infty}^{\infty}\frac{4\sinh(t+T)}{\cosh^3(t+T)}m(\theta_1 + t/\varepsilon, \theta_2 + \gamma t/\varepsilon)\,dt$$

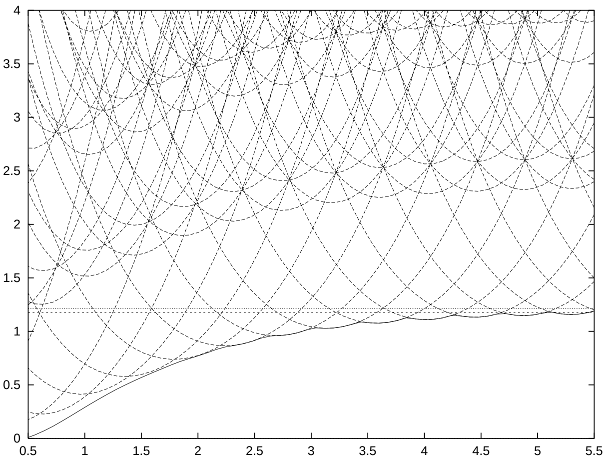

Figure 1: Each dashed line represents a Fourier coefficient of the Melnikov function as a function of ε: $-\sqrt{\varepsilon}\log_{10}|M_{k_1k_2}(\varepsilon)|$ versus $-\log_{10}\varepsilon$. The solid lines represents the maximum of the modulus of the Melnikov function in the same scale.

and the last integral is analytic with respect to θ_1, θ_2 and T. For the Fourier coefficients of the Melnikov function we have

$$M_{k_1,k_2}(\varepsilon) = -\varepsilon^p \int_{-\infty}^{\infty} \frac{4\sinh t}{\cosh^3 t} e^{i(k_1+\gamma k_2)t/\varepsilon}\, dt \ m_{k_1k_2}.$$

Calculating the integral by residues we obtain

$$M_{k_1,k_2}(\varepsilon) = -\frac{2\pi i \varepsilon^p (k_1+\gamma k_2)^2}{\varepsilon^2 \sinh\left(\frac{\pi(k_1+\gamma k_2)}{2\varepsilon}\right)} \ m_{k_1k_2}. \tag{2.1}$$

All these coefficients are exponentially small with respect to ε, but the constant in the exponent depends, in an essential way, on the coefficient index. The dependence on the largest Fourier coefficients on ε is represented in Figure 1 in logarithmic scale for a perturbation with $|m_{k_1k_2}| = 1$. In this figure for a fixed first coordinate the lower is a point the larger is the value. The scale is chosen in such a way that a horizontal line corresponds to the function $\exp(-C/\sqrt{\varepsilon})$ for some constant C.

The most important resonant terms correspond to Fibonacci numbers, that is $|k_1| = F_{n+1}$, $|k_2| = F_n$. Taking into account (1.5), (1.6) and (6.3) we bound these coefficients from below and from above by the terms of the form

$$\text{const}\, \frac{\varepsilon^{p-2}}{F_n^2} \exp\left(-\frac{\pi C_F}{2\varepsilon F_n} - (r_1\gamma + r_2)F_n\right).$$

For a fixed value of ε the first term in the exponent is an increasing function of F_n and the second one is decreasing. In order to describe this competition it is convenient to rewrite the last formula as

$$\text{const}\,\frac{\varepsilon^p}{\varepsilon^2 F_n^2}\exp\left(-\frac{C_0\cosh\left(\frac12\log(\varepsilon F_n^2)-\frac12\log\phi_0\right)}{\sqrt{\varepsilon}}\right),$$

where

$$C_0=\sqrt{2\pi C_F(r_1\gamma+r_2)},\qquad \phi_0=\frac{\pi C_F}{2(r_1\gamma+r_2)}.$$

For a fixed ε, the largest term corresponds to the minimal value of the numerator or, equivalently, minimizes $|\log(\varepsilon F_n^2)-\log\phi_0|$. This happens for F_n closest to

$$F^*(\varepsilon)=\sqrt{\frac{\phi_0}{\varepsilon}}.$$

That is, the index of the most important terms in the Fourier series for the Melnikov function grows as $1/\sqrt{\varepsilon}$. Except when $F^*(\varepsilon)$ lies exactly in the centre of an interval $[F_n,F_{n+1}]$ there is only one Fibonacci number closest to $F^*(\varepsilon)$. The index of the leading term changes when ε crosses this value. In a small neighbourhood of this value two terms are of the same order. In fact, the number of leading terms is two or four, respectively, since we have to take into account complex conjugate coefficients, $(-k_1,-k_2)$.

Since $F_n=C_F(\gamma^{n+1}+(-1)^n\gamma^{-n-1})$ we have

$$\log(\varepsilon F_n^2)=\log\varepsilon+2(n+1)\log\gamma+\log C_F+2\log\left(1+(-1)^n\gamma^{-2n-2}\right).$$

The value of this expression repeats with the error of the order $O(\gamma^{-2n-2})=O(F_n^{-2})=O(\varepsilon)$ when we increase n by 1 and simultaneously decrease $\log\varepsilon$ by $2\log\gamma$. Thus we obtain that $\sup_{k_1k_2}|M_{k_1k_2}(\varepsilon)|$ can be estimated from below and from above by

$$\text{const}\,\varepsilon^{p-1}\exp\left(-\frac{c(\log\varepsilon)}{\sqrt{\varepsilon}}\right),$$

where the function c was defined by (1.12). In the exponent the numerator oscillates between C_0 and $C_0\cosh\left(\log\sqrt{\gamma}\right)$ with the period $2\log\gamma$ in $\log\varepsilon$.

In particular, this gives the lower bound for the maximum of the Melnikov function modulus, since a Fourier coefficient of a function cannot be larger than the maximum value of the function.

The Fourier coefficients, which are not related to the Fibonacci numbers, can be estimated in the same manner, but with a constant larger than C_F. That implies that they are exponentially small with respect to the Fibonacci ones for small values of ε. The proof of the fact that the sum of these terms is also exponentially small is straightforward, and we omit it since it literally repeats the proof of Lemma 4. □

As we have established that for most small values of ε only the terms with $(k_1,k_2)=\pm(F_{n(\varepsilon)+1},-F_{n(\varepsilon)})$ are important, the Melnikov function is essentially

$$M(\theta_1,\theta_2;\varepsilon)\approx 2\left|M_{F_{n(\varepsilon)+1},-F_{n(\varepsilon)}}(\varepsilon)\right|\sin\left(F_{n(\varepsilon)+1}\theta_1-F_{n(\varepsilon)}\theta_2+\varphi(\varepsilon)\right).$$

The zeros of the Melnikov function correspond to homoclinic trajectories. The above formula implies that the zeros of the Melnikov function form two lines on the torus. As already noticed by C. Simó [Sim94], the averaged slopes of those lines approach to γ when $\varepsilon \to 0$.

3 Normal Form and Local Manifolds

As we have seen during the analysis of the Melnikov function, the size of the splitting depends essentially on the widths of the analyticity strip (r_1, r_2) of the angular variables θ_1, θ_2, as well as on the width of the analyticity strip of the separatrix $(x_0(t), y_0(t))$. Therefore, to detect the splitting in the quasiperiodic case the loss of domain in the angular variables must be very small (i.e., $\mathrm{O}(\varepsilon^\alpha)$, where α depends on the Diophantine properties of the frequencies). This makes another difference with the periodic case, where the size of the splitting does *not* depend on the width of the analyticity strip of the angular variable θ, but only on the width of the analyticity strip of the separatrix $(x_0(t), y_0(t))$. When dealing with the frequencies $(1/\varepsilon, \gamma/\varepsilon)$ one needs a reduction of $\mathrm{O}(\sqrt{\varepsilon})$ at most. Hence, during the proof of the convergence of the normal form one has to bound carefully the loss of domain (with respect to the angular variables) in order to achieve such a small reduction.

Finally, we want to stress that if the amount of reduction is something bigger, one can only produce upper bounds for the splitting of separatrices.

Theorem 2 (Normal Form Theorem) *Let $\varepsilon \in (0, \varepsilon_0)$. In a neighbourhood of the hyperbolic torus $\mathcal{T}$ there is a canonical change of variables $(x, y) \mapsto (X, Y)$, which depends 2π-periodically on θ_1 and θ_2, such that the Hamiltonian* (1.1) *takes the form*

$$H(XY, \varepsilon) = H_0(XY) + \varepsilon^{p-1} H_1(XY, \varepsilon),$$

where H_0 is the normal form Hamiltonian for the unperturbed pendulum. Moreover, the change of variables has the form

$$\begin{aligned} x &= x^{(0)}(X, Y) + \varepsilon^{p-1} x^{(1)}(X, Y, \theta_1, \theta_2, \varepsilon), \\ y &= y^{(0)}(X, Y) + \varepsilon^{p-1} y^{(1)}(X, Y, \theta_1, \theta_2, \varepsilon), \end{aligned} \tag{3.1}$$

where $(x^{(0)}, y^{(0)})$ are normal form coordinates for the unperturbed pendulum.

The functions H_0, H_1, $x^{(0)}$, $y^{(0)}$, $x^{(1)}$ and $y^{(1)}$ are analytic and uniformly bounded in the complex domain defined by

$$|X|^2 + |Y|^2 < r_0^2, \qquad |\operatorname{Im} \theta_1| < r_1 - \sqrt{\varepsilon}, \qquad |\operatorname{Im} \theta_2| < r_2 - \sqrt{\varepsilon},$$

for r_1 and r_2 from (1.5) *and some positive constant $r_0 > 0$.*

We prove this theorem in a more general form in section 8.

In the normal form coordinates the stable whisker is given by the equation $X = 0$ and the unstable one by the equation $Y = 0$. Let $\lambda = H'(0, \varepsilon)$. Normal Form Theorem provides a convenient parameterization for the invariant manifolds:

$$\begin{aligned} x = x^s(T, \theta_1, \theta_2, \varepsilon) \equiv x(0, e^{-\lambda T}, \theta_1, \theta_2, \varepsilon), \\ y = y^s(T, \theta_1, \theta_2, \varepsilon) \equiv y(0, e^{-\lambda T}, \theta_1, \theta_2, \varepsilon), \end{aligned} \tag{3.2}$$

and

$$\begin{aligned} x = x^u(T, \theta_1, \theta_2, \varepsilon) \equiv x(e^{\lambda T}, 0, \theta_1, \theta_2, \varepsilon), \\ y = y^u(T, \theta_1, \theta_2, \varepsilon) \equiv y(e^{\lambda T}, 0, \theta_1, \theta_2, \varepsilon), \end{aligned} \tag{3.3}$$

where we have used the change (3.1). Theorem 2 also implies that there is a positive ε-independent number T_0, such that

$$\begin{aligned} |x^u(T, \theta_1, \theta_2, \varepsilon) - x_0(T)| \le C\varepsilon^{p-1} \\ |y^u(T, \theta_1, \theta_2, \varepsilon) - y_0(T)| \le C\varepsilon^{p-1} \end{aligned} \quad \text{for} \quad T \le -T_0, \tag{3.4}$$

and

$$\begin{aligned} |x^s(T, \theta_1, \theta_2, \varepsilon) - x_0(T)| \le C\varepsilon^{p-1} \\ |y^s(T, \theta_1, \theta_2, \varepsilon) - y_0(T)| \le C\varepsilon^{p-1} \end{aligned} \quad \text{for} \quad T \ge T_0. \tag{3.5}$$

4 Extension Theorem

By the Normal Form Theorem, the unstable manifold is $O(\varepsilon^{p-1})$-close to the unperturbed separatrix. Next theorem extends this local approximation to a global one. Since the unperturbed separatrix $(x_0(T), y_0(T))$ has a singularity on $T = \pm\pi/2$, we will restrict ourselves to $|\operatorname{Im} T| \le \pi/2 - \sqrt{\varepsilon}$, i.e., up to a distance to the singularity $T = \pm\pi/2$ of the same order as the loss of domain in the angular variables. Besides, the extension time $t + T$ will be chosen big enough in order that the unperturbed separatrix reaches again the domain of convergence of the normal form.

Theorem 3 (Extension Theorem) *Let $(x_0(t), y_0(t))$ be the unperturbed homoclinic trajectory given in (1.10), let $\alpha \in (0, 1)$, $s = 2\alpha$, and assume $p - s - 2\alpha > 0$. Then, there exists $\varepsilon_0 > 0$ such that the following extension property holds:*

For any positive constants C and T_0 there exists a constant C_1, such that for any $\varepsilon \in (0, \varepsilon_0)$, every solution of system (1.3) that satisfies the initial conditions

$$|x(t_0) - x_0(t_0 + T)| \le C\varepsilon^{p-s}, \quad |y(t_0) - y_0(t_0 + T)| \le C\varepsilon^{p-s}, \tag{4.1}$$

$$|\operatorname{Im} \theta_1(t_0)| \le r_1 - \varepsilon^\alpha, \quad |\operatorname{Im} \theta_2(t_0)| \le r_2 - \varepsilon^\alpha,$$

for some $T \in \mathbb{C}$, $t_0 \in \mathbb{R}$ with

$$|\operatorname{Im} T| \le \pi/2 - \varepsilon^\alpha, \qquad -T_0 \le t_0 + \operatorname{Re} T \le 0,$$

can be extended for $-T_0 \le t + \operatorname{Re} T \le T_0$ and verifies there

$$|x(t) - x_0(t + T)| \le C_1\varepsilon^{p-s-2\alpha}, \quad |y(t) - y_0(t + T)| \le C_1\varepsilon^{p-s-2\alpha}.$$

The proof is given in section 9. From now on we fix $\alpha = 1/2$. Theorem 3 can be applied to get an approximation for the stable and the unstable manifold. As we will see in section 7, constructing the approximation for the invariant manifolds in such a complex domain enable us to bound the error on the real axis in a way sufficiently precise to detect the splitting.

Corollary 1 *The following estimate holds*

$$h_0(x^u, y^u) - h_0(x^s, y^s) = M(\theta_1 - T/\varepsilon, \theta_2 - \gamma T/\varepsilon; \varepsilon) + O\left(\varepsilon^{2(p-2)}\right), \tag{4.2}$$

where the value of h_0 is evaluated at points of the invariant manifolds corresponding to $(T, \theta_1, \theta_2, \varepsilon)$,

$$\operatorname{Re} T \in (T_0 - R, T_0), \quad |\operatorname{Im} T| \le \pi/2 - \sqrt{\varepsilon}, \quad |\operatorname{Im} \theta_k| \le r_k - \sqrt{\varepsilon}, \quad k = 1, 2, \tag{4.3}$$

for any positive constants T_0 and R, $R < T_0$. The constant in the estimate (4.2) *depends on these two constants.*

Proof of Corollary 1. Since $\dot{h}_0 = \{h_0, h\}$, we have

$$h_0(x^u, y^u) = \int_{-\infty}^{0} \{h_0, h\}(x^u, y^u, \theta + \omega t/\varepsilon, \varepsilon)\, dt$$

and

$$h_0(x^s, y^s) = -\int_{0}^{+\infty} \{h_0, h\}(x^s, y^s, \theta + \omega t/\varepsilon, \varepsilon)\, dt.$$

Inside the above integrals the functions x^u, y^u, x^s, y^s are evaluated on $(T+t, \theta+\omega t/\varepsilon, \varepsilon)$, with $\theta + \omega t/\varepsilon = (\theta_1 + t/\varepsilon, \theta_2 + \gamma t/\varepsilon)$. We can write the difference in the form

$$\begin{aligned} h_0(x^u, y^u) - h_0(x^s, y^s) &= \int_{-\infty}^{+\infty} \{h_0, h\}(x_0(t+T), y_0(t+T), \theta + \omega t/\varepsilon, \varepsilon)\, dt \\ &+ \int_{-\infty}^{0} \left(\{h_0, h\}(x^u, y^u, \theta + \omega t/\varepsilon, \varepsilon) - \{h_0, h\}(x_0, y_0, \theta + \omega t/\varepsilon, \varepsilon)\right) dt \\ &+ \int_{0}^{+\infty} \left(\{h_0, h\}(x^s, y^s, \theta + \omega t/\varepsilon, \varepsilon) - \{h_0, h\}(x_0, y_0, \theta + \omega t/\varepsilon, \varepsilon)\right) dt. \end{aligned}$$

We have three integrals in this expression. The first one is the Melnikov integral, and we have to bound the second and third one. We note that

$$\{h_0, h\}(x, y, \theta_1, \theta_2, \varepsilon) = \varepsilon^p\, y \sin x\, m(\theta_1, \theta_2)$$

and, consequently, since by (1.7) $|m(\theta_1, \theta_2)| \le K/\varepsilon$ holds on (4.3), we get

$$\begin{aligned} &\left|\{h_0, h\}(x^{u,s}, y^{u,s}, \theta_1, \theta_2, \varepsilon) - \{h_0, h\}(x_0, y_0, \theta_1, \theta_2, \varepsilon)\right| \\ &\qquad \le\ K\varepsilon^{p-1}\Big(|y^{u,s} - y_0|\,|\sin x^{u,s}| + |y_0|\,|\sin x^{u,s} - \sin x_0|\Big). \end{aligned}$$

We note that $\sin x^{u,s}$ decrease exponentially as t goes to $\pm\infty$, respectively, as well as y_0 as t goes to both $\pm\infty$. Then the extension theorem and the estimates (3.4–3.5) imply, that the second and third integrals are bounded by $O\left(\varepsilon^{2(p-2)}\right)$ and $O\left(\varepsilon^{2(p-1)}\right)$, respectively. Indeed, for the third integral we only have to use the estimate (3.4) to get a $O(\varepsilon^{2(p-1)})$-bound, and for the second integral, we only have to use the extension theorem to get a $O(\varepsilon^{2(p-2)})$-bound. (See lemma 10 for related bounds.) □

5 First return

A trajectory with initial conditions on the local unstable manifold leaves the domain of the normal form. Such a trajectory remains close to the homoclinic trajectory of the unperturbed pendulum at least during the time sufficient for the unperturbed trajectory to come back to the domain of the normal form. This part of the unperturbed separatrix, and, consequently, the corresponding part of the unstable manifold are close to the local stable manifold.

In order to describe the difference between the unstable and local stable manifolds it is convenient to take H and $T = -\log Y/\left(\varepsilon H'(XY)\right)$ as canonical coordinates near the stable separatrix. The equation of the local stable manifold is $H = 0$. In this coordinate system the unstable separatrix is a graph of a quasiperiodic function. This function is approximately the Melnikov function with the error $O\left(\varepsilon^{2p-4}\right)$. We use Fourier series arguments to show, that on the real value of the arguments the remainder is exponentially small. It is less than the amplitude of the Melnikov function, as shown in the next theorem, which contains the Main Theorem.

Theorem 4 *There exist positive constants T_0 and R, $R < T_0$, such that in the coordinate system $(H, T, \theta_1, \theta_2)$ the unstable manifold can be represented as the graph of the function $H = H^u(T, \theta_1, \theta_2; \varepsilon)$, where the function H^u depends 2π-periodically on θ_1 and θ_2. In the domain*

$$\begin{aligned} \operatorname{Re} T \in (T_0 - R, T_0), \qquad & |\operatorname{Im} T| \le \frac{\pi}{2} - \sqrt{\varepsilon}, \\ |\operatorname{Im}\theta_1| < r_1 - \sqrt{\varepsilon}, \qquad & |\operatorname{Im}\theta_2| < r_2 - \sqrt{\varepsilon}, \end{aligned} \tag{5.1}$$

this function is analytic and close to the Melnikov function:

$$H^u(T, \theta_1, \theta_2; \varepsilon) = M(\theta_1 - T/\varepsilon, \theta_2 - \gamma T/\varepsilon; \varepsilon) + O\left(\varepsilon^{2p-4}\right). \tag{5.2}$$

Moreover,

$$H^u(T, \theta_1, \theta_2; \varepsilon) = H^u(0, \theta_1 - T/\varepsilon, \theta_2 - \gamma T/\varepsilon; \varepsilon), \tag{5.3}$$

and its mean value is zero:

$$\iint_{\mathbb{T}^2} H^u(0, \theta_1, \theta_2; \varepsilon)\, d\theta_1 d\theta_2 = 0. \tag{5.4}$$

Furthermore, for $p > 3$ and real T, θ_1 and θ_2,

$$|H^u(T, \theta_1, \theta_2; \varepsilon) - M\left(\theta_1 - T/\varepsilon, \theta_2 - \gamma T/\varepsilon; \varepsilon\right)| \le \operatorname{const} \varepsilon^{2p-4} \exp\left(-\frac{c(\log\varepsilon)}{\sqrt{\varepsilon}}\right), \tag{5.5}$$

where $c(\delta)$ is defined in (1.12). If the condition (1.6) *is fulfilled then the Melnikov function is larger than the right hand side of (5.5).*

Proof. The equation of the unstable manifold is

$$x = x^u(T^u, \theta_1, \theta_2, \varepsilon), \qquad y = y^u(T^u, \theta_1, \theta_2, \varepsilon).$$

The Extension Theorem implies that they are ε^{p-2}-close to the pendulum separatrix for $|\operatorname{Re} T^u| < T_0$. Choosing T_0 large enough we ensure that the segment of the unstable separatrix, which correspond to $T_0 - R < \operatorname{Re} T^u < T_0$ belongs to the domain of the normal form. Then we can represent this segment in the parametric form

$$H = \tilde{H}^u(T^u, \theta_1, \theta_2, \varepsilon), \qquad T = \tilde{T}(T^u, \theta_1, \theta_2, \varepsilon) \tag{5.6}$$

evaluating H and T at a point $x = x^u(T^u, \theta_1, \theta_2, \varepsilon)$, $y = y^u(T^u, \theta_1, \theta_2, \varepsilon)$. Denote by X^u and Y^u the value of the normal form coordinates at the corresponding point. As it was pointed out previously, the stable manifold is given by $X^s = 0$ and $Y^s(T, \theta_1, \theta_2) = e^{-\lambda T}$. Using the Normal Form theorem, we obtain

$$\begin{aligned}
&\tilde{H}^u(T^u, \theta_1, \theta_2, \varepsilon)\\
&\quad= H(X^uY^u, \varepsilon) = H(X^uY^u, \varepsilon) - H(X^sY^s, \varepsilon)\\
&\quad= H_0(X^uY^u) - H_0(X^sY^s) + \varepsilon^{p-1}\Big(H_1(X^uY^u, \varepsilon) - H_1(X^sY^s, \varepsilon)\Big)\\
&\quad= H_0(X^uY^u) - H_0(X^sY^s) + O(\varepsilon^{2(p-2)})\\
&\quad= H_0(X^{(0)}(x^u, y^u)Y^{(0)}(x^u, y^u)) - H_0(X^{(0)}(x^s, y^s)Y^{(0)}(x^s, y^s)) + O(\varepsilon^{2(p-2)})\\
&\quad= h_0(x^u, y^u) - h_0(x^s, y^s) + O(\varepsilon^{2(p-2)}),
\end{aligned}$$

where $X^{(0)}$ and $Y^{(0)}$ denote the normal form coordinates of the unperturbed pendulum. Here the parameterizations of both invariant manifolds are taken at a point $(T^u, \theta_1, \theta_2, \varepsilon)$. Together with the estimate (4.2) this implies

$$\tilde{H}^u(T^u, \theta_1, \theta_2, \varepsilon) = M\left(\theta_1 - T^u/\varepsilon, \theta_2 - \gamma T^u/\varepsilon; \varepsilon\right) + O\left(\varepsilon^{2(p-2)}\right). \tag{5.7}$$

Now we have to eliminate the parameter T^u. From

$$\begin{aligned}
&\tilde{T}(T^u, \theta_1, \theta_2, \varepsilon) = T(X^u(T^u, \theta_1, \theta_2, \varepsilon), Y^u(T^u, \theta_1, \theta_2, \varepsilon))\\
&\quad= T(X^{(0)}(x^u, y^u), Y^{(0)}(x^u, y^u)) + O(\varepsilon^{p-2})\\
&\quad= T\left(X^{(0)}(x_0(T^u), y_0(T^u)), Y^{(0)}(x_0(T^u), y_0(T^u))\right) + O(\varepsilon^{p-2})\\
&\quad= T^u + O(\varepsilon^{p-2})
\end{aligned}$$

on the domain (5.1), we get, by Cauchy estimates, that

$$\frac{\partial \tilde{T}}{\partial T^u}(T^u, \theta_1, \theta_2, \varepsilon) = 1 + O(\varepsilon^{p-5/2}).$$

If $p > 5/2$, the Implicit Function Theorem allows us to eliminate $\tilde{T}$ from the first equation in (5.6) and to obtain the estimate (5.2).

Suppose for a moment that the mean value of the function $H^u(T,\theta_1,\theta_2;\varepsilon)$ with respect to the angle variables is equal to zero. Then the estimate (5.5) for real T, θ_1 and θ_2, is a consequence of the quasiperiodicity of H^u, the estimate (5.2) and of Lemma 4 of section 7. In the last lemma one has to replace r_1 and r_2 by $r_1-\sqrt{\varepsilon}$ and $r_2-\sqrt{\varepsilon}$, respectively, and take $\rho=\pi/2-\sqrt{\varepsilon}$.

Lemma 1 shows that for $p>3$ the amplitude of the Melnikov function is larger then the error term in (5.5). The exponentially small upper bound for the error is proved for $p>5/2$. So what we have for $5/2<p\le 3$ is a very sharp upper bound for the splitting.

To use Lemma 4 we have to prove that the mean value of the function $H^u(T,\theta_1,\theta_2;\varepsilon)$ with respect to the angle variables is equal to zero. Indeed, in the variables H,T,θ_1,θ_2 the equations of motion have the form

$$\dot H=0,\qquad \dot T=1,\qquad \dot\theta_1=\frac{1}{\varepsilon},\qquad \dot\theta_2=\frac{\gamma}{\varepsilon},$$

Since H is an integral of motion we obtain (5.3).

The proof of the equality (5.4) is completely analogous to the case of periodic perturbation. Consider the part of the phase space bounded by a KAM torus and the segments of the stable and unstable separatrices. Since the flow is Hamiltonian the volume of this subset is time-invariant, that is, the volume of the trajectories which enter the subset equals to the volume of the trajectories which leave it. The trajectories may enter or leave this subset only through the "turnstile" formed by the split separatrices. The algebraic value of the volume, which passes through the "turnstile" during a small time interval Δt, may be evaluated in the coordinate system (H,T,θ_1,θ_2) as

$$\Delta t\iint_{\mathbb{T}^2} H^u(T,\theta_1,\theta_2;\varepsilon)\,d\theta_1 d\theta_2.$$

By (5.3), this integral does not depend on T, and thus it should be zero.

In other words, the equality (5.4) means that in average there is no diffusion in the direction of the T-axis. $\square$

6 Rational Approximation of γ

In this section we discuss the approximations of the number $\gamma=\frac{\sqrt{5}+1}{2}$ by rational numbers. The best approximation is given in terms of Fibonacci numbers, which are defined by the following recurrent formula

$$F_0=1,\qquad F_1=1\qquad F_{n+1}=F_n+F_{n-1},\quad n>1.$$

It is easy to check the following:

$$F_{n-1}=\frac{\gamma^n-(-1)^n\gamma^{-n}}{\gamma+\gamma^{-1}},\tag{6.1}$$

and

$$F_n - \gamma F_{n-1} = \frac{(-1)^n}{\gamma^n} = \frac{(-1)^n}{F_n + \gamma^{-1} F_{n-1}}. \tag{6.2}$$

For large values of n this implies

$$F_n - \gamma F_{n-1} = (-1)^n \frac{C_F}{F_{n-1}} + O\left(\frac{1}{F_{n-1}^3}\right), \qquad C_F = \frac{1}{\gamma + \gamma^{-1}}. \tag{6.3}$$

The estimation of the following lemma is not sharp, but it is sufficient for our proposes.

Lemma 2 *If $N \in \mathbb{N}$ is not a Fibonacci number, then for all integer k*

$$|k - \gamma N| > \frac{\gamma C_F}{N}. \tag{6.4}$$

Proof. The first Fibonacci numbers are $F_1 = 1$, $F_2 = 2$, $F_3 = 3$ and $F_4 = 5$. Let us define

$$d_n = \min_{1 \le j < F_n} \min_k |k - \gamma j|.$$

Equation (6.2) implies that $d_2 = \gamma^{-2}$ and $d_3 = \gamma^{-3}$.

Suppose that $d'_n = \gamma^{-n'}$ for all n', $2 \le n' \le n$. First, consider an integer number j, $F_n < j < F_{n+1}$, and let $j' = j - F_n$. Obviously, $1 \le j' < F_{n+1} - F_n = F_{n-1}$ and we have for all k

$$\begin{aligned} |k - \gamma j| &= |k - F_{n+1} - \gamma j' + (F_{n+1} - \gamma F_n)| \\ &\ge |k - F_{n+1} - \gamma j'| - |F_{n_0+1} - \gamma F_n| \\ &\ge d_{n-1} - \gamma^{-n-1} = \gamma^{-n+1} - \gamma^{-n-1} = \gamma^{-n}. \end{aligned} \tag{6.5}$$

Then take $j = F_n$,

$$\min_k |k - \gamma j| = |F_{n+1} - \gamma F_n| = \gamma^{-n-1}.$$

Comparing with the previous inequality, we see that the minimum is reached at the Fibonacci numbers. So we have $d_{n+1} = \gamma^{-n-1}$. Consequently, by induction we obtain that

$$d_n = \gamma^{-n} \quad \text{for all} \quad n \ge 2.$$

Now consider a non Fibonacci number N, let $F_n < N < F_{n+1}$. Formula (6.1) implies that if $N > F_n$, then $N > \frac{\gamma+1}{\gamma+2}\gamma^n$. The estimates (6.5) with $j = N$ implies that

$$|k - \gamma N| > \gamma^{-n} \ge \frac{\gamma+1}{\gamma+2}\frac{1}{N},$$

which is equivalent to the desired estimate (6.4). $\square$

7 Exponentially small upper bounds

The two following lemmas provide a tool to pass from estimate (5.2) to the sharp estimate (5.5). The proof of the first lemma is standard.

Lemma 3 *Let $F(\theta_1 + s/\varepsilon, \theta_2 + \gamma s/\varepsilon)$ be a 2π-periodic function of the variables θ_1, θ_2, analytic in the product of strips $|\mathrm{Im}\,\theta_1| \le r_1$, $|\mathrm{Im}\,\theta_2| \le r_2$ and $|\mathrm{Im}\,s| \le \rho$, and $|F| \le A$ for these values of the variables. Then for all k_1, $k_2 \in \mathbb{Z}$*

$$|F_{k_1k_2}| \le Ae^{-|k_1|r_1-|k_2|r_2}e^{-\rho|k_1+\gamma k_2|/\varepsilon}. \tag{7.1}$$

Consider the $(2\log\gamma)$-periodic function $c_{\rho,r_1,r_2}(\delta)$ defined on the interval $[\log\varepsilon^* - \log\gamma, \log\varepsilon^* + \log\gamma]$ by

$$c_{\rho,r_1,r_2}(\delta) = C_0\cosh\left(\frac{\delta - \log\varepsilon^*}{2}\right), \tag{7.2}$$

$$\varepsilon^* = \frac{\rho(\gamma+\gamma^{-1})}{(\gamma r_1 + r_2)\gamma^2}, \qquad C_0 = 2\sqrt{\frac{(\gamma r_1 + r_2)\rho}{\gamma+\gamma^{-1}}}. \tag{7.3}$$

and continued by $2\log\gamma$-periodicity.

The following lemma gives the exponentially small upper bound for the function F for real values of the variables.

Lemma 4 *Let F satisfy the conditions of Lemma 3. If $\gamma = (\sqrt{5}+1)/2$ is the golden mean number and the mean value of the function F is zero, then*

$$|F(\theta_1,\theta_2)| \le \mathrm{const}\, A\exp\left(-\frac{c_{\rho,r_1,r_2}(\log\varepsilon)}{\sqrt{\varepsilon}}\right) \tag{7.4}$$

on the real values of its arguments. The constant depends continuously on r_1 and r_2 on $r_1 > 0$ and $r_2 > 0$.

Proof. If the arguments of the function F are real, then

$$|F(\theta_1,\theta_2)| \le \sum_{|k_1|+|k_2|\ne 0} |F_{k_1k_2}| \le A \sum_{|k_1|+|k_2|\ne 0} \exp(-|k_1|r_1 - |k_2|r_2 - \rho|k_1+\gamma k_2|/\varepsilon), \tag{7.5}$$

due to the estimate (7.1). In order to estimate the last sum in (7.5) we separate it into two parts. The first one contains non-resonant terms, that is, all the terms such that $|k_1 + \gamma k_2| \ge 1/2$. We easily obtain the upper estimate

$$\sum_{|k_1+\gamma k_2|\ge 1/2} e^{-|k_1|r_1-|k_2|r_2-\rho|k_1+\gamma k_2|/\varepsilon}$$

$$< e^{-\rho/(2\varepsilon)} \sum_{|k_1|+|k_2|\ne 0} e^{-|k_1|r_1-|k_2|r_2} = \frac{2(e^{-r_1}+e^{r_2}-e^{-r_1-r_2})e^{-\rho/(2\varepsilon)}}{(1-e^{-r_1})(1-e^{-r_2})}. \tag{7.6}$$

For the resonant terms we have $|k_1+\gamma k_2|<1/2$. Obviously, for every k_2 there exists exactly one integer $k_1=k_1(k_2)$ such that this inequality holds. Since the coefficients of the sum (7.5) are even with respect to (k_1,k_2) we can assume that k_2 is positive and, at the end, multiply the estimates by 2.

The sum of the resonant terms with $k_2\ge\varepsilon^{-1}$ can be easily estimated:

$$\sum_{k_2\ge\varepsilon^{-1}} e^{-|k_1|r_1-|k_2|r_2-\rho|k_1+\gamma k_2|/\varepsilon} \le \sum_{k_2\ge\varepsilon^{-1}} e^{-|k_1|r_1-|k_2|r_2}$$
$$\le \sum_{k_2\ge\varepsilon^{-1}} e^{r_1/2-(\gamma r_1+r_2)k_2} \le \frac{e^{r_1/2}e^{-(\gamma r_1+r_2)/\varepsilon}}{1-e^{-(\gamma r_1+r_2)}}. \tag{7.7}$$

Now we estimate the resonant terms with $1\le k_2<\varepsilon^{-1}$. The number of such terms is large, but finite. We will show, that all of them, except at most 4, can be estimated by $O(e^{-C_1/\sqrt{\varepsilon}})$ with a constant $C_1>\max_\delta c_{\rho,r_1,r_2}(\delta)=C_0\left(\gamma^{1/2}+\gamma^{-1/2}\right)/2$.

Let B denote the following expression from the exponent of the right hand of (7.5), obtained after substituting $|k_1|=\gamma k_2$:

$$B(k_2,\varepsilon)=(\gamma r_1+r_2)k_2\sqrt{\varepsilon}+\frac{\rho|k_1+\gamma k_2|}{\sqrt{\varepsilon}}.$$

It is sufficient to provide an appropriate lower bound for this function.

If k_2 is not a Fibonacci number, then we use Lemma 2:

$$B(k_2,\varepsilon)\ge(\gamma r_1+r_2)k_2\sqrt{\varepsilon}+\frac{\rho\gamma C_F}{k_2\sqrt{\varepsilon}}\ge 2\sqrt{(\gamma r_1+r_2)\rho\gamma C_F}\equiv C_1=\sqrt{\gamma}C_0.$$

If k_2 is a Fibonacci number, then instead of (6.4), we use (6.2)

$$|k_1+\gamma k_2|=\frac{1}{|k_1+\gamma^{-1}k_2|}\ge\frac{1}{\gamma k_2+1+\gamma^{-1}k_2}=\frac{C_F}{k_2+C_F}$$

to obtain

$$B(k_2,\varepsilon)\ge(\gamma r_1+r_2)k_2\sqrt{\varepsilon}+\frac{\rho C_F}{(k_2+C_F)\sqrt{\varepsilon}}.$$

Provided ε is small, $0<\varepsilon<\varepsilon_0$, there are two positive numbers K_1 and K_2, such that the right hand side of the last inequality is larger than C_1 for k_2 outside the interval $(K_1/\sqrt{\varepsilon},K_2/\sqrt{\varepsilon})$. Moreover, this interval contains at most 2 Fibonacci numbers, that is,

$$B(k_2,\varepsilon)\ge C_1 \tag{7.8}$$

for all except at most 2 terms. For these exceptional terms

$$B(k_2,\varepsilon)\ge(\gamma r_1+r_2)k_2\sqrt{\varepsilon}+\frac{\rho C_F}{k_2\sqrt{\varepsilon}}-\sqrt{\varepsilon}\frac{\rho C_F^2}{K_1(K_1+C_F\sqrt{\varepsilon})},$$

and it is convenient to rewrite

$$B(k_2,\varepsilon) \geq C_0 \cosh\left(\log(k_2\sqrt{\varepsilon}) - \log\sqrt{\frac{\rho C_F}{\gamma r_1 + r_2}}\right) - O(\sqrt{\varepsilon}).$$

The above $O(\sqrt{\varepsilon})$ term affects only the constant in front of the estimate (7.4), since the terms in the sum of the right hand of (7.5) are of the form $\exp(-B(k_2,\varepsilon)/\sqrt{\varepsilon})$. Since k_2 is a Fibonacci number, $k_2 = F_n$ for some n and taking into account (6.1), we obtain

$$B(k_2,\varepsilon) \geq C_0 \cosh\left(\frac{1}{2}\log\varepsilon + n\log\gamma + \log\frac{\gamma+1}{\gamma+2} - \log\sqrt{\frac{\rho C_F}{\gamma r_1 + r_2}}\right) - O(\sqrt{\varepsilon}).$$

The envelope of this family of curves is the function $c_{\rho,r_1,r_2}(\delta)$ defined by equation (7.2).

Thus, in the sum of the resonant terms there is one leading term which is exponentially larger than the others except in the neighbourhoods of $\varepsilon = \varepsilon^*\gamma^n$, when the index of the leading term changes, and there are two terms of the same order. Moreover, we have established that for all resonant terms, with $k_2 < \varepsilon^{-1}$,

$$B(k_2,\varepsilon) \geq c_{\rho,r_1,r_2}(\log\varepsilon) - O(\sqrt{\varepsilon}).$$

Together with the estimates (7.6), (7.7) and (7.8), this completes the proof. □

8 Normal Form Theorem

This theorem gives a convergent normal form in a neighbourhood of a hyperbolic torus of a one degree of freedom Hamiltonian system under quasiperiodic time-dependent perturbations (with two frequencies). The main contribution is that the convergence is ensured in a wide domain of the angle variables θ of the perturbation: the loss of domain is only $O(\sqrt{\varepsilon})$ in the complex direction of θ.

Theorem 5 (Normal Form theorem) *Let K be a Hamiltonian of the form*

$$K(x,y,\theta,p,\varepsilon) = \frac{\omega}{\varepsilon}\cdot p + h_0(x,y,\varepsilon) + \varepsilon^q h_1(x,y,\theta,\varepsilon), \tag{8.1}$$

with regard to the symplectic form $\mathrm{d}x\wedge\mathrm{d}y + \mathrm{d}\theta\wedge\mathrm{d}p$, with h_0, h_1 analytic in the variables x, y, θ and with continuous (and bounded) dependence on ε, on the set $|x|,|y| < r_0$, $|\mathrm{Im}\,\theta_i| < \rho_i - \sqrt{\varepsilon}$ $(i=1,2)$, $0<\varepsilon<\varepsilon_0$, for some positive constants r_0, $\rho=(\rho_1,\rho_2)$ and ε_0. Assume also:

1. *There exists $c>0$ such that $|k\cdot\omega| \geq c/|k|$, $\forall k\in\mathbb{Z}^2\setminus\{0\}$.*
2. *The origin is a saddle point of the Hamiltonian $h_0(x,y,0)$.*
3. $h_1(0,0,\theta,\varepsilon) = \partial_x h_1(0,0,\theta,\varepsilon) = \partial_y h_1(0,0,\theta,\varepsilon) = 0.$

Then, there exists ε_1 $(0 < \varepsilon_1 < \varepsilon_0)$*,* r_1 $(0 < r_1 < r_0)$ *and a canonical change of variables*

$$\begin{aligned} x &= x^{(0)}(X, Y, \varepsilon) + \varepsilon^q x^{(1)}(X, Y, \theta, \varepsilon), \\ y &= y^{(0)}(X, Y, \varepsilon) + \varepsilon^q y^{(1)}(X, Y, \theta, \varepsilon), \\ \theta &= \theta, \\ p &= p(X, Y, \theta, P, \varepsilon), \end{aligned}$$

analytic in the variables X*,* Y*,* θ*,* P*, and bounded and continuous in* ε*, on the set* $|x|, |y| < r_1$*,* $|\mathrm{Im}\, \theta_i| < \rho_i - 2\sqrt{\varepsilon}$ $(i = 1, 2)$*,* $0 < \varepsilon < \varepsilon_1$*, which transforms Hamiltonian* (8.1) *into its normal form:*

$$\mathcal{K}(X, Y, P, \varepsilon) = \frac{\omega}{\varepsilon} \cdot P + H_0(XY, \varepsilon) + \varepsilon^q H_1(XY, \varepsilon).$$

Moreover, the canonical change of variables

$$\begin{aligned} x &= x^{(0)}(X, Y, \varepsilon), \\ y &= y^{(0)}(X, Y, \varepsilon), \end{aligned}$$

transforms the unperturbed Hamiltonian h_0 *into its normal form* H_0*.*

8.1 Idea of the Proof

The proof is based on a quadratically convergent scheme, similar to the one used in the proof of KAM Theorem (see [Arn63]).

The first step is to put the Hamiltonian in the action-angle variables of H_0:

$$\frac{\omega}{\varepsilon} \cdot p + H_0(z, \varepsilon) + \varepsilon^q H_1(z, \varphi, \theta, \varepsilon),$$

where the couples (z, φ) and (p, θ) correspond to canonically conjugated variables. Here $H_0(z, 0)$ is the normal form of the unperturbed Hamiltonian.

The next step is to start a sequence of changes of variables to kill the term $\varepsilon^q H_1$, in the same way it is done in the proof of the KAM Theorem. The main difference is due to the fact that the "small" divisors we will obtain are of the form

$$\varepsilon \ell \lambda_0(z, \varepsilon) + \sqrt{-1} k \cdot \omega, \qquad \ell \in \mathbb{Z},\ k \in \mathbb{Z}^2,\ |\ell| + |k| \neq 0,$$

where both $\lambda_0(z, \varepsilon) := \partial_z H_0(z, \varepsilon)$ and ω are real. This makes that all the divisors with $\ell \neq 0$ are separated from zero, so they do not produce convergence problems. The only small divisors correspond to the case $\ell = 0$, so we will ask ω to satisfy a suitable Diophantine condition. In this case, *as* ω *is fixed in all the iterative process* (see Lemma 7), the initial Diophantine condition will be satisfied in all the steps of the iterative scheme. This also implies that we can have convergence on an open set of the phase space. This makes a difference with standard KAM problems: there it is usual that frequencies depend on actions and then they have to be controlled at each step of the proof. This leads to take out a Cantor-like set of actions, so the convergence is only proved on sets with empty interior. Finally, let us comment that the quadratic convergence allows us to be very strict in the amount of domain lost at each step, so we have been able to show that the loss of domain in the θ variables is bounded by $\sqrt{\varepsilon}$.

We want to stress that the good properties of this case are due to the fact that the unperturbed problem has a saddle point with one degree of freedom. If the saddle is replaced by a centre (with one or more degrees of freedom) we obtain a standard KAM Theorem valid only on a set with empty interior (see [JS96]). If the unperturbed Hamiltonian has a saddle point with more than one degrees of freedom we obtain new small divisors (the resonances between the eigenvalues of the saddle when $k = 0$), that require to be controlled by using the actions of the Hamiltonian. This produces a KAM Theorem on the conservation of hyperbolic invariant manifolds.

There is another detail worth to comment: if we proceed exactly in the way mentioned above, we have technical problems due to the lack of definition of action-angle variables at the origin (we want to show convergence on a neighbourhood of that point!). To avoid this difficulty, we will work all the time with spatial (cartesian) coordinates, but grouping them as if they were the action and angle ones. One can see that (some of) these groupings have poles at the origin (as expected), but they also have factors that cancel those singularities (as expected too). So, bounding them together we can show that the whole thing is well defined and it is convergent in a neighbourhood of the origin.

Now, let us go for the details.

8.2 Technical lemmas

This section contains the lemmas used during the proof of Theorem 5.

Lemma 5 *Let $\omega = (\omega_1, \omega_2) \in \mathbb{R}^2$ such that $|k \cdot \omega| \geq c/|k|$, $\forall k \in \mathbb{Z}^2 \setminus \{0\}$, for some constant $c > 0$. Then there exists a constant $\beta = \beta(\omega) > 0$ such that the following inequality holds for all $\alpha \in (0, 1]$:*

$$\sum_{k \in \mathbb{Z}^2 \setminus \{0\}} \frac{\mathrm{e}^{-\alpha |k|}}{|k \cdot \omega|} \leq \frac{\beta}{\alpha^2}.$$

Proof. Assume, for instance, that $|\omega_1| \leq |\omega_2|$. Given an integer $k_1 \neq 0$, there exists a unique integer $k_2 = k_2^*(k_1)$ such that $|k_1 \omega_1/\omega_2 + k_2| < 1/2$, and in particular, such that $|k_1\omega_1 + k_2\omega_2| < |\omega_2|/2$. Moreover, $|k_2^*(k_1)| \leq |k_1\omega_1/\omega_2| + |k_1\omega_1/\omega_2 + k_2^*(k_1)| < |k_1| + 1/2$. Therefore,

$$\begin{aligned}
\sum_{k \in \mathbb{Z}^2 \setminus \{0\}} \frac{\mathrm{e}^{-\alpha|k|}}{|k \cdot \omega|} &= \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \neq k_2^*(k_1)} \frac{\mathrm{e}^{-\alpha|k|}}{|k \cdot \omega|} + \sum_{k_1 \neq 0,\, k_2 = k_2^*(k_1)} \frac{\mathrm{e}^{-\alpha|k|}}{|k \cdot \omega|} \\
&< \frac{2}{|\omega_2|} \sum_{k \in \mathbb{Z}^2} \mathrm{e}^{-\alpha|k|} + \frac{1}{c} \sum_{k_1 \neq 0} \mathrm{e}^{-\alpha|k_1|} \left(2\,|k_1| + \frac{1}{2}\right)
\end{aligned}$$

$$= \frac{2}{|\omega_2|}\left(\coth\frac{\alpha}{2}\right)^2 + \frac{2}{c}\sum_{n=1}^{\infty} \mathrm{e}^{-\alpha n}\left(2n+\frac{1}{2}\right)$$
$$= \left(\frac{2}{|\omega_2|} + \frac{5-\mathrm{e}^{-\alpha}}{4c}\right)\left(\coth\frac{\alpha}{2}\right)^2 \le \frac{\beta}{\alpha^2}. \qquad \square$$

Lemma 6 *Let us consider the change of variables* $x = x(X, Y, \theta, \varepsilon)$, $y = y(X, Y, \theta, \varepsilon)$ *given implicitly by*

$$\left.\begin{aligned} X &= x + \varepsilon^m \frac{\partial S}{\partial Y}(x, Y, \theta, \varepsilon), \\ y &= Y + \varepsilon^m \frac{\partial S}{\partial x}(x, Y, \theta, \varepsilon), \end{aligned}\right\} \tag{8.2}$$

where $S(z_1, z_2, \theta, \varepsilon)$ *is defined on the set* $\varepsilon \in (0, \varepsilon_0]$, $|z_i| < r$, $|\operatorname{Im}\theta_i| \le \rho_i$, $i = 1, 2$, *with* $z = (z_1, z_2) \in \mathbb{C}^2$ *and* $\operatorname{Re}\theta = (\operatorname{Re}\theta_1, \operatorname{Re}\theta_2) \in \mathbb{T}^2$*. Moreover, it depends on* z_1, z_2 *and* θ *in an analytic way, and has continuous and bounded dependence on* ε*. Let us also assume that* $|S(z_1, z_2, \theta, \varepsilon)| \le M$ *on this set and let* η *be such that* $0 < \eta < 1$*. Then, if*

$$\varepsilon_0^m \le \frac{r^2\eta^2(1-\eta)}{2M}, \tag{8.3}$$

equation (8.2) *defines the change* $x = x(X, Y, \theta, \varepsilon)$, $y = y(X, Y, \theta, \varepsilon)$ *analytic in the variables* X, Y, θ *and bounded and continuous in* ε, *on the set* $|X| < r(1-\eta)^2$, $|Y| < r(1-\eta)^2$, $|\operatorname{Im}\theta_i| \le \rho_i$*. Besides, the following bounds hold:*

$$|X - x| \le \varepsilon^m \frac{M}{r\eta}, \qquad |Y - y| \le \varepsilon^m \frac{M}{r\eta}. \tag{8.4}$$

The proof of this lemma is omitted, since it is straightforward. Now let us introduce some notations to be used in this section.

Let $f(\theta)$ be a periodic function, $\theta \in \mathbb{T}^2$, analytic on a (complex) strip of width $\rho = (\rho_1, \rho_2)$, that is, analytic on $|\operatorname{Im}\theta| < \rho$ and continuous on the boundary. We denote by $\|f\|_\rho$ the sup norm on that set, this is,

$$\|f\|_\rho = \sup_{|\operatorname{Im}\theta| \le \rho} |f(\theta)|,$$

where $|\operatorname{Im}\theta| \le \rho$ means $|\operatorname{Im}\theta_1| \le \rho_1$ and $|\operatorname{Im}\theta_2| \le \rho_2$. Moreover, if $F(z, \theta)$ is an analytic periodic function with respect to $\theta \in \mathbb{T}^2$ on a strip of width ρ, and analytic with respect to z on $|z| \le r$, we define the norm

$$\|F\|_{r,\rho} = \sup_{|z| \le r} \|F(z, \cdot)\|_\rho = \sup_{\substack{|z| \le r \\ |\operatorname{Im}\theta| \le \rho}} |F(z, \theta)|.$$

If F depends only on z, we simply denote $\|F\|_{r,0} = \sup_{|z| \le r} |F(z)|$. Here we have assumed that $z \in \mathbb{C}$ or $z \in \mathbb{C}^2$. Of course, in this last case the notation $|z| \le r$ means $|z_1| \le r$ and $|z_2| \le r$.

For the sake of the simplicity (and without loss of generality), we will assume that $\varepsilon_0 \leq 1$, $\lambda \leq 1$, $r \leq 1$, $A_0 \geq 1$ and $A_1 \geq 1$. This will be used along this section to avoid cumbersome bounds.

Lemma 7 (Inductive lemma) *Let us consider the Hamiltonian*

$$\frac{\omega}{\varepsilon} \cdot p + K_0(xy, \varepsilon) + \varepsilon^m K_1(x, y, \theta, \varepsilon), \tag{8.5}$$

where

1. $K_0(z, \varepsilon)$ *is an analytic function of* z *on* $|z| \leq r^2$ *and* $0 < \varepsilon \leq \varepsilon_0$.
2. $|\partial_z K_0(z, \varepsilon)| \geq \lambda > 0$, *on* $|z| \leq r^2$, $0 < \varepsilon \leq \varepsilon_0$.
3. $K_0(z, \varepsilon)$ *and* $K_1(x, y, \theta, \varepsilon)$ *depend on* ε *in a continuous and bounded way, if* $0 < \varepsilon \leq \varepsilon_0$.
4. $K_1(x, y, \theta, \varepsilon)$ *is analytic on* $|x| \leq r$, $|y| \leq r$, $|\operatorname{Im} \theta_i| \leq \rho_i$, *i=1,2.*
5. $K_1(0, 0, \theta, \varepsilon) = \partial_x K_1(0, 0, \theta, \varepsilon) = \partial_y K_1(0, 0, \theta, \varepsilon) = 0$.

Moreover, let us define $A_0 = \|K_0\|_{r,0}$ *and* $A_1 = \|K_1\|_{r,\rho}$, *and let us consider* $0 < \eta < 1/2$ *and* $0 < \delta < 1$. *Then, for* ε_0 *small enough (which is detailed below), there exists a canonical change of variables, given implicitly by a generating function* $\theta \cdot P + Yx + \varepsilon^m S(x, Y, \theta, \varepsilon)$ *such that the new Hamiltonian takes the form*

$$\frac{\omega}{\varepsilon} \cdot P + \widehat{K}_0(XY, \varepsilon) + \varepsilon^{2m} \widehat{K}_1(X, Y, \theta, \varepsilon), \tag{8.6}$$

where

1. *The "normal form"* $\widehat{K}_0(Z, \varepsilon)$ *is an analytic function of* Z *on* $|Z| \leq r^2(1-\eta)^6$, $0 < \varepsilon \leq \varepsilon_0$.
2. $|\partial_Z \widehat{K}_0(Z, \varepsilon)| \geq \lambda - \varepsilon_0^m A_1/\eta > 0$, *on* $|Z| \leq r^2(1-\eta)^6$.
3. $\widehat{K}_0(Z, \varepsilon)$ *and* $\widehat{K}_1(X, Y, \theta, \varepsilon)$ *depend on* ε *in a continuous and bounded way, if* $0 < \varepsilon \leq \varepsilon_0$.
4. $\widehat{K}_1(X, Y, \theta, \varepsilon)$ *is analytic on* $|X| \leq r(1-\eta)^3$, $|Y| \leq r(1-\eta)^3$, $|\operatorname{Im} \theta_i| \leq \widehat{\rho}_i = \rho_i - \delta\sqrt{\varepsilon}$, *i=1,2.*
5. $\widehat{K}_1(0, 0, \theta, \varepsilon) = \partial_x \widehat{K}_1(0, 0, \theta, \varepsilon) = \partial_y \widehat{K}_1(0, 0, \theta, \varepsilon) = 0$.

If we define $\widehat{A}_0 = \|\widehat{K}_0\|_{r(1-\eta)^3,0}$ *and* $\widehat{A}_1 = \|\widehat{K}_1\|_{r(1-\eta)^3,\widehat{\rho}}$, *the following bounds hold:*

$$\widehat{A}_0 \leq A_0 + \varepsilon_0^m \frac{A_1}{\eta}, \qquad \widehat{A}_1 \leq C \frac{A_0 A_1^2}{\lambda^2 r^4 \delta^4 \eta^8},$$

where C is a constant that only depends on ω. The generating function is bounded by

$$\|S(x, Y, \theta, \varepsilon)\|_{r(1-\eta), \widehat{\rho}} \leq \frac{BA_1}{\lambda \delta^2 \eta^2}, \tag{8.7}$$

where the constant B only depends on ω.

Finally, in order to apply this lemma, ε_0 must satisfy (8.3), where M is the bound that appears in (8.7).

Proof. In order to facilitate the reading, the proof has been split in several parts.

1.- *Rearranging the initial Hamiltonian.*

Let us start by expanding K_1 in Taylor series with respect to x and y,

$$K_1(x, y, \theta, \varepsilon) = \sum_{i+j \geq 2} a_{ij}(\theta, \varepsilon) x^i y^j,$$

where $a_{ij}(\theta, \varepsilon)$ is a periodic function of θ, that can be expanded in Fourier series,

$$a_{ij}(\theta, \varepsilon) = \sum_{k \in \mathbb{Z}^2} a_{ij}^k(\varepsilon) e^{k \cdot \theta \sqrt{-1}}.$$

Moreover, the following bounds hold:

$$\|a_{ij}\|_\rho \leq \frac{A_1}{r^{i+j}}, \qquad |a_{ij}^k| \leq \frac{A_1}{r^{i+j}} e^{-|k_1|\rho_1 - |k_2|\rho_2}, \qquad 0 < \varepsilon \leq \varepsilon_0.$$

Let us write K_1 in the following form:

$$K_1(x, y, \theta, \varepsilon) = \sum_{i+j \geq 2} \sum_{k \in \mathbb{Z}^2} a_{ij}^k(\varepsilon) (xy)^j x^{i-j} e^{k \cdot \theta \sqrt{-1}}.$$

Now, we define $\ell = i - j$. The conditions $i \geq 0$, $j \geq 0$ and $i + j \geq 2$ are simply replaced by $\ell + j \geq 0$, $j \geq 0$ and $2\ell + j \geq 2$ in terms of ℓ and j, so

$$K_1(x, y, \theta, \varepsilon) = \sum_{k \in \mathbb{Z}^2, \, \ell \in \mathbb{Z}} \left[\sum_{j \geq \max\{0, 1-\ell/2, -\ell\}} a_{\ell+j, j}^k(\varepsilon) (xy)^j \right] x^\ell e^{k \cdot \theta \sqrt{-1}},$$

and defining $a^{k,\ell}(xy, \varepsilon)$ as the expression between brackets, one obtains

$$K_1(x, y, \theta, \varepsilon) = \sum_{k \in \mathbb{Z}^2, \, \ell \in \mathbb{Z}} a^{k,\ell}(xy, \varepsilon) x^\ell e^{k \cdot \theta \sqrt{-1}}.$$

For the moment, we will handle these expressions as formal series and, once the change of variables be (formally) computed we will check that everything is well defined on the suitable domains.

2.- *Computing the generating function.*

We will look for a change of variables given by a generating function of the type

$$P \cdot \theta + Yx + \varepsilon^m S(x, Y, \theta, \varepsilon).$$

So, the corresponding change is

$$\left.\begin{aligned} p &= P + \varepsilon^m \frac{\partial S}{\partial \theta}, & \theta &= \theta, \\ y &= Y + \varepsilon^m \frac{\partial S}{\partial x}, & X &= x + \varepsilon^m \frac{\partial S}{\partial Y}. \end{aligned}\right\} \tag{8.8}$$

We want that this change transforms (8.5) into (8.6), so we insert (8.8) into (8.5), asking the result to be equal to (8.6):

$$\begin{aligned} \frac{\omega}{\varepsilon} \cdot P + \varepsilon^m \frac{\omega}{\varepsilon} \cdot \frac{\partial S}{\partial \theta} + K_0 \left(x \left(Y + \varepsilon^m \frac{\partial S}{\partial x} \right), \varepsilon \right) + \varepsilon^m K_1 \left(x, Y + \varepsilon^m \frac{\partial S}{\partial x}, \theta, \varepsilon \right) \\ = \frac{\omega}{\varepsilon} \cdot P + \widehat{K}_0 \left(\left(x + \varepsilon^m \frac{\partial S}{\partial Y} \right) Y, \varepsilon \right) + \varepsilon^{2m} \widehat{K}_1 \left(x + \varepsilon^m \frac{\partial S}{\partial Y}, Y, \theta, \varepsilon \right). \end{aligned}$$

This equation is automatically satisfied at order 0 in ε, if we choose $\widehat{K}_0(z,\varepsilon) = K_0(z,\varepsilon) + O(\varepsilon^m)$. If we take the terms of order ε^m and we ask them to cancel out we obtain an equation that determines S:

$$\frac{\omega}{\varepsilon} \cdot \frac{\partial S}{\partial \theta} + \lambda_0(xY) \left[x \frac{\partial S}{\partial x} - Y \frac{\partial S}{\partial Y} \right] + K_1(x, Y, \theta, \varepsilon) = 0, \tag{8.9}$$

where $\lambda_0(z,\varepsilon) = \partial_z K_0(z,\varepsilon)$. We will try to solve this equation in the sense of the formal series. We look for S of the same type as K_1:

$$S(x, Y, \theta, \varepsilon) = \sum_{k \in \mathbb{Z}^2,\, \ell \in \mathbb{Z}} s^{k,\ell}(xY, \varepsilon) x^\ell e^{k \cdot \theta \sqrt{-1}},$$

being

$$s^{k,\ell}(xY, \varepsilon) = \sum_{j \geq \max\{0, 1-\ell/2, -\ell\}} s^k_{\ell+j,j}(\varepsilon)(xY)^j.$$

The next step is to substitute S into (8.9), to obtain

$$\begin{aligned} \sum_{k \in \mathbb{Z}^2,\, \ell \in \mathbb{Z}} \sqrt{-1} s^{k,\ell}(xY, \varepsilon) x^\ell k \cdot \frac{\omega}{\varepsilon} e^{k \cdot \theta \sqrt{-1}} \\ + \lambda_0(xY) \sum_{k \in \mathbb{Z}^2,\, \ell \in \mathbb{Z}} \ell s^{k,\ell}(xY, \varepsilon) x^\ell e^{k \cdot \theta \sqrt{-1}} + \sum_{k \in \mathbb{Z}^2,\, \ell \in \mathbb{Z}} a^{k,\ell}(xY, \varepsilon) x^\ell e^{k \cdot \theta \sqrt{-1}} = 0. \end{aligned}$$

Equating terms, we obtain the following set of equations:

$$\sqrt{-1} k \cdot \frac{\omega}{\varepsilon} s^{k,\ell}(xY, \varepsilon) + \lambda_0(xY, \varepsilon) \ell s^{k,\ell}(xY, \varepsilon) + a^{k,\ell}(xY, \varepsilon) = 0, \quad k \in \mathbb{Z}^2,\ \ell \in \mathbb{Z}.$$

So, if $k \neq 0$ or $\ell \neq 0$, the generating function S is defined (formally) by the coefficients

$$s^{k,\ell}(xY, \varepsilon) = -\frac{\varepsilon a^{k,\ell}(xY, \varepsilon)}{\varepsilon \ell \lambda_0(xY, \varepsilon) + \sqrt{-1} k \cdot \omega}.$$

Hence, this implies that we can not kill $a^{k,\ell}$ with $k=0$ and $\ell=0$ and then equation (8.9) has no solution. So, we solve equation (8.9) but with $a^{0,0}(xY,\varepsilon)$ instead of 0 in the right-hand side (this is not a problem because $a^{0,0}(xY,\varepsilon)$ is already in normal form). Since $s^{0,0}$ can be chosen arbitrarily, we simply take $s^{0,0}(xY,\varepsilon)=0$.

3.- *Bounds on the generating function.*

Let us reduce the analyticity strip with respect to θ from $\rho=(\rho_1,\rho_2)$ to $\widehat{\rho}=(\rho_1-\delta\sqrt{\varepsilon},\rho_2-\delta\sqrt{\varepsilon})$ and the analyticity ball with respect to x and y from r to $r(1-\eta)$.

Now, let us bound S:

$$\begin{aligned}\|S\|_{r(1-\eta),\widehat{\rho}} &\leq \sum_{\substack{k\in\mathbb{Z}^2,\,\ell\in\mathbb{Z}\\ |k|+|\ell|\neq 0}} \|s^{k,\ell}(xY,\varepsilon)x^\ell\|_{r(1-\eta),0}e^{(\rho_1-\delta\sqrt{\varepsilon})|k_1|+(\rho_2-\delta\sqrt{\varepsilon})|k_2|}\\ &\leq \sum_{\substack{k\in\mathbb{Z}^2,\,\ell\in\mathbb{Z}\\ |k|+|\ell|\neq 0}} \frac{\varepsilon\|a^{k,\ell}(xY,\varepsilon)x^\ell\|_{r(1-\eta),0}}{|\varepsilon\ell\lambda_0(xY)+\sqrt{-1}k\cdot\omega|}e^{(\rho_1-\delta\sqrt{\varepsilon})|k_1|+(\rho_2-\delta\sqrt{\varepsilon})|k_2|}.\end{aligned} \tag{8.10}$$

To bound $\|a^{k,\ell}(xy,\varepsilon)x^\ell\|_{r(1-\eta),0}$ we recall that

$$a^{k,\ell}(xY,\varepsilon)=\sum_{j\geq\max\{0,1-\ell/2,-\ell\}} a^k_{\ell+j,j}(xY)^j.$$

Hence,

$$\begin{aligned}\|a^{k,\ell}(xY,\varepsilon)x^\ell\|_{r(1-\eta),0} &\leq \sum_{j\geq\max\{0,1-\ell/2,-\ell\}} |a^k_{\ell+j,j}||xY|^j|x|^\ell\\ &\leq A_1e^{-\rho_1|k_1|-\rho_2|k_2|}(1-\eta)^\ell\sum_{j\geq\max\{0,1-\ell/2,-\ell\}}(1-\eta)^{2j}.\end{aligned}$$

Here we distinguish several possibilities, according to the values of ℓ:

a) $\ell\leq -2$, that implies $j\geq-\ell$. Since $\sum_{j\geq-\ell}(1-\eta)^{2j}<(1-\eta)^{-2\ell}/\eta$, one has

$$\|a^{k,\ell}(xY,\varepsilon)x^\ell\|_{r(1-\eta),0}\leq A_1e^{-\rho_1|k_1|-\rho_2|k_2|}(1-\eta)^{-\ell}/\eta.$$

b) $\ell=-1$, that implies $j\geq 2$. Since $\sum_{j\geq 2}(1-\eta)^{2j}<(1-\eta)^4/\eta$,

$$\|a^{k,\ell}(xY,\varepsilon)x^\ell\|_{r(1-\eta),0}\leq A_1e^{-\rho_1|k_1|-\rho_2|k_2|}(1-\eta)^3/\eta.$$

c) $\ell=0,1$ that implies $j\geq 1$. Since $\sum_{j\geq 1}(1-\eta)^{2j}<(1-\eta)^2/\eta$,

$$\|a^{k,\ell}(xY,\varepsilon)x^\ell\|_{r(1-\eta),0}\leq A_1e^{-\rho_1|k_1|-\rho_2|k_2|}(1-\eta)^{2+\ell}/\eta. \tag{8.11}$$

d) $\ell\geq 2$ that implies $j\geq 0$. Since $\sum_{j\geq 0}(1-\eta)^{2j}<1/\eta$,

$$\|a^{k,\ell}(xY,\varepsilon)x^\ell\|_{r(1-\eta),0}\leq A_1e^{-\rho_1|k_1|-\rho_2|k_2|}(1-\eta)^\ell/\eta.$$

Now, putting these bounds into (8.10), separating the cases $k=0$ and $k\neq 0$, and taking into account that $|\varepsilon\ell\lambda_0(xY)+\sqrt{-1}k\cdot\omega|\geq|k\cdot\omega|$ one obtains

$$\begin{aligned}\|S\|_{r(1-\eta),\widehat{\rho}} \leq & \sum_{\ell\neq 0}\frac{\|a^{0,\ell}(xY,\varepsilon)x^\ell\|_{r(1-\eta),0}}{|\ell|\lambda} \\ & + \sum_{k\neq 0}\sum_{\ell\in\mathbb{Z}}\frac{\varepsilon\|a^{k,l}(xY,\varepsilon)x^\ell\|_{r(1-\eta),0}e^{(\rho_1-\delta\sqrt{\varepsilon})|k_1|+(\rho_2-\delta\sqrt{\varepsilon})|k_2|}}{|\varepsilon\ell\lambda_0(xY)+\sqrt{-1}k\cdot\omega|}.\end{aligned}$$

To bound the first sum we start using a), b), c) and d) for the different values of ℓ, and then we apply the bound

$$\sum_{\ell\geq 2}\frac{(1-\eta)^\ell}{\ell}\leq|\ln\eta|-(1-\eta),$$

to obtain

$$\sum_{\ell\neq 0}\frac{\|a^{0,\ell}(xY,\varepsilon)x^\ell\|_{r(1-\eta),0}}{|\ell|\lambda}<\frac{2A_1(1+\lambda\beta)}{\lambda\delta^2\eta^2}.$$

With a similar scheme one can show that

$$\sum_{k\neq 0}\sum_{\ell\in\mathbb{Z}}\frac{\varepsilon\|a^{k,l}(xY,\varepsilon)x^\ell\|_{r(1-\eta),0}e^{(\rho_1-\delta\sqrt{\varepsilon})|k_1|+(\rho_2-\delta\sqrt{\varepsilon})|k_2|}}{|\varepsilon\ell\lambda_0(xY)+\sqrt{-1}k\cdot\omega|}<\frac{2\varepsilon A_1}{\eta^2}\sum_{k\neq 0}\frac{e^{-\delta\sqrt{\varepsilon}|k|}}{|k\cdot\omega|}.$$

As $\delta\sqrt{\varepsilon}<1$, we can apply Lemma 5 to control this sum. So, the final bound on the generating function is

$$\|S\|_{r(1-\eta),\widehat{\rho}}\leq\frac{2A_1(1+\lambda\beta)}{\lambda\delta^2\eta^2}\leq\frac{BA_1}{\lambda\delta^2\eta^2}. \tag{8.12}$$

where β only depends on ω and $B=2(1+\beta)$.

4.- *The transformed Hamiltonian.*

Up to now, we have found a function $S(x,Y,\theta,\varepsilon)$ such that equation (8.9) holds, but with $a^{0,0}(xY,\varepsilon)$ instead of 0 in the right-hand side:

$$\begin{aligned}\frac{\omega}{\varepsilon}\cdot\frac{\partial S}{\partial\theta}(x,Y,\theta,\varepsilon)+\lambda_0(xY,\varepsilon)\left[x\frac{\partial S}{\partial x}(x,Y,\theta,\varepsilon)-Y\frac{\partial S}{\partial Y}(x,Y,\theta,\varepsilon)\right] \\ +K_1(x,Y,\theta,\varepsilon)=a^{0,0}(xY,\varepsilon).\end{aligned} \tag{8.13}$$

This function S satisfies the bound (8.12) and it generates implicitly the canonical change (8.8), whose explicit expression is of the form

$$p=P+\varepsilon^m f_0(X,Y,\theta,\varepsilon),\ x=X+\varepsilon^m f_1(X,Y,\theta,\varepsilon),\ y=Y+\varepsilon^m f_2(X,Y,\theta,\varepsilon). \tag{8.14}$$

Applying this change to the initial Hamiltonian (8.5) and expanding it in power series one obtains

$$\begin{aligned}&\frac{\omega}{\varepsilon}\cdot P+\varepsilon^m\frac{\omega}{\varepsilon}\cdot f_0(X,Y,\theta,\varepsilon)+K_0(XY,\varepsilon)\\ &\quad+\varepsilon^m\lambda_0(XY,\varepsilon)(Xf_2(X,Y,\theta,\varepsilon)+Yf_1(X,Y,\theta,\varepsilon))+\varepsilon^m K_1(X,Y,\theta,\varepsilon)+O(\varepsilon^{2m}).\end{aligned}$$

Let us define $\varepsilon^{2m}\widetilde{K}_1(X,Y,\theta,\varepsilon)$ as the remainder $O(\varepsilon^{2m})$ in this formula. Then, $\varepsilon^{2m}\widetilde{K}_1$ can be obtained in terms of the initial Hamiltonian by computing the remainder of the corresponding Taylor expansion with respect to ε^m. Indeed, let us call R_1 the contribution to $\varepsilon^{2m}\widetilde{K}_1$ that comes from K_1 and R_2 the one that comes from K_0 (this is, $\varepsilon^{2m}\widetilde{K}_1 = R_1 + R_2$). Then, it is straightforward to check that

$$\begin{aligned} R_1 &= \varepsilon^{2m} f_1 \frac{\partial K_1}{\partial x}(X+\tau\varepsilon^m f_1, Y+\tau\varepsilon^m f_2) \\ &\quad + \varepsilon^{2m} f_2 \frac{\partial K_1}{\partial y}(X+\tau\varepsilon^m f_1, Y+\tau\varepsilon^m f_2), \end{aligned} \tag{8.15}$$

$$\begin{aligned} R_2 &= \frac{\varepsilon^{2m}}{2} K_0''(XY+\tau\varepsilon^m[Xf_2+Yf_1]+\tau^2\varepsilon^{2m}f_1f_2)(Xf_2+Yf_1+2\tau\varepsilon^m f_1 f_2)^2 \\ &\quad + \varepsilon^{2m} K_0'(XY+\tau\varepsilon^m[Xf_2+Yf_1]+\tau^2\varepsilon^{2m}f_1f_2)f_1f_2, \end{aligned} \tag{8.16}$$

where $0<\tau<1$. Moreover, let us define

$$\begin{aligned} R_p &= \varepsilon^m \frac{\omega}{\varepsilon}\cdot\left(f_0(X,Y,\theta,\varepsilon) - \frac{\partial S}{\partial \theta}(X,Y,\theta,\varepsilon)\right), \\ R_x &= \varepsilon^m \lambda_0(XY,\varepsilon) Y\left(\frac{\partial S}{\partial Y}(X,Y,\theta,\varepsilon) + f_1(X,Y,\theta,\varepsilon)\right), \\ R_y &= \varepsilon^m \lambda_0(XY,\varepsilon) X\left(f_2(X,Y,\theta,\varepsilon) - \frac{\partial S}{\partial x}(X,Y,\theta,\varepsilon)\right). \end{aligned}$$

So, with these notations and using (8.13) (replacing the value x by X), the Hamiltonian can be rewritten as

$$\frac{\omega}{\varepsilon}\cdot P + K_0(XY,\varepsilon) + \varepsilon^m a^{0,0}(XY,\varepsilon) + \varepsilon^{2m}\widehat{K}_1(X,Y,\theta,\varepsilon),$$

where $\varepsilon^{2m}\widehat{K}_1(X,Y,\theta,\varepsilon) = R_1 + R_2 + R_p + R_x + R_y$.

5.- *Bounds on the transformed Hamiltonian.*

The bound on $a^{0,0}$ has already been done in (8.11):

$$\|a^{0,0}\|_{r(1-\eta),0} \le \frac{A_1(1-\eta)^2}{\eta} < \frac{A_1}{\eta}.$$

The bounds on R_1 and R_2 are obtained bounding directly (8.15) and (8.16), since bounds on f_1 and f_2 are given by (8.4) but taking into account that the actual function $S(z_1,z_2,\theta,\varepsilon)$ is defined on $|z_1|\le r(1-\eta)$ and $|z_2|\le r(1-\eta)$. Of course, the constant M that appears in (8.4) must be replaced by the bound (8.12). The bounds on $\partial_x K_1$ and $\partial_y K_1$ are produced applying Cauchy estimates, reducing the analyticity domain of K_1 from r to $r(1-\eta)$ (we can reduce it more, but this is enough and it produces simpler bounds). Hence, for $|X|\le r(1-\eta)^3$, $|Y|\le r(1-\eta)^3$, $|\operatorname{Im}\theta_i|\le\widehat{\rho}_i$, $i=1,2$ and $0<\varepsilon\le\varepsilon_0$ (this is the domain where the change is defined, see Lemma 6) we have

$$\|R_1\|_{r(1-\eta)^3,\widehat{\rho}} \le \varepsilon^{2m}\frac{4BA_1^2}{\lambda\delta^2\eta^4 r^2}.$$

Similarly (but with a little more work), one can derive the corresponding bound for R_2:

$$\|R_2\|_{r(1-\eta)^3,\widehat{\rho}} \le \varepsilon^{2m} \frac{72B^2A_0A_1^2}{\lambda^2 r^4\delta^4\eta^8}.$$

To bound R_p, R_x and R_y we note that

$$\begin{aligned} f_0(X,Y,\theta,\varepsilon) &= \frac{\partial S}{\partial\theta}(x(X,Y,\theta,\varepsilon),Y,\theta,\varepsilon), && (8.17)\\ f_1(X,Y,\theta,\varepsilon) &= -\frac{\partial S}{\partial Y}(x(X,Y,\theta,\varepsilon),Y,\theta,\varepsilon), && (8.18)\\ f_2(X,Y,\theta,\varepsilon) &= \frac{\partial S}{\partial x}(x(X,Y,\theta,\varepsilon),Y,\theta,\varepsilon), && (8.19) \end{aligned}$$

where $x(X,Y,\theta,\varepsilon)$ is the change of variables for the x coordinate given in (8.14). Our purpose is to apply the mean value Theorem plus (8.4) to derive the desired bounds.

Before continuing, let us give a bound to be used later:

$$\left\| \frac{\omega}{\varepsilon} \cdot \frac{\partial^2 S}{\partial x\partial\theta} \right\| \le C\frac{A_0A_1}{r^4\lambda\delta^2\eta^5}, \tag{8.20}$$

where C is a constant that only depends on the initial Hamiltonian. This has been obtained from (8.13), differentiating both sides with respect to x and then applying Cauchy estimates. Now it is not difficult to produce bounds for R_p, R_x and R_y, applying the mean value theorem:

$$\begin{aligned} \|R_p\|_{r(1-\eta)^3,\widehat{\rho}} &\le C_p\varepsilon^{2m}\frac{A_0A_1^2}{r^4\lambda^2\delta^4\eta^8},\\ \|R_x\|_{r(1-\eta)^3,\widehat{\rho}} &\le C_x\varepsilon^{2m}\frac{A_0A_1^2}{r^2\lambda^2\delta^4\eta^8},\\ \|R_y\|_{r(1-\eta)^3,\widehat{\rho}} &\le C_y\varepsilon^{2m}\frac{A_0A_1^2}{r^2\lambda^2\delta^4\eta^8}, \end{aligned}$$

where constants C_p, C_x and C_y only depend on the initial Hamiltonian. Of course, (8.20) has been used to bound $\|R_p\|_{r(1-\eta)^3}$. The final bound is obtained by taking the biggest factors of these two bounds (we recall that $r \le 1$, $\lambda \le 1$ and $A_0 \ge 1$). □

8.3 Proof of Theorem 5

The first step is to transform the autonomous Hamiltonian $h_0(x,y,\varepsilon)$ into its normal form. This can be done because it is an integrable Hamiltonian. The domain of definition of h_0 may be reduced, but in that case we rename this new domain in order to keep the same notation.

Now, assume we have done n changes of variables like the ones of Lemma 7, so the Hamiltonian has the form

$$H^{(n)}(x,y,\theta,p,\varepsilon) = \frac{\omega}{\varepsilon}\cdot p + H_0^{(n)}(xy,\varepsilon) + \varepsilon^{q2^n} H_1^{(n)}(x,y,\theta,\varepsilon), \tag{8.21}$$

where we have kept the same notation for the variables. This Hamiltonian is defined on some set $|x| \le r_n$, $|y| \le r_n$, $|\operatorname{Im}\theta_i| \le \rho_i^{(n)}$, $i = 1, 2$ and $0 < \varepsilon \le \varepsilon_0$. On this set, we define the following constants (bounds):

$$A_0^{(n)} = \|K_0\|_{r_n}, \qquad A_1^{(n)} = \|K_1\|_{r_n, \rho^{(n)}}, \qquad \lambda_n = \inf_{|z| \le r_n^2} |\partial_z K_0(z, \varepsilon)|.$$

Now, let us define the domains. Let δ_n be $\delta_0/(n+1)^2$, with $\delta_0 = 6/\pi^2$. Then, the reduction of the analyticity strip with respect to θ done to $H^{(n)}$ in order to compute the generating function will be $\delta_n\sqrt{\varepsilon}$ (see Lemma 7). Note that in this way the total reduction of the domain is exactly $\sqrt{\varepsilon}$. For the domain with respect to the spatial variables x and y we define the sequence $\{r_n\}$ as $r_n = r_{n-1}(1 - \eta_{n-1})^3$, where $\eta_n = 1 - \exp(-\overline{\eta}/(n+1)^2)$ and $\overline{\eta} = (2\ln 2)/\pi^2$. As the total reduction of domain is given by

$$\left[\prod_{n \ge 1} (1 - \eta_{n-1})\right]^3,$$

one can check that, with this selection of η_n and $\overline{\eta}$, this makes a reduction from r_0 to $r_0/2$. From Lemma 7 we obtain the following inequalities:

$$A_0^{(n)} \le A_0^{(n-1)} + \varepsilon_0^{q2^{n-1}} \frac{A_1^{(n-1)}}{\eta_{n-1}}, \tag{8.22}$$

$$A_1^{(n)} \le C \frac{A_0^{(n-1)} \left(A_1^{(n-1)}\right)^2}{\lambda_{n-1}^2 r_{n-1}^4 \delta_{n-1}^4 \eta_{n-1}^8}, \tag{8.23}$$

$$\lambda_n \ge \lambda_{n-1} - \varepsilon_0^{q2^{n-1}} \frac{A_1^{(n-1)}}{\eta_{n-1}}. \tag{8.24}$$

Let us assume that $\lambda_i \ge \lambda_0/2$ and $A_0^{(i)} \le 2A_0^{(0)}$, $i = 1, \ldots, n-1$. Then, from (8.23) one obtains

$$A_1^{(n)} \le E n^{24} \left(A_1^{(n-1)}\right)^2,$$

where E is a constant that only depends on the initial Hamiltonian. Taking logarithms to both sides of the expression above one gets (see [JS92, lemma 5] for the details):

$$A_1^{(n)} \le E^{1+2+\cdots+2^{n-1}} \left[\prod_{i=0}^{n-1} (n-i)^{2^i}\right]^{24} \left(A_1^{(0)}\right)^{2^n} \le \frac{1}{E}\left[\left(\frac{5}{3}\right)^{24} E A_1^{(0)}\right]^{2^n}.$$

Now it is easy to finish the proof. It is immediate that there exist a sufficiently small (but different from zero) value of ε_0 such that

$$\varepsilon_0^{q2^n} A_1^{(n)} \le \frac{D}{80}(1/2)^{2^n}, \tag{8.25}$$

where $D = \max\{A_0, \lambda_0/2\}$. This value ε_0 does not depend on n and only depends on the initial Hamiltonian. It has been chosen in this way because it satisfies

$$\sum_{n\geq 1} \varepsilon_0^{q2^{n-1}} \frac{A_1^{(n-1)}}{\eta_{n-1}} \leq D.$$

Then, from (8.22) and (8.24) one gets $\lambda_n \geq \lambda_0/2$ and $A_0^{(n)} \leq 2A_0^{(0)}$. Then, using induction, the proof of the convergence of the normal form is finished.

To show the convergence of the sequence of changes of variables we use (8.4), where the bound M on the generating function is given by (8.7). Then, at step n ($n \geq 1$), we are applying a change of variables whose distance to the identity is bounded by

$$\varepsilon_0^{q2^{n-1}} \frac{BA_1^{(n-1)}}{r_{n-1}\lambda_{n-1}\delta_{n-1}^2\eta_{n-1}^3}.$$

Now, using (8.25), the convergence of the sequence of changes of variables is straightforward. □

9 Proof of the Extension Theorem

Before proceeding to the proof of this theorem, we will introduce some notations as well as some auxiliary lemmas.

In the sequel, $\alpha > 0$, t_0, $T_0 > 0$, will be the real parameters introduced in the Extension Theorem, T the *complex* parameter in the strip $|\mathrm{Im}\, T| \leq \pi/2 - \varepsilon^\alpha$ and t will be the *real* time. K will denote a generic positive constant independent of ε, and τ will denote $|t + T - \pi i/2|$.

In order to bound the solution $(x(t), y(t))$, we will compare it with the homoclinic solution of the unperturbed system, and thus we introduce the functions:

$$\xi(t) := x(t) - x_0(t+T), \quad \eta(t) := \dot{\xi}(t) = y(t) - y_0(t+T),$$

which satisfy the system of differential equations with respect to the variable t:

$$\begin{aligned} \dot{\xi} &= \eta, \\ \dot{\eta} &= \sin(x_0(t+T)+\xi) - \sin(x_0(t+T)) \\ &\quad + \varepsilon^p \sin(x_0(t+T)+\xi)\cdot m(\theta_1(t), \theta_2(t)). \end{aligned}$$

In order to study this system is very convenient to write it as:

$$\dot{z} = A(t+T)z + \varepsilon^p G(x_0(t+T), t) + F(\xi, t+T, t), \tag{9.1}$$

where $z = (\xi, \eta)^\top$, $A(u)$ is the matrix

$$A(u) = \begin{pmatrix} 0 & 1 \\ \cos(x_0(u)) & 0 \end{pmatrix},$$

and the functions $G = (0, g)^\top$ and $F = (0, f)^\top$, that depend also on ε, are given by:

$$\begin{aligned} g(x,t) &= \sin x \cdot m(\theta_1(t), \theta_2(t)), \\ f(\xi, u, t) &= \sin(x_0(u) + \xi) - \sin(x_0(u)) - \cos(x_0(u)) \cdot \xi \\ &\quad + \varepsilon^p \left[g(x_0(u) + \xi, t) - g(x_0(u), t) \right]. \end{aligned} \tag{9.2}$$

From the initial condition (4.1), our goal is to bound solutions $z(t)$ of system (9.1) with $z(t_0) = \mathrm{O}(\varepsilon^{p-s})$. To this purpose, first of all we seek for a fundamental matrix of the corresponding homogeneous linear system

$$\frac{dz}{du} = A(u) z, \tag{9.3}$$

which can be integrated using the fact that $(y_0(u), \dot{y}_0(u))$ is a solution of equation (9.3). Another independent solution can be obtained in the form $\xi(u) = y_0(u) W(u)$, $\eta(u) = \dot{\xi}(u)$, with

$$W(u) = \int_b^u \frac{d\sigma}{y_0(\sigma)^2},$$

b being an arbitrary complex number. It is very important to choose adequately this parameter b to get a function $W(u)$ as regular as possible, near the singularities of y_0.

We will consider first the case $0 \le \operatorname{Im} T \le \pi/2$ and we choose $b = \pi i/2$. In this way, at the point $u = \pi i/2$, since $y_0(u)$ has a simple pole, $W(u)$ has a triple zero and $y_0(u) W(u)$ has a double zero.

Introducing:

$$\begin{aligned} \Psi(u) &= y_0(u) = \dot{x}_0(u), \\ \Phi(u) &= y_0(u) W(u) = \Psi(u) W(u), \end{aligned}$$

a fundamental matrix of the linear equation (9.3) is

$$M(u) = \begin{pmatrix} \Psi(u) & \Phi(u) \\ \Psi'(u) & \Phi'(u) \end{pmatrix},$$

which has determinant 1: $\Psi(u)\Phi'(u) - \Phi(u)\Psi'(u) = y_0(u)^2 W'(u) = 1$.

Expanding the functions Ψ, Φ near $u = \pi i/2$ we get:

$$\begin{aligned} \Psi(u) &= \frac{-2i}{(u - \pi i/2)} \left(1 + \mathrm{O}((u - \pi i/2)^2)\right), \\ \Phi(u) &= \frac{i}{6}(u - \pi i/2)^2 \left(1 + \mathrm{O}((u - \pi i/2)^2)\right), \end{aligned} \tag{9.4}$$

and that means that the fundamental matrix $M(u)$ behaves near $u = \pi i/2$ as:

$$\begin{pmatrix} \dfrac{-2i}{(u - \pi i/2)} & \dfrac{i}{6}(u - \pi i/2)^2 \\ \dfrac{2i}{(u - \pi i/2)^2} & \dfrac{i}{3}(u - \pi i/2) \end{pmatrix}.$$

In passing, from formula (9.4) one gets easily the following bounds that will be used later on.

Lemma 8 *For $0 \le \operatorname{Im} T \le \pi/2$ and $0 \le \tau := |t+T-\pi i/2| \le T_0$, the following bounds hold:*

$$\begin{aligned} |\Psi(t+T)| &\le \frac{K}{\tau}, & |\Phi(t+T)| &\le K\tau^2, \\ |\Psi'(t+T)| &\le \frac{K}{\tau^2}, & |\Phi'(t+T)| &\le K\tau. \end{aligned}$$

Since the fundamental solution $\varphi(u,\sigma)$ of the linear equation (9.3) satisfying the initial condition $\varphi(u,u) = \mathrm{Id}$ is given by $\varphi(u,\sigma) = M(u)M^{-1}(\sigma)$, we can now easily write the solution $z(t) = (\xi(t), \eta(t))$ of system (9.1) with initial condition $z(t_0)$ as:

$$z(t) = z_{\mathrm{in}}(t_0,t) + \int_{t_0}^{t} \varphi(t+T,\sigma+T) N(\xi(\sigma),\sigma+T,\sigma)\, d\sigma,$$

with

$$z_{\mathrm{in}}(t_0,t) = \varphi(t+T,t_0+T)z(t_0) = M(t+T)M^{-1}(t_0+T)z(t_0),$$

and

$$N(\xi,u,t) = \varepsilon^p G(x_0(u),t) + F(\xi,u,t).$$

The integral equation for $z(t)$ can be also written as

$$\begin{aligned} z(t) &= M(t+T)M^{-1}(t_0+T)z(t_0) \\ &\quad + M(t+T)\int_{t_0}^{t} M^{-1}(\sigma+T) N\left(\xi(\sigma),\sigma+T,\frac{\sigma}{\varepsilon}\right) d\sigma \\ &= z^1(t) + M(t+T)\int_{t_0}^{t} M^{-1}(\sigma+T) F\left(\xi(\sigma),\sigma+T,\frac{\sigma}{\varepsilon}\right) d\sigma, \end{aligned}$$

with

$$\begin{aligned} z^1(t) &:= M(t+T) \left[M^{-1}(t_0+T)z(t_0) \right. \\ &\qquad \left. + \varepsilon^p \int_{t_0}^{t} M^{-1}(\sigma+T) G\left(x_0(\sigma+T),\frac{\sigma}{\varepsilon}\right) d\sigma \right]. \end{aligned}$$

Writing this equation for $z(t)$ in components, we obtain

$$\begin{aligned} \xi(t) &= \xi_1(t) - \Psi(t+T)\int_{t_0}^{t} \Phi(\sigma+T) f(\xi(\sigma),\sigma+T,\sigma)\, d\sigma \\ &\quad + \Phi(t+T)\int_{t_0}^{t} \Psi(\sigma+T) f(\xi(\sigma),\sigma+T,\sigma)\, d\sigma, \qquad (9.5) \\ \eta(t) = \dot\xi(t) &= \dot\xi_1(t) - \Psi'(t+T)\int_{t_0}^{t} \Phi(\sigma+T) f(\xi(\sigma),\sigma+T,\sigma)\, d\sigma \\ &\quad + \Phi'(t+T)\int_{t_0}^{t} \Psi(\sigma+T) f(\xi(\sigma),\sigma+T,\sigma)\, d\sigma, \qquad (9.6) \end{aligned}$$

where

$$\begin{aligned}
\xi_1(t) \;&= \Psi(t+T) \quad \Big[\Phi'(t_0+T)\xi(t_0) - \Phi(t_0+T)\eta(t_0) \\
&\qquad - \varepsilon^p \int_{t_0}^{t} \Phi(\sigma+T) g(x_0(\sigma+T),\sigma)\, d\sigma\Big] \\
&\quad + \Phi(t+T) \quad \Big[\Psi'(t_0+T)\xi(t_0) - \Psi(t_0+T)\eta(t_0) \\
&\qquad + \varepsilon^p \int_{t_0}^{t} \Psi(\sigma+T) g(x_0(\sigma+T),\sigma)\, d\sigma\Big]. \qquad (9.7)
\end{aligned}$$

We have now a suitable expression (9.5) for ξ, and we need to bound the functions f, g, defined in equations (9.2), that appear therein.

Lemma 9 *If $\varepsilon^\alpha \le \tau \le T_0 + \pi/2$, it follows that*

$$|g(x_0(t+T),t)| \le \frac{K\varepsilon^{-s}}{\tau^2}. \qquad (9.8)$$

Moreover, if $|\xi_j| \le \lambda/\tau^\beta \le 1$, $j=1,2$, then

$$|f(\xi_1,t+T,t) - f(\xi_2,t+T,t)| \le K\left(\frac{\lambda}{\tau^{\beta+2}} + \frac{\varepsilon^{p-s}}{\tau^2}\right)|\xi_1 - \xi_2|. \qquad (9.9)$$

Proof of the lemma. We first recall that, at $\pm\pi i/2$, $y_0(u) = 2/\cosh u$ has a simple pole, $\sin x_0(u) = \dot y_0(u)$ has a double pole, as well as $\cos x_0(u) = 1 - y_0(u)^2/2$. In the bound of g it appears also the bound (1.7) of $m(\theta_1,\theta_2)$ for the complex values (1.8) of θ_1,θ_2. In order to bound f it is enough to apply Taylor's Theorem to the function $\sin x$. $\Box$

Finally, here comes a technical lemma that will be needed later on. Its proof is straightforward (and can be found in [DS92, lemma 7.1] for $\beta = 3$).

Lemma 10 *Let t, t_0 real, T complex, such that*

$$|\mathrm{Im}\, T| < \pi/2, \quad -T_0 \le t_0 + \mathrm{Re}\, T \le t + \mathrm{Re}\, T \le T_0.$$

Then, given $\beta \in \mathbb{R}$, the following inequality holds:

$$\int_{t_0}^{t} \frac{d\sigma}{|\sigma + T - \pi i/2|^\beta} \le K \cdot \rho_{[t_0,t]}^{-(\beta-1)}(T),$$

with $K = K(T_0,\beta) > 0$, and

$$\rho_{[t_0,t]}^{-\beta}(T) := \begin{cases} \sup \dfrac{1}{|\sigma + T - \pi i/2|^\beta}, & \text{if } \beta \ne 0, \\ \sup |\ln(|\sigma + T - \pi i/2|)|, & \text{if } \beta = 0, \end{cases}$$

where the supremum is taken for $\sigma \in [t_0,t]$.

The proof of the Extension Theorem, for the moment for $T \in D^+ = \{T \in \mathbb{C} : 0 < \operatorname{Im} T \le \pi/2 - \varepsilon^\alpha\}$, is based on the next two propositions. In the first one, the solutions of system (9.1) with initial conditions (4.1) will be extended up to $t = t_1(T)$. In the second proposition, we take $t = t_1(T)$ as the initial time.

We divide the complex strip D^+ in two parts

$$\begin{aligned} D^{\rm up} &= \{T \in \mathbb{C} : \pi/2 - \varepsilon^{2\alpha/3} \le \operatorname{Im} T \le \pi/2 - \varepsilon^\alpha\}, &(9.10)\\ D^{\rm down} &= \{T \in \mathbb{C} : 0 \le \operatorname{Im} T \le \pi/2 - \varepsilon^{2\alpha/3}\} &(9.11)\end{aligned}$$

and define the separation point $t_1(T)$ by

$$t_1(T) + \operatorname{Re} T = \begin{cases} \varepsilon^{2\alpha/3}, & \text{for} \quad T \in D^{\rm up}, \\ 0, & \text{for} \quad T \in D^{\rm down}. \end{cases} \tag{9.12}$$

Proposition 1 *Let $z = (\xi(t), \eta(t))$ a solution of system (9.1) with initial conditions satisfying*

$$|\xi(t_0)| \le C\varepsilon^{p-s}, \quad |\eta(t_0)| \le C\varepsilon^{p-s}. \tag{9.13}$$

Then, if $p - s - 2\alpha > 0$, there exists $\varepsilon_0 > 0$ such that, for $0 < \varepsilon < \varepsilon_0$, $(\xi(t), \eta(t))$ can be extended for $t \in [t_0, t_1(T)]$ and satisfies there the following bound:

$$|\Phi'(t+T)\xi(t)| + |\Phi(t+T)\eta(t)| \le K\varepsilon^{p-s}. \tag{9.14}$$

Proof of the proposition. We shall use the method of successive approximations. We begin the iteration process with $\xi_0(t) = 0$, and consider, for $n \ge 0$, the recurrence suggested by equation (9.5):

$$\begin{aligned} \xi_{n+1}(t) \;&= \xi_1(t) \quad - \Psi(t+T)\int_{t_0}^t \Phi(\sigma+T) f(\xi_n(\sigma), \sigma+T, \sigma)\, d\sigma \\ &\qquad + \Phi(t+T)\int_{t_0}^t \Psi(\sigma+T) f(\xi_n(\sigma), \sigma+T, \sigma)\, d\sigma. \end{aligned} \tag{9.15}$$

The first iterate is $\xi_1(t)$, as given by equation (9.7), and can be bounded easily, using the initial conditions (9.13), and Lemmas 8, 9, and 10:

$$|\xi_1(t)| \le \frac{K}{\tau}\left[\varepsilon^{p-s} + \varepsilon^{p-s}\right] + K\tau^2\left[\varepsilon^{p-s} + \varepsilon^{p-s}\rho^{-2}_{[t_0,t]}(T)\right]. \tag{9.16}$$

An analogous bound for $\Phi'(t+T)\xi_1(t)$ follows immediately

$$|\Phi'(t+T)\xi_1(t)| \le K\left[\varepsilon^{p-s} + \varepsilon^{p-s}\right] + K\tau^3\left[\varepsilon^{p-s} + \varepsilon^{p-s}\rho^{-2}_{[t_0,t]}(T)\right].$$

Now, for $T \in D^{\rm down}$ we have $\rho^{-2}_{[t_0,t]}(T) \le K\tau^{-2\alpha}$, and for $T \in D^{\rm up}$ we have

$$\begin{aligned} \rho^{-2}_{[t_0,t]}(T) \le K\tau^{-2}, &\quad \text{for} \quad t_0 \le t \le -\operatorname{Re} T, \\ \rho^{-2}_{[t_0,t]}(T) \le K\varepsilon^{-2\alpha}, \quad \tau^3 \le K\varepsilon^{2\alpha}, &\quad \text{for} \quad -\operatorname{Re} T \le t \le t_1(T), \end{aligned}$$

and, consequently, in all the strip D^+: $T \in D^+$, and for $t \in [t_0, t_1(T)]$, it follows that

$$\tau^3 \rho_{[t_0,t]}^{-2}(T) \le K. \tag{9.17}$$

Let us remark that the value of $t_1(T)$ has been chosen just in order that bound (9.17) holds. In this way, we can bound $|\Phi'(t+T)\xi_1(t)|$ in a uniform way: $|\Phi'(t+T)\xi_1(t)| \le K\varepsilon^{p-s}$, or, in other words, $|\xi_1(t)| \le K\varepsilon^{p-s}/\tau$.

To begin the iteration process, we introduce the norm

$$\|\xi\| := \sup |\Phi'(t+T)\xi(t)|,$$

where the supremum is taken for $T \in D^+$ and $t \in [t_0, t_1(T)]$. The above bound on $\Phi'(t+T)\xi_1(t)$ reads now as

$$\|\xi_1\| \le K\varepsilon^{p-s}.$$

Assuming that $\|\xi_{n-1}\|, \|\xi_n\| \le K\varepsilon^{p-s}$, we now consider

$$\begin{aligned} \xi_{n+1}(t) - \xi_n(t) &= -\Psi(t+T)\int_{t_0}^{t} \Phi(\sigma+T)\,[f_n - f_{n-1}]\,d\sigma \\ &\quad + \Phi(t+T)\int_{t_0}^{t} \Psi(\sigma+T)\,[f_n - f_{n-1}]\,d\sigma, \end{aligned}$$

where f_k denotes $f(\xi_k(\sigma), \sigma+T, \sigma)$. Applying Lemmas 8, 9 (with $\lambda = K\varepsilon^{p-s}$ and $\beta = 1$) and 10, as well as inequality (9.17), we obtain

$$\begin{aligned} &|\Phi'(t+T)(\xi_{n+1}(t) - \xi_n(t))| \\ &\le K\int_{t_0}^{t} \tau(\sigma)^2 \frac{\varepsilon^{p-s}}{\tau(\sigma)^3}\,|\xi_n(\sigma) - \xi_{n-1}(\sigma)|\,d\sigma \\ &\quad + K\tau^3\int_{t_0}^{t} \frac{1}{\tau(\sigma)} \cdot \frac{\varepsilon^{p-s}}{\tau(\sigma)^3}\,|\xi_n(\sigma) - \xi_{n-1}(\sigma)|\,d\sigma \\ &\le K\left[\int_{t_0}^{t} \frac{\varepsilon^{p-s}d\sigma}{\tau(\sigma)^2} + \tau^3\int_{t_0}^{t} \frac{\varepsilon^{p-s}d\sigma}{\tau(\sigma)^5}\right]\|\xi_n - \xi_{n-1}\| \\ &\le K\left[\varepsilon^{p-s}\rho_{[t_0,t]}^{-1}(T) + \varepsilon^{p-s}\tau^3\rho_{[t_0,t]}^{-4}(T)\right]\|\xi_n - \xi_{n-1}\| \\ &\le K\varepsilon^{p-s-2\alpha}\,\|\xi_n - \xi_{n-1}\|, \end{aligned}$$

where $\tau(\sigma)$ denotes $|\sigma + T - \pi i/2|$. Since $p - s - 2\alpha > 0$, if we choose now $\varepsilon_0 = \varepsilon_0(K, p-s-2\alpha)$ small enough, it follows, by induction, that the following inequalities

$$\|\xi_k\| \le 2\,\|\xi_1\| \le K\varepsilon^{p-s}, \quad \|\xi_{k+1} - \xi_k\| \le \frac{1}{2}\,\|\xi_k - \xi_{k-1}\|, \tag{9.18}$$

are valid for $k \ge 1$, $0 < \varepsilon \le \varepsilon_0$, and consequently $(\xi_k)_{n\ge 0}$ converges uniformly for $T \in D^+$ and $t \in [t_0, t_1(T)]$ to the first component $\xi(t)$ of a solution of system (9.1), satisfying

$$|\Phi'(t+T)\xi(t)| \le K\varepsilon^{p-s}.$$

For the second component $\eta(t)$, we simply use its integral equation (9.6), and it is straightforward to check that

$$|\Phi(t+T)\eta(t)| \leq K\varepsilon^{p-s},$$

and consequently the bound (9.14) is proved. □

From bound (9.14) we get the following global estimates

$$|\xi(t)| \leq K\varepsilon^{p-s-\alpha}, \quad |\eta(t)| \leq K\varepsilon^{p-s-2\alpha},$$

for $t \in [t_0, t_1(T)]$. On the final point $t = t_1(T)$, bound (9.14) gives a better estimate

$$|\xi(t_1(T))| \leq K\frac{\varepsilon^{p-s}}{\tau_1} \leq K\varepsilon^{p-s-2\alpha/3}, \quad |\eta(t_1(T))| \leq K\frac{\varepsilon^{p-s}}{\tau_1^2} \leq K\varepsilon^{p-s-4\alpha/3}, \tag{9.19}$$

where we have denoted $\tau_1 = |t_1(T) + T - \pi i/2|$, and we have used that $\tau_1 \geq \varepsilon^{2\alpha/3}$. These are the initial conditions for the next proposition.

Proposition 2 *Let $(\xi(t), \eta(t))$ a solution of system (9.1) with initial conditions satisfying (9.19). Then, if $p - s - 2\alpha > 0$, there exists $\varepsilon_0 > 0$ such that, for $0 < \varepsilon < \varepsilon_0$, $(\xi(t), \eta(t))$ can be extended for $t \in [t_1(T), T_0 - \mathrm{Re}\, T]$ and satisfies there the following bound:*

$$|\Psi'(t+T)\xi(t)| + |\Psi(t+T)\eta(t)| \leq K\varepsilon^{p-s-2\alpha}. \tag{9.20}$$

Proof of the proposition. We shall use exactly the same method of successive approximations as in proposition 1, but replacing the initial condition t_0 by $t_1(T)$:

$$\begin{aligned} \xi_{n+1}(t) \ = \xi_1(t) \quad & - \Psi(t+T)\int_{t_1(T)}^{t} \Phi(\sigma+T) f(\xi_n(\sigma), \sigma+T, \sigma)\, d\sigma \\ & + \Phi(t+T)\int_{t_1(T)}^{t} \Psi(\sigma+T) f(\xi_n(\sigma), \sigma+T, \sigma)\, d\sigma. \end{aligned} \tag{9.21}$$

The first iteration gives $\xi_1(t)$ as provided by equation (9.7), but with $t_1(T)$ instead of t_0. Proceeding like in proposition 1, but using now the initial conditions (9.19), we can bound the first iterate $\xi_1(t)$ as in (9.16):

$$|\xi_1(t)| \ \leq \ \frac{K}{\tau}\left[\varepsilon^{p-s} + \varepsilon^{p-s}\right] + K\tau^2\left[\frac{\varepsilon^{p-s}}{\tau_1^3} + \varepsilon^{p-s}\rho^{-2}_{[t_1(T),t]}(T)\right].$$

Now the following inequalities hold

$$\rho^{-\beta}_{[t_1(T),t]}(T) \leq K\tau_1^{-\beta}, \quad \rho^{0}_{[t_1(T),t]}(T) \leq K\left|\ln \tau_1\right|, \quad \rho^{\beta}_{[t_1(T),t]}(T) \leq K\tau^{\beta}, \tag{9.22}$$

and consequently we can bound $\Psi'(t+T)\xi_1(t)$ as

$$|\Psi'(t+T)\xi_1(t)| \leq K\left(\frac{\varepsilon^{p-s}}{\tau^3} + \frac{\varepsilon^{p-s}}{\tau_1^3}\right) \leq \varepsilon^{p-s-2\alpha},$$

where we have used that $\tau \geq \tau_1 \geq \varepsilon^{2\alpha/3}$. In view of this bound, we define now the norm

$$\|\xi\| := \sup |\Psi'(t+T)\xi(t)| \,,$$

with the supremum taken for for $T \in D^+$ and $t \in [t_1(T), T_0 - \operatorname{Re} T]$. With this new terminology we have proved that

$$\|\xi_1\| \leq K\varepsilon^{p-s-2\alpha},$$

and therefore

$$|\xi_1(t)| \leq K\varepsilon^{p-s-2\alpha}\tau^2.$$

For the successive iterates we apply Lemmas 8, 9 (with $\lambda = K\varepsilon^{p-s-2\alpha}$ and $\beta = -2$) and 10, as well as inequalities (9.22), obtaining

$$\begin{aligned}
&|\Psi'(t+T)(\xi_{n+1}(t) - \xi_n(t))| \\
\leq\ & \frac{K}{\tau^3}\int_{t_1(T)}^t \tau(\sigma)^2 \left(\varepsilon^{p-s-2\alpha} + \frac{\varepsilon^{p-s}}{\tau(\sigma)^2}\right) |\xi_n(\sigma) - \xi_{n-1}(\sigma)| \, d\sigma \\
&+ K\int_{t_1(T)}^t \frac{1}{\tau(\sigma)} \left(\varepsilon^{p-s-2\alpha} + \frac{\varepsilon^{p-s}}{\tau(\sigma)^2}\right) |\xi_n(\sigma) - \xi_{n-1}(\sigma)| \, d\sigma \\
\leq\ & K\left[\frac{1}{\tau^3}\int_{t_1(T)}^t \left(\varepsilon^{p-s-2\alpha}\tau(\sigma)^4 + \varepsilon^{p-s}\tau(\sigma)^2\right) d\sigma \right. \\
&\left. + \int_{t_1(T)}^t \left(\varepsilon^{p-s-2\alpha}\tau(\sigma) + \frac{\varepsilon^{p-s}}{\tau(\sigma)}\right) d\sigma\right] \|\xi_n - \xi_{n-1}\| \\
\leq\ & K\left[\frac{\varepsilon^{p-s-2\alpha}}{\tau^3}\rho^5 + \frac{\varepsilon^{p-s}}{\tau^3}\rho^3 + \varepsilon^{p-s-2\alpha}\rho^2 + \varepsilon^{p-s}\rho^0\right] \|\xi_n - \xi_{n-1}\| \\
\leq\ & K\varepsilon^{p-s-2\alpha} \|\xi_n - \xi_{n-1}\| \,,
\end{aligned}$$

where $\tau(\sigma)$ has denoted $|\sigma + T - \pi i/2|$, and ρ^β has denoted $\rho^\beta_{[t_1(T),t]}(T)$. Since $p - s - 2\alpha > 0$, choosing now $\varepsilon_0 = \varepsilon_0(K, p-s-2\alpha)$ small enough, it follows by induction that for $n \geq 1$

$$\|\xi_n\| \leq K\varepsilon^{p-s-2\alpha}, \quad \|\xi_{n+1} - \xi_n\| \leq \frac{1}{2}\|\xi_n - \xi_{n-1}\| \,,$$

and consequently $(\xi_n)_{n\geq 0}$ converges uniformly for $T \in D^+$ and $t \in [t_1(T), T_0 - \operatorname{Re} T]$ to the first component $\xi(t)$ of a solution of system (9.1), satisfying the required bound.

As in proposition 2, we can now bound $\eta(t)$ from its integral equation (9.5), and we finally obtain bound (9.20). □

Proof of the Extension Theorem. First consider $0 \leq \operatorname{Im} T \leq \pi/2 - \varepsilon^\alpha$. Putting propositions 1 and 2 together, as well as the bound (9.20) produced by this last proposition, we immediately obtain the Extension Theorem for $0 \leq \operatorname{Im} T \leq \pi/2 - \varepsilon^\alpha$, with the required estimates.

For $-\pi/2 + \varepsilon^\alpha \leq \operatorname{Im} T \leq 0$ we only have to choose $b = -\pi i/2$ in the definition of $W(u)$, in order to get a second solution $\Phi(u)$ of the linear system (9.3) with a double zero at $u = -\pi i/2$. Lemma 8, as well as propositions 1 and 2 are also valid for $-\pi/2 + \varepsilon^\alpha \leq \operatorname{Im} T \leq 0$, and consequently the Extension Theorem follows for $|\operatorname{Im} T| \leq \pi/2 - \varepsilon^\alpha$. □

Acknowledgments

We are indebted to C. Simó for relevant discussions and remarks. Three of the authors (A. D, A. J. and T.M. S) have been partially supported by the Spanish grant DGICYT PB94–0215, the EC grant ERBCHRXCT940460, and the Catalan grant CIRIT 1996SGR–00105. One of the authors (V. G.) was supported by a CICYT grant.

References

[Arn63] V.I. Arnold. Proof of A.N. Kolmogorov's theorem on the preservation of quasi-periodic motions under small perturbations of the Hamiltonian. *Russian Math. Surveys*, 18:9–36, 1963.

[Arn64] V.I. Arnold. Instability of dynamical systems with several degrees of freedom. *Soviet Math. Dokl.*, 5:581–585, 1964.

[BCF96] G. Benettin, A. Carati, and F. Fassò. On the conservation of adiabatic invariants for a system of coupled rotators. Preprint, 1996.

[BCG95] G. Benettin, A. Carati, and G. Gallavotti. A rigorous implementation of the Jeans-Landau-Teller approximation for adiabatic invariants. Preprint, archived in `mp_arc@math.utexas.edu`, #95-412, August 1995.

[CG94] L. Chierchia and G. Gallavotti. Drift and diffusion in phase space. *Ann. Inst. H. PoincaréPhys. Théor.*, 60(1):1–144, 1994.

[DGJS97a] A. Delshams, V.G. Gelfreich, A. Jorba, and T.M. Seara. Lower and upper bounds for the splitting of separatrices of the pendulum under a fast quasiperiodic forcing. *Electron. Res. Announ. Amer. Math. Soc.*, 3, 1997.

[DGJS97b] A. Delshams, V.G. Gelfreich, A. Jorba, and T.M. Seara. Splitting of separatrices for (fast) quasiperiodic forcing. In C. Simó, editor, *Hamiltonian Systems with Three or More Degrees of Freedom*, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. Held in S'Agaró, Spain, 29–30 June 1995. Kluwer Acad. Publ., Dordrecht, Holland, to appear in 1997.

[DJS91] C. Díez, A. Jorba, and C. Simó. A dynamical equivalent to the equilateral libration points of the real Earth-Moon system. *Celestial Mech.*, 50:13–29, 1991.

[DS92] A. Delshams and T.M. Seara. An asymptotic expression for the splitting of separatrices of the rapidly forced pendulum. *Comm. Math. Phys.*, 150:433–463, 1992.

[DS96] A. Delshams and T.M. Seara. Splitting of separatrices in Hamiltonian systems with one and a half degrees of freedom. Preprint, archived in `mp_arc@math.utexas.edu`, #96-537, October 1996.

[Fon93] E. Fontich. Exponentially small upper bounds for the splitting of separatrices for high frequency periodic perturbations. *Nonlinear Anal.*, 20(6):733–744, 1993.

[Fon95] E. Fontich. Rapidly forced planar vector fields and splitting of separatrices. *J. Differential Equations*, 119(2):310–335, 1995.

[Gal94] G. Gallavotti. Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review. *Rev. Math. Phys.*, 6(3):343–411, 1994.

[Gel90] V.G. Gelfreich. Separatrices splitting for the rapidly forced pendulum. (unpublished), 1990.

[Gel93] V.G. Gelfreich. Separatrices splitting for the rapidly forced pendulum. In S. Kuksin, V.F. Lazutkin, and J. Pöschel, editors, *Proceedings of the Dynamical Systems Semester*, pages 47–67. Held in St. Petersburg, Russia, 17–30 November, 1991. Birkhäuser, Basel-Boston-Stuttgart, 1993.

[Gel97] V.G. Gelfreich. Reference systems for splitting of separatrices. *Nonlinearity*, 10(1):175–193, 1997.

[GJMS91] G. Gómez, A. Jorba, J. Masdemont, and C. Simó. A quasiperiodic solution as a substitute of l_4 in the Earth-Moon system. In *Proceedings of the 3rd International Sysmposium on Spacecraft Flight Dynamics*, pages 35–41, ESTEC, Noordwijk, Holand, 1991. ESA Publications Division.

[HMS88] P. Holmes, J. Marsden, and J. Scheurle. Exponentially small splittings of separatrices with applications to KAM theory and degenerate bifurcations. *Contemp. Math.*, 81:213–244, 1988.

[JS92] A. Jorba and C. Simó. On the reducibility of linear differential equations with quasiperiodic coefficients. *J. Differential Equations*, 98:111–124, 1992.

[JS96] A. Jorba and C. Simó. On quasiperiodic perturbations of elliptic equilibrium points. *SIAM J. Math. Anal.*, 27(6):1704–1737, 1996.

[Khi63] A.Ya. Khintchine. *Continued Fractions*. P. Noordhoff, Ltd., Groningen, The Netherlands, 1963.

[Lan91] S. Lang. *Introduction to Diophantine Approximations*. Springer, New York, second edition, 1991.

[Laz84] V.F. Lazutkin. Splitting of separatrices for the Chirikov's standard map. Preprint VINITI No. 6372–84 (in Russian), 1984.

[Loc92] P. Lochak. Canonical perturbation theory via simultaneous approximation. *Russian Math. Surveys*, 47(6):57–133, 1992.

[Mos56] J. Moser. The analytic invariants of an area-preserving mapping near a hyperbolic fixed point. *Comm. Pure Appl. Math.*, 9:673–692, 1956.

[Nei84] A.I. Neishtadt. The separation of motions in systems with rapidly rotating phase. *J. Appl. Math. Mech.*, 48(2):133–139, 1984.

[Poi99] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste*, volume 1, 2, 3. Gauthier-Villars, Paris, 1892–1899.

[RW95] M. Rudnev and S. Wiggins. KAM theory near multiplicity one resonant surfaces in perturbations of a-priori stable hamiltonian systems. Preprint, archived in `mp_arc@math.utexas.edu`, #95-537, 1995.

[RW97] M. Rudnev and S. Wiggins. Existence of exponentially small separatrix splittings and homoclinic connections between whiskered tori in weakly hyperbolic near-integrable hamiltonian systems. Preprint, archived in `mp_arc@math.utexas.edu`, #97-4, 1997.

[Sim94] C. Simó. Averaging under fast quasiperiodic forcing. In J. Seimenis, editor, *Hamiltonian Mechanics: Integrability and Chaotic Behaviour*, volume 331 of *NATO Adv. Sci. Inst. Ser. B Phys.*, pages 13–34. Held in Toruń, Polland, 28 June–2 July 1993. Plenum, New York, 1994.

[Tre94] D.V. Treschev. Separatrix splitting for a pendulum with rapidly oscillating suspension point. Preprint 273, Centre de Recerca Matemàtica, Barcelona, December 1994. To appear in Russ. J. Math. Phys.