# Exponentially small separatrix splittings and almost invisible homoclinic bifurcations in some billiard tables ⋆

Rafael Ramírez-Ros

*Departament de Matemàtica Aplicada 1, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain*

**Abstract**

We present a numerical study of some billiard tables depending on a *perturbative parameter* $\epsilon \geq 0$ and a *hyperbolicity parameter* $h > 0$. These tables are ellipses for $\epsilon = 0$ and circumferences in the limit $h \to 0^+$. Elliptic billiard tables are integrable and have four separatrices connecting their hyperbolic two-periodic points.

These connections break up when $\epsilon > 0$. As $h \to 0^+$, the area of the main lobes of the resulting turnstile (which can be interpreted as the difference of the lengths of the symmetric primary homoclinic billiard trajectories) behaves like an exponential term $\epsilon\,\mathrm{e}^{-\pi^2/h}$ times an asymptotic series $\sum_{j\geq 0} \alpha_j^\epsilon h^{2j}$ such that $\alpha_0^\epsilon \neq 0$. This series is Gevrey-1 of type $\rho = 1/2\pi^2$, so that its Borel transform is convergent on a disk of radius $2\pi^2$. In the limit $\epsilon \to 0$, the series $\sum_{j\geq 0} \alpha_j^0 h^{2j}$ is an analytic function which can be explicitly computed with a discrete Melnikov method. The asymptotic series $\sum_{j\geq 0} \omega_j^\epsilon h^{2j}$ associated to the second exponential term $\epsilon\,\mathrm{e}^{-2\pi^2/h}$ has the same properties. Finally, we have detected some *almost invisible* homoclinic bifurcations that take place in an exponentially small region of the parameter space.

Our computations have been performed in multiple-precision arithmetic (namely, with several thousands decimal digits) and rely strongly on the expansion of the local invariant curves up to very high orders (namely, with several hundreds Taylor coefficients). Our programs have been written using the PARI system.

*Key words:* Billiards, separatrix splitting, exponentially small phenomena, numeric experiments, homoclinic bifurcations
*PACS:* 05.45.Pq, 42.65.Sf, 45.50.Tn
*1991 MSC:* 37J45, 37G25, 34E05, 34M37

⋆ Expanded version of a talk presented at the Workshop on Differential Equations dedicated to the memory of Vladimir F. Lazutkin (Saint-Petesburg, August 2002).

*Email address:* `rafael@vilma.upc.es` (Rafael Ramírez-Ros).
*URL:* `www.ma1.upc.es/~rafael/` (Rafael Ramírez-Ros).

## 1 Introduction

G. D. Birkhoff [2] introduced the problem of *convex billiard tables* as a way to describe the motion of a free particle inside a closed convex curve such that it reflects at the boundary according to the law "angle of incidence equals to angle of reflection". He realized that this billiard motion can be modeled by an area-preserving (in fact, twist) diffeomorphism defined on the annulus. If the curve is an ellipse, its billiard map has separatrices, and if it is close to a circumference, the square of its billiard map is close to the identity in a portion of the phase space. The splitting of separatrices of area-preserving maps close to the identity is one of the most paradigmatic fields related to *exponentially small* phenomena. The field reached its maturity with the work of V. F. Lazutkin. He gave an asymptotic formula for the splitting size in the *standard map* and provided the basic lines of a proof [15]. His work influenced strongly the research in the field and many papers are based on his ideas. Lazutkin's formula was fully proved fifteen years later by V. G. Gelfreich [10]. The current state-of-the-art of the field is described in [12]. See also §2.

Our goal is to present several numerical results about the exponentially small splitting of separatrices that takes place for the billiard maps associated to some perturbed almost circular ellipses. This question was posed more than ten years ago by M. B. Tabanov [23, §IX] in the case of a concrete quartic perturbation. The general case was posed in [4].

To keep the technicalities of this introduction to a minimum, we shall describe the main results of this paper outside of the frame of twist maps, but then we need to explain what is the length of certain kind of billiard trajectories. That is very clear for periodic trajectories, but we need to consider homoclinic ones.

The *chords* of a curve are the segments perpendicular to the curve at their ends. To any chord is associated a two-periodic trajectory. If this trajectory is hyperbolic, the chord is called *hyperbolic*. For instance, the major axis of an ellipse is a hyperbolic chord. (In this paper, ellipses never are circunferences.)

Let $\mathcal{T}$ be a billiard trajectory homoclinic to a hyperbolic chord of a curve $C$, whose bi-infinite sequence of impact points is $\ldots, c_{-1}, c_0, c_1, \ldots$. Let $\ell$ be the length of the chord. By hyperbolicity, the sequence of impact points tends at an exponential rate to the ends of the chord. Thus the series

$$\mathrm{Length}\mathcal{T} := \sum_{k\in\mathbb{Z}} \left(|c_{k+1} - c_k| - \ell\right)$$

converges to a negative quantity called the *(homoclinic) length* of $\mathcal{T}$.

Billiards inside ellipses are called *elliptic*. Elliptic billiards are integrable and have separatrices. The trajectories contained on the separatrices of an elliptic billiard correspond to rays passing through the foci of the ellipse and they converge to the major axis of the ellipse. Using this geometric characterization and a straightforward telescopic argument, one realizes that the length of *any* homoclinic trajectory inside an ellipse is equal to minus the focal distance of the ellipse. The fact that all the homoclinic lengths coincide is a direct consequence of the existence of the separatrices. We shall take the difference of lengths of some perturbed homoclinic trajectories as a symplectic invariant measure of the splitting of the separatrices under perturbations of the ellipse.

Now, to be more precise, we consider the perturbed ellipse

$$C_\epsilon = \left\{(x,y) \in \mathbb{R}^2 : x^2/a^2 + y^2/b^2 + \epsilon y^{2n}/\gamma^{2n} = 1\right\} \tag{1}$$

where $0 < b < a$, $\epsilon \geq 0$, $n \geq 2$, $e = (1 - b^2/a^2)^{1/2}$, and $\gamma = b/e$. Here $\epsilon$ is the *perturbative parameter*, $2n$ is the *degree of the perturbation*, $a$ and $b$ are the *semi-axes lengths* of the unperturbed ellipse $C_0$, and $e$ is the *eccentricity* of $C_0$. The diameter of $C_\epsilon$ has length $2a$, its ends are the points $(\pm a, 0)$ and is always hyperbolic. The *hyperbolicity parameter* $h > 0$ determined by

$$a/b = \cosh(h/2), \qquad a/\gamma = \sinh(h/2), \qquad e = \tanh(h/2)$$

quantifies how much hyperbolic it is, since $h$ is the *characteristic exponent* of the two-periodic trajectory along it, see §3.2. The intrinsic parameters of the problem we are dealing with are $\epsilon$ and $h$. Hence, we shall express any quantity in terms of them. The curve $C_\epsilon$ is the ellipse $x^2/a^2 + y^2/b^2 = 1$ for $\epsilon = 0$, and it tends to the circumference $x^2 + y^2 = a^2$ in the limit $h \to 0^+$ for any fixed $\epsilon$.

The smooth convex curve $C_\epsilon$ has two axial symmetries, which play a key rôle. A billiard trajectory inside $C_\epsilon$ is *x-axial* (respectively, *y-axial*) when it is symmetric with regard to the $x$-axis (respectively, $y$-axis). Inside $C_0$ there are four $x$-axial and four $y$-axial homoclinic billiard trajectories (see figure 2), which persist under symmetric perturbations [7]. Let $\mathcal{T}_\epsilon^+$ (respectively, $\mathcal{T}_\epsilon^-$) be any of these $y$-axial (respectively, $x$-axial) persistent homoclinic trajectories inside $C_\epsilon$. The following asymptotic expansion for the difference of homoclinic lengths $\Delta_{\mathrm{Length}} := \mathrm{Length}\mathcal{T}_\epsilon^- - \mathrm{Length}\mathcal{T}_\epsilon^+$ has been numerically established

$$\Delta_{\mathrm{Length}} \stackrel{\mathrm{as}}{=} a\epsilon\, \mathrm{e}^{-\pi^2/h} \sum_{j\geq 0} \alpha_j^\epsilon h^{2j} \qquad (h \to 0^+, \epsilon \text{ fixed}).$$

The symbol $\stackrel{\mathrm{as}}{=}$ has been used to emphasize the asymptotic nature of this series. In other words, if we retain only finitely many terms of the right-hand side. the error will be of the order of the first discarded term.

The dominant asymptotic coefficient $\alpha_0^\epsilon$ does not vanish in the range $0 \leq \epsilon \leq 1$ for $n = 2, 3, 4$. Therefore, the homoclinic lengths no longer coincide and the separatrices really split under the perturbations of degree four, six and eight. In addition, although the first coefficients of the asymptotic series $\sum_{j\geq 0} \alpha_j^\epsilon h^{2j}$ decrease, an accurate computation of several hundreds of its coefficients shows that it is Gevrey [1], and so it diverges for all $h \neq 0$. In fact, it is Gevrey-1 of type $\rho = 1/2\pi^2$, since the following asymptotic expansion holds

$$\bar{\alpha}_j^\epsilon := \frac{(2\pi^2)^{2j} \alpha_j^\epsilon}{\epsilon(2j+2)!} \stackrel{\text{as}}{=} \bar{\alpha}_\infty^\epsilon + \sum_{l\geq 2} \bar{\beta}_l^\epsilon j^{-l} \qquad (j \to +\infty).$$

We have computed the limit $\bar{\alpha}_\infty^\epsilon$ and the first coefficients $\bar{\beta}_l^\epsilon$ with more than forty decimal digits for $n = 2, 3, 4$ and for $\epsilon = 10^{-k}$ with $k = 1, 2, \ldots, 9, 10, 50$. The values obtained for $\epsilon = 10^{-50}$ are a very accurate approximation of the limits $\bar{\alpha}_\infty^0 = \lim_{\epsilon\to 0} \bar{\alpha}_\infty^\epsilon$ and $\bar{\beta}_l^0 = \lim_{\epsilon\to 0} \bar{\beta}_l^\epsilon$. Surprisingly, if the perturbation is quartic: $n = 2$, then $\bar{\alpha}_\infty^0 = -8$ and the first limit values $\bar{\beta}_l^0$ are rational combinations of powers of $\pi^4$, namely

$$\bar{\beta}_2^0 = 14\pi^4/3, \quad \bar{\beta}_3^0 = -2\pi^4, \quad \bar{\beta}_4^0 = 2\pi^4/3 - 8\pi^8/5, \quad \bar{\beta}_5^0 = -9\pi^8/5.$$

There are not so nice results for the perturbations of degree six and eight.

The experiments presented up to now are very similar to the ones contained in [6] on perturbed weakly hyperbolic McMillan maps. We note that the study of billiards is computationally more expensive (by a factor ten) than the study of McMillan maps. In spite of this drawback, the experiments in the current paper are more accurate due to three factors: 1) the hardware is faster (we have used a Beowulf cluster with several tens of processors, instead of a single computer); 2) the software is better (we have written our programs using the PARI system [1], instead of hand-made routines); and 3) the algorithms have been tuned in several tricky ways. The realization of these improvements has been an incentive to push the experiments beyond the first exponential term. C. Simó [21] has performed similar experiments for the *standard map*, the *Hénon map*, and other celebrated maps.

To explain this new challenge, let $\omega^\pm$ be the Lazutkin's homoclinic invariants (introduced in [9], see also §2.1) of the above-mentioned symmetric trajectories $\mathcal{T}_\epsilon^\pm$. As usual, $\omega^+$ and $\omega^-$ have the same asymptotic behavior, but with

[1] A series $\sum_j f_j x^j$ is *Gevrey-r of type $\rho$* when there exist constants $C, \ell > 0$ such that $|f_j| \leq C\rho^{rj}\Gamma(rj+\ell)$, where $\Gamma(z)$ stands for the Gamma function. When $r = 1$, the Borel transform $\sum_j f_j s^{j-1}/(j-1)!$ converges in the disk $\{s \in \mathbb{C} : |s| < 1/\rho\}$.

opposite signs. In fact,

$$\omega^{\pm} \stackrel{\text{as}}{=} \pm 2\pi^2 a\epsilon h^{-2}\, \mathrm{e}^{-\pi^2/h} \sum_{j\geq 0} \alpha_j^\epsilon h^{2j} \qquad (h \to 0^+, \epsilon \text{ fixed})$$

where $\sum_{j\geq 0} \alpha_j^\epsilon h^{2j}$ is the same series that appeared in the expansion of $\Delta_{\text{Length}}$. Thus, it is quite natural to study the asymptotic behavior of $\Omega := \omega^+ + \omega^-$. There are reasons to guess that $\Omega$ has order $\mathrm{e}^{-2\pi^2/h}$, see §2.5. We have checked that this guest is correct, since

$$\Omega := \omega^+ + \omega^- \stackrel{\text{as}}{=} 16\pi^2 a\epsilon h^{-2}\, \mathrm{e}^{-2\pi^2/h} \sum_{j\geq 0} \omega_j^\epsilon h^{2j} \qquad (h \to 0^+, \epsilon \text{ fixed})$$

for some new series $\sum_{j\geq 0} \omega_j^\epsilon h^{2j}$, which is also Gevrey-1 of type $\rho = 1/2\pi^2$, although the analysis of this asymptotic series has been more delicate.

As $\epsilon \to 0$, the Gevrey coefficients $\alpha_j^\epsilon$ and $\omega_j^\epsilon$ tend to the Taylor coefficients $\alpha_j^0$ and $\omega_j^0$ of a couple of analytic functions (in the variable $h$) which can be explicitly computed with a discrete Melnikov method. For instance,

$$\alpha_0^0 = (-1)^n 4\pi^2 \sum_{j=0}^{n-2} \frac{(-1)^j \pi^{2j}}{(2j+1)!}, \qquad \omega_0^0 = (-1)^n 8\pi^2 \sum_{j=0}^{n-2} \frac{(-1)^j (2\pi)^{2j}}{(2j+1)!}.$$

Obviously, the errors $|\alpha_j^\epsilon - \alpha_j^0| = \mathrm{O}(\epsilon)$ and $|\omega_j^\epsilon - \omega_j^0| = \mathrm{O}(\epsilon)$ are not uniform in the index $j$, since the Taylor coefficients verify some *potential bounds*, whereas the Gevrey coefficients *grow* at a factorial rate, see remark 3.

A similar understanding of the next exponential terms seems unreachable with the present techniques, and not for computing limitations. For instance, in the same way that the sum of the homoclinic invariants $\omega^+$ and $\omega^-$ cancels the first exponential term jointly with its whole asymptotic series, so that it leads naturally to the study of the second exponential term, the combination

$$h^2(\omega^+ - \omega^-)/4\pi^2 - \Delta_{\text{Length}}$$

cancels simultaneously the first and second exponential terms (jointly with their asymptotic series). But, although this quantity seems to be of order $\epsilon h^{-1}\, \mathrm{e}^{-3\pi^2/h}$ as $h \to 0^+$, we have been not able to filter any information about the asymptotic series that this third exponential term could have. There are intrinsic mathematical limitations that obstruct this study.

The last question that we have tackled out is the analysis of the homoclinic bifurcations that take place inside the perturbation of degree six

$$C_d = \left\{ (x,y) \in \mathbb{R}^2 : x^2/a^2 + y^2/b^2 + \epsilon(y^2/\gamma^2 + d) y^4/\gamma^4 = 1 \right\}.$$

Once fixed $\epsilon$ and $h$, we look for changes in the topology of the perturbed invariant curves when the *bifurcation parameter* $d$ varies. It turns out that there exist three bifurcation values $d_+ < d_0 < d_-$ such that

$$d = d_\pm \Rightarrow \omega^\pm = 0, \qquad d = d_0 \Rightarrow \Delta_{\mathrm{Length}} = 0.$$

At the bifurcation value $d = d_+$ (respectively, $d = d_-$) the separatrices have a cubic tangency at the four $y$-axial (respectively, $x$-axial) persistent homoclinic orbits and the number of primary homoclinic orbits changes. In the weakly hyperbolic case, these bifurcations are *almost invisible*, since they take place in an exponentially small (with respect to $h$) range of $d$. Concretely,

$$\Delta_{\mathrm{Bifur}} := d_- - d_+ \overset{\mathrm{as}}{=} \mathrm{e}^{-\pi^2/h} \sum_{j\geq 0} \delta_j^\epsilon h^{2j} \qquad (h \to 0^+, \epsilon \text{ fixed})$$

for some asymptotic series $\sum_{j\geq 0} \delta_j^\epsilon h^{2j}$ such that $\delta_0^\epsilon = 8\pi^2 + \mathrm{O}(\epsilon)$. (It is worth mentioning that there are other homoclinic bifurcations for big values of $h$, but they fall out of the scope of this work since they are perfectly visible.)

The experiments get complicated by problems of stability, precision, and time. This forces us to use a multiple-precision arithmetic, to expand the invariant curves up to high order, and to take advantage of symmetries. The methods goes back to C. Simó [19], and were first applied in [8]. They are also explained in [6]. The main obstacle is the computation of exponentially small quantities with much more precision than the usual one; namely, with a relative error less than $10^{-1500}$ in the most extreme cases. Hence, the use of a multiple-precision arithmetic is unavoidable, due to the requirement of a very high precision in the final result, and the cancellation of order $\mathrm{e}^{-2\pi^2/h}$ produced when adding the Lazutkin's homoclinic invariants. For sample, when $h = 0.002$, the computation requires an arithmetic with 6500 digits and the expansion of the invariant curves up to order 1100. It takes two to three hours, depending on the degree $n$, on a single Xeon 2800 processor running Linux. The computations have been launched in a Beowulf cluster of Xeon processors. The programs have been written using the PARI system [1].

We complete this introduction with a note on the organization of this paper. In section 2, we introduce some concepts about the splitting of separatrices for area-preserving maps and we also recall some previous celebrated results about exponentially small phenomena. In section 3, we present the convex billiard tables studied through the rest of the paper. Afterwards, we describe in section 4 the results on the exponentially small splittings that take place under monomial perturbations. The study of the primary homoclinic bifurcations under a binomial perturbation is contained in section 5. The details on the Melnikov and numerical computations are relegated to appendixes §A and §B, respectively.

## 2 Notations and a bit of history

We introduce some quantities used to measure the size of the splitting and we reproduce several celebrated results on their exponentially smallness. For the sake of space the exposition is very compact, sometimes even without precise statements of the theorems. We are mainly interested in the ideas. More details can be found in the survey [12].

### *2.1 Homoclinic invariants*

Let $f : M \to M$ be an analytic exact area-preserving diffeomorphism defined on an exact symplectic surface $M$. Let $\mho = -\,\mathrm{d}\theta$ be the exact symplectic two-form such that $f^*\mho = \mho$. Then there exists a function $L : M \to \mathbb{R}$ such that $f^*\theta - \theta = \mathrm{d}L$. This function is the *generating function* of the map $f$.

Let $m_\infty$ be a saddle point of $f$. Then the eigenvalues of $df(m_\infty)$ are of the form $\lambda$ and $\lambda^{-1}$ for some real $\lambda$ such that $|\lambda| > 1$. Squaring the map if necessary, we can assume that $\lambda > 1$. The quantity $h > 0$ such that $\lambda = \mathrm{e}^h$ is the *characteristic exponent* of the saddle point.

It is well-known that a saddle point $m_\infty$ has *stable* and *unstable curves*

$$W^{\mathrm{s}}(m_\infty) = \{m \in M : \lim_{n\to+\infty} \operatorname{dist}(f^n(m), m_\infty) = 0\}\,,$$
$$W^{\mathrm{u}}(m_\infty) = \{m \in M : \lim_{n\to-\infty} \operatorname{dist}(f^n(m), m_\infty) = 0\}\,.$$

They are analytic immersions of the real line in $M$ without self-intersections and there exist some analytic parameterizations $m^{\mathrm{u,s}} : \mathbb{R} \to W^{\mathrm{u,s}}$ such that

$$m^{\mathrm{u,s}}(0) = m_\infty, \qquad f(m^{\mathrm{u}}(r)) = m^{\mathrm{u}}(\lambda r), \qquad f(m^{\mathrm{s}}(r)) = m^{\mathrm{s}}(r/\lambda).$$

It is convenient to pass to the parameter $t = \ln r$ (respectively, $t = -\ln r$) on the unstable (respectively, stable) curve by setting $\psi^{\mathrm{u}}(t) = m^{\mathrm{u}}(\mathrm{e}^t)$ and $\psi^{\mathrm{s}}(t) = m^{\mathrm{s}}(\mathrm{e}^{-t})$. Clearly, the functions $\psi^{\mathrm{u,s}}(t)$ satisfy the conditions

$$\lim_{t\to-\infty} \psi^{\mathrm{u}}(t) = m_\infty, \qquad \lim_{t\to+\infty} \psi^{\mathrm{s}}(t) = m_\infty, \qquad f(\psi^{\mathrm{u,s}}(t)) = \psi^{\mathrm{u,s}}(t+h).$$

Finally, let $O = (m_n)_{n\in\mathbb{Z}}$ be a homoclinic orbit to $m_\infty$ passing by a point $m_0$ (that is, $m_n = f^n(m_0)$ and $\lim_{n\to\pm\infty} m_n = m_\infty$) such that $m^{\mathrm{u,s}}(r^{\mathrm{u,s}}) = \psi^{\mathrm{u,s}}(t^{\mathrm{u,s}}) = m_0$ for some parameters $r^{\mathrm{u,s}} > 0$ and $t^{\mathrm{u,s}} = \ln r^{\mathrm{u,s}}$.

Then the *Lazutkin's homoclinic invariant* of $m_0$ is defined as the quantity

$$\omega(m_0) := \mho(\dot{\psi}^{\mathrm{u}}(t^{\mathrm{u}}), \dot{\psi}^{\mathrm{s}}(t^{\mathrm{s}})) = r^{\mathrm{u}} r^{\mathrm{s}} \mho(\dot{m}^{\mathrm{u}}(r^{\mathrm{u}}), \dot{m}^{\mathrm{s}}(r^{\mathrm{s}})).$$

It does not depend on the point of the homoclinic orbit: $\omega(m_n) = \omega(m_0)$ for all $n$, so that we can write $\omega = \omega(O)$. Moreover, $\omega$ is invariant by symplectic changes of variables and is proportional to the splitting angle. In particular, $\omega(O) = 0$ if and only if the invariant curves are tangent along $O$. Therefore, it is a very suitable quantity to measure the splitting size on a homoclinic orbit.

On the other hand, the *homoclinic action* of the orbit $O$ is the quantity

$$W[O] := \sum_{n\in\mathbb{Z}} (L(m_n) - L(m_\infty)).$$

By hyperbolicity this series is absolutely convergent. Now, let us suppose that $O^+ = (m_n^+)_{n\in\mathbb{Z}}$ and $O^- = (m_n^-)_{n\in\mathbb{Z}}$ are a couple of homoclinic orbits such that the pieces of the curves $W^{\mathrm{u,s}}$ between $m_0^+$ and $m_0^-$ do not contain other points of these orbits. These pieces enclose a region called a *lobe*. The area $A$ of this lobe is also invariant by symplectic changes of variables. It is another classical measure of the splitting size. Using the *MacKay-Meiss-Percival action principle* [17], we get that

$$A = W[O^-] - W[O^+] = \sum_{n\in\mathbb{Z}} (L(m_n^-) - L(m_n^+)).$$

Finally, we recall that a homoclinic orbit $O = (m_n)_{n\in\mathbb{Z}}$ is called *primary* when the pieces of the curves $W^{\mathrm{u,s}}$ between the saddle point $m_\infty$ and the homoclinic point $m_0$ have only their ends in common. These orbits are very important, because if there is one homoclinic orbit, there are an infinity of them, but only a finite portion of them are primary. And this portion suffices to understand the structure of the homoclinic tangle. The homoclinic orbits mentioned in this paper are always primary.

### *2.2 The rôle of the reversors*

In general, all celebrated area-preserving maps have a very useful property which simplifies very much the search of homoclinic points. Let us explain it.

We suppose that the diffeomorphism $f$ has a reversor $R$, that is, an involutive map such that $f \circ R = R \circ f^{-1}$. We also assume that the *symmetry line*

$$\mathrm{Fix}\{R\} := \{m \in M : R(m) = m\}$$

is a smooth curve in the surface $M$. Finally, we suppose that $m_\infty \in \mathrm{Fix}\{R\}$. Then the reversor interchanges the invariant curves: $R(W^{\mathrm{u,s}}) = W^{\mathrm{s,u}}$.

Under these assumptions, the points on the intersection of the invariant curves with the symmetry line are a special kind of homoclinic points, usually called *symmetric*. From a numerical point of view, the symmetric homoclinic points are easier to compute, since the symmetry lines often have closed expressions. Furthermore, if a reversible diffeomorphism has an invariant curve transverse to a symmetry line, then the associated symmetric homoclinic points persist under small reversible perturbations of the diffeomorphism.

On the other hand, if $R$ is a reversor of $f$, then $f \circ R$ is another reversor. Therefore, the reversors (and so their symmetric homoclinic orbits) appear in natural couples. That is, if a reversible map has a symmetric homoclinic orbit $O^+$ passing by a point $m_0^+ \in \mathrm{Fix}\{R\}$, usually it has another symmetric homoclinic orbit $O^-$ passing by a point $m_0^- \in \mathrm{Fix}\{f \circ R\}$. Then the Lazutkin's homoclinic invariants $\omega^\pm = \omega(O^\pm)$ and the lobe area $A$ enclosed by these orbits are the most natural quantities to measure the splitting size. We present some examples in the next subsections.

### *2.3 The standard map and the Hénon map*

Probably, the most celebrated example is the *standard map* defined by

$$SM : \mathbb{T}^2 \to \mathbb{T}^2, \qquad SM(x, y) = (x + y + \epsilon \sin x, y + \epsilon \sin x).$$

It is area-preserving and reversible. The map $R(x, y) = (2\pi - x, y + \epsilon \sin x)$ is one of the reversors, being $\{x = \pi\}$ its symmetry line. If $\epsilon > 0$, the origin is hyperbolic and its characteristic exponent $h$ is determined by $\epsilon = 4 \sinh^2(h/2)$. Let $\omega^+$ and $\omega^-$ be the Lazutkin's invariants of the symmetric homoclinic orbits passing by the first intersection of the invariant curves with the symmetry lines of the reversors $R$ and $SM \circ R$, respectively. Let $A$ be the area of the lobe enclosed between the couple of symmetric homoclinic orbits. Then the following asymptotic formulae hold

$$\omega^\pm \stackrel{\mathrm{as}}{=} \pm 4\pi h^{-2}\, \mathrm{e}^{-\pi^2/h} \sum_{j\geq 0} \omega_j h^{2j}, \qquad A \stackrel{\mathrm{as}}{=} 2\pi^{-1}\, \mathrm{e}^{-\pi^2/h} \sum_{j\geq 0} \omega_j h^{2j}$$

for some real coefficients $\omega_j$ that can be determined through some auxiliary problems independent of $\epsilon$. These asymptotic formulae were first stated in [9]. A complete proof can be found in [10]. The fact that the quantities $2\pi^2 h^{-2} A$, $\omega^+$, and $-\omega^-$ have the same asymptotic expansion can be understood with an argument based on the splitting function, see §2.5.

The first asymptotic coefficient $\omega_0 \simeq 1118.827706$ is the *Lazutkin's constant*. Several hundreds of asymptotic coefficients $\omega_k$ were computed by C. Simó. His experiments suggest that the series $\sum_{j\geq 0}\omega_j h^{2j}$ is Gevrey-1 of type $\rho = 1/2\pi^2$.

Another celebrated area-preserving map is the *Hénon map*

$$HM : \mathbb{R}^2 \to \mathbb{R}^2, \qquad HM(x,y) = (x + y + \epsilon x(1-x), y + \epsilon x(1-x)).$$

As before, the origin is a hyperbolic fixed point if $\epsilon > 0$. The relation between the characteristic exponent $h$ and $\epsilon$ is the same than in the previous sample. The Hénon map is reversible, $R(x,y) = (x-y, -y)$ is a distinguished reversor, and $\{y = 0\}$ is the symmetry line. The Lazutkin's invariants of the associated couple of symmetric homoclinic orbits verify the asymptotic formulae

$$\omega^{\pm} \overset{\mathrm{as}}{=} \pm 4\pi h^{-6}\,\mathrm{e}^{-2\pi^2/h}\sum_{j\geq 0}\omega_j h^{2j}.$$

This asymptotic series has nothing to do with the one of the standard map. I do not know references with a complete proof of this asymptotic expansion. The first asymptotic coefficient $\omega_0 \simeq 2474425.5935525$ can be found in [3]. The fact that $\omega_0 \neq 0$ was analytically established in [11]. Several hundreds of digits of $\omega_0$ are listed in [20]. Some numerical experiments performed by C. Simó suggest that this new asymptotic series is also Gevrey-1.

### *2.4 McMillan maps and Melnikov methods*

Now we present a qualitatively different kind of example. Concretely, we study perturbations of some integrable standard-like maps first introduced by McMillan [18]. Although these perturbed maps are by far less celebrated than the standard or the Hénon maps, they have the following interesting property. They depend on two parameters: the perturbation strength $\epsilon$ and the characteristic exponent $h$ of the origin. For $\epsilon = 0$, they are integrable with a separatrix to the origin, whereas they asymptote to flows with homoclinic connections as $h \to 0^+$. Moreover, some explicit exponentially small estimates of the splitting size can be obtained using a discrete version of the Melnikov method. It represents a strong coincidence with the billiard maps here studied.

To begin with, let us consider the area-preserving map

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \qquad f(x,y) = \left(y, -x + \frac{2\mu_0 y}{1+y^2} + \epsilon V'(y)\right)$$

where $V(y) = \sum_{n\geq 1} V_n y^{2n}$ is *any* even entire perturbation. If $\mu := \mu_0 + \epsilon V_1 > 1$, the origin is a hyperbolic point whose characteristic exponent $h$ is determined by the relation $\mu = \cosh h$. This map is reversible and $R(x, y) = (y, x)$ is a reversor whose symmetry line is the bisectrix $\{y = x\}$. Let $A$ be the area of the region enclosed by the invariant curves between the couple of symmetric homoclinic orbits contained in the first quadrant. The $\mathrm{O}(\epsilon)$-term of $A$ can be computed using standard ideas in Melnikov theory:

$$A = A(h, \epsilon) = \epsilon A_1(h) + \mathrm{O}(\epsilon^2), \qquad A_1(h) = \mathrm{e}^{-\pi^2/h}\left(8\pi \widehat{V}(2\pi) + \mathrm{O}(h^2)\right)$$

where $\widehat{V}(\xi) = \sum_{n\geq 1} V_n \xi^{2n-1}/(2n-1)!$ is the Borel transform of $V(y)$, see [5]. Thus, the Melnikov term $\epsilon A_1(h)$ gives the right asymptotic behavior of $A$ when $h$ is fixed, $A_1(h) \neq 0$, and $\epsilon \to 0$. On the contrary, when $h \to 0^+$ the Melnikov prediction is, at a first glance, useless, unless $\epsilon$ is exponentially small in $h$.

The papers [5] and [6] are devoted to clarify the asymptotic behavior of $A$ as $h \to 0^+$. In the first paper, it is analytically established that if $p > 6$ then

$$A = \epsilon\, \mathrm{e}^{-\pi^2/h}\left(8\pi \widehat{V}(2\pi) + \mathrm{O}(h^2)\right) \qquad (h \to 0^+, \epsilon = h^p).$$

On the other hand, the main conclusion of the numeric experiments presented in the second paper is that if $V'(y) = y$ or $V'(y) = y^3$ then

$$A \stackrel{\mathrm{as}}{=} \epsilon\, \mathrm{e}^{-\pi^2/h} \sum_{j\geq 0} \alpha_j^\epsilon h^{2j} \qquad (h \to 0^+, \epsilon \text{ fixed})$$

where $\sum_{j\geq 0} \alpha_j^\epsilon h^{2j}$ is Gevrey-1 of type $\rho = 1/2\pi^2$ and $\alpha_0^\epsilon = 8\pi\widehat{V}(2\pi) + \mathrm{O}(\epsilon)$.

### *2.5 The splitting function and the splitting potential*

We describe briefly some ideas originally proposed by Lazutkin, which were the first step to prove the exponentially smallness of some splitting quantities. Besides, they help to understand the close relation among several splitting quantities. Although these ideas are semi-heuristic in the most general frame, their validity has been established rigorously in some concrete cases.

After Lazutkin, the standard way to prove the exponential smallness in $h$ of the above-mentioned quantities is to construct a real analytic $h$-periodic function $\Psi(t)$, called the *splitting function*, whose main properties are:

(1) Its roots are in 1-to-1 correspondence with the primary homoclinic points.
(2) The area of the lobe enclosed by the separatrices between two primary homoclinic points is equal to its integral between the corresponding roots.

(3) The Lazutkin's invariant of a primary homoclinic orbit is equal to its first derivative at the corresponding root.
(4) It has zero mean: $\int_0^h \Psi(t)\,\mathrm{d}t = 0$.

Next, this function is analytically extended to a complex strip of the form $\Pi_\delta = \{t \in \mathbb{C} : |\Im t| < \delta\}$ for some positive width $\delta$ and it is bounded in a bit narrower strips, usually of width $\delta - h$. Hence, its Fourier coefficients are exponentially small, and so the splitting quantities also are. In many cases, these quantities have order $h^\beta\,\mathrm{e}^{-2\pi\delta/h}$, for some $\beta$. We note that $\delta = \pi$ for the Hénon map, whereas $\delta = \pi/2$ for the standard map, the McMillan perturbed maps, and our billiard maps.

Using that $\Psi(t)$ has zero mean, we find another $h$-periodic function $\Theta(t)$, called the *splitting potential*, such that $\Psi(t) = \Theta'(t)$. We can chose $\Theta(t)$ in such a way that it has zero mean. Besides, when the map is reversible, the splitting potential is even and, normalizing it, one can impose that the symmetric primary homoclinic orbits are located at $t = 0$ and $t = h/2$. Hence,

$$A = \Theta(h/2) - \Theta(0), \qquad \omega^+ = \Theta''(0), \qquad \omega^- = \Theta''(h/2).$$

We write the Fourier expansion of the splitting potential as

$$\Theta(t) = -\tfrac{1}{2}\textstyle\sum_{n\geq 1}\Theta_n \cos(Tnt),$$

where $T = 2\pi/h$. In what follows, we suppose that $\Theta_n$ has order $\mathrm{e}^{-2\pi n\delta/h}$, which is rather natural, since the splitting potential is $h$-periodic and can be extended to the strip $\Pi_\delta$. Then we can approximate the first Fourier coefficients of $\Theta(t)$ in terms of the splitting quantities $\omega^\pm$ and $A$. Firstly, we note that

$$A = \sum_{n\geq 0}\Theta_{2n+1}, \qquad \Sigma = 4\sum_{n\geq 0} n^2\Theta_{2n}, \qquad \Delta = \sum_{n\geq 0}(2n+1)^2\Theta_{2n+1},$$

where $\Sigma := (\omega^+ + \omega^-)/T^2$ and $\Delta := (\omega^+ - \omega^-)/T^2$. Therefore, we get that

$$\Theta_1 \approx (9A - \Delta)/8, \qquad \Theta_2 \approx \Sigma/4, \qquad \Theta_3 \approx (\Delta - A)/8.$$

These digressions clarify why the quantities $2\pi^2 h^{-2} A$, $\omega^+$, and $-\omega^-$ have the same asymptotic behavior as $h \to 0^+$. They also suggest that the sum $\Omega := \omega^+ + \omega^- \approx 16\pi^2 h^{-2}\Theta_2$ and the combination $h^2(\omega^+ - \omega^-)/4\pi^2 - A$ have sizes of order $\mathrm{e}^{-4\pi\delta/h}$ and $\mathrm{e}^{-6\pi\delta/h}$, respectively. Thus, we have found candidates to study the second and third exponential terms that govern the splitting.

## 3 Billiard tables

We collect some classical results on convex and elliptic billiards. Most of them can be found in the monographes [14,22], although we have taken the notation from [13, §9.2]. We also present the perturbations we shall study later on.

### 3.1 Convex billiard tables

Let $C$ be a closed convex curve in the Euclidean plane $\mathbb{R}^2$. Let $\gamma : \mathbb{T} \to C$ be a parameterization of this curve, where $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$ stands for the configuration space. Finally, let us consider the phase space

$$\mathbb{A} := \mathbb{T} \times (0, \pi) = \{(\theta, r) : \theta \in \mathbb{T}, 0 < r < \pi\}. \tag{2}$$

Then we can model the billiard dynamics inside $C$ by means of the map $f(\theta, r) = (\theta', r')$ defined as follows. If the particle hits the curve at a point $c = \gamma(\theta) \in C$ under an angle of incidence $r$, the next point is $c' = \gamma(\theta') \in C$ and the next angle of incidence is $r'$. This map has the following properties:

- *Regularity.* If the curve is smooth or analytic, so is the map.
- *Hyperbolicity.* Let $c_\pm = \gamma(\theta_\pm)$ be the ends of a chord of $C$ and $\ell = |c_+ - c_-|$. Let $\kappa_\pm$ be the curvature of $C$ at $c_\pm$. Then the chord is hyperbolic if and only if $(\kappa_+ + \kappa_-)\ell > 4$. When this condition holds, the characteristic exponent $h > 0$ of the associated two-periodic points is determined by the relation

$$2\cosh(h/2) = \sqrt{(\kappa_+ + \kappa_-)\ell}.$$

- *Reversibility.* If $C$ is symmetric with regard to both axes of coordinates, then the map is reversible, $R(\theta, r) = (\pi - \theta, r)$ is a reversor, and

$$\mathrm{Fix}\{R\} = \{(\theta, r) \in \mathbb{A} : \theta = \pi/2 \pmod{\pi}\}.$$

- *Exactness.* Let $\rho = |\dot\gamma(\theta)| \cos r$ and $\rho' = |\dot\gamma(\theta')| \cos r'$. The map is exact in the $(\theta, \rho)$ coordinates: $\rho' \, \mathrm{d}\theta' - \rho \, \mathrm{d}\theta = \mathrm{d}L(\theta, \theta')$, where $L(\theta, \theta') = |\gamma(\theta) - \gamma(\theta')|$ is the *generating function* or *Lagrangian.*
- *Twist character.* Given $c = \gamma(\theta)$, the function $(0, \pi) \ni r \mapsto \theta'(\theta, r) \in \mathbb{T} \setminus \{\theta\}$ is a diffeomorphism. Hence, given two different impact points $c$ and $c'$, there exists a unique billiard trajectory from $c$ to $c'$.
- *Lagrangian formulation.* The billiard dynamics can be expressed by means of implicit difference equations of second order: given three impact points $c_-, c, c_+ \in C$ such that $c_- = \gamma(\theta_-)$, $c = \gamma(\theta)$ and $c_+ = \gamma(\theta_+)$, there exists a

billiard trajectory from $c_-$ to $c_+$ passing by $c$ if and only if

$$\partial_2 L(\theta_-, \theta) + \partial_1 L(\theta, \theta_+) = 0.$$

- *Variational formulation.* The billiard trajectories $\mathcal{T} = (c_k)_{k\in\mathbb{Z}} \in C^{\mathbb{Z}}$ are in correspondence with the (formal) critical points of the *action functional*

$$W : \mathbb{T}^{\mathbb{Z}} \to \mathbb{R}, \qquad W[(\theta_k)_{k\in\mathbb{Z}}] = \sum_{k\in\mathbb{Z}} L(\theta_{k-1}, \theta_k)$$

by means of the relation $c_k = \gamma(\theta_k)$. We shall write $W[\mathcal{T}] = W[(\theta_k)_{k\in\mathbb{Z}}]$.
- *Actions and lengths.* If $\mathcal{T}^+$ and $\mathcal{T}^-$ are two billiard trajectories homoclinic to a hyperbolic chord of $C$, the area of the region enclosed by the invariant curves between them is equal to

$$A = W[\mathcal{T}^-] - W[\mathcal{T}^+] = \mathrm{Length}\mathcal{T}^- - \mathrm{Length}\mathcal{T}^+ =: \Delta_{\mathrm{Length}}.$$

### *3.2 Elliptic billiard tables*

Let $f : \mathbb{A} \to \mathbb{A}$ be the billiard map inside the ellipse

$$C = \left\{(x, y) \in \mathbb{R}^2 : x^2/a^2 + y^2/b^2 = 1\right\} = \{\gamma(\theta) : \theta \in \mathbb{T}\} \qquad 0 < b < a,$$

where $\gamma(\theta) = (a\cos\theta, b\sin\theta)$ is the parameterization we shall use.

The chord whose ends are the vertexes $c_\pm = (\pm a, 0)$ is hyperbolic, because $\ell = |c_+ - c_-| = 2a$ and $\kappa_\pm = a/b^2$, so that $(\kappa_+ + \kappa_-)\ell = 4a^2/b^2 > 4$. Besides, its characteristic exponent $h$ is determined by the relation $\cosh(h/2) = a/b$, as mentioned in the introduction.

We recall now a nice geometric property of the ellipses. Let

$$C(\lambda) = \left\{(x, y) \in \mathbb{R}^2 : \frac{x^2}{a^2 - \lambda^2} + \frac{y^2}{b^2 - \lambda^2} = 1\right\} \qquad \lambda^2 \neq b^2, a^2$$

be the family of *confocal conics* to the ellipse. It is clear that $C(\lambda)$ is an ellipse for $\lambda^2 < b^2$ and a hyperbola for $b^2 < \lambda^2 < a^2$. No real conic exists for $\lambda^2 > a^2$. Concerning the degenerate cases $\lambda^2 \in \{b^2, a^2\}$, we first note that for $\lambda \to b^-$ (respectively, $\lambda \to b^+$) the conic $C(\lambda)$ flattens into the region of the x-axis enclosed by (respectively, outside) the foci of the ellipse $C = C(0)$. On the other hand, the hyperbola flattens into the whole y-axis when $\lambda \to a^-$.

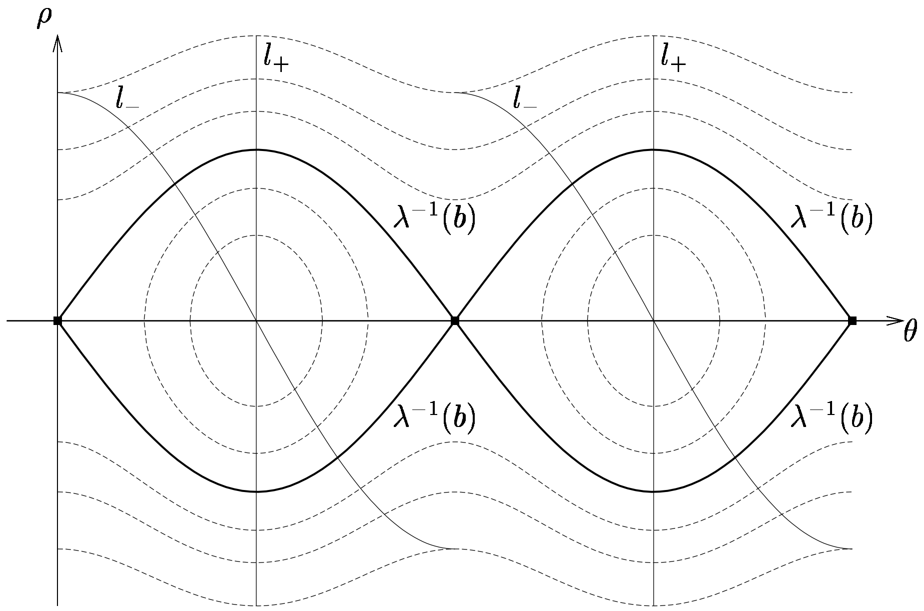


Fig. 1. Phase portrait of the elliptic billiard map in $(\theta, \rho)$ coordinates. The solid squares are the hyperbolic two-periodic points. The thick lines are the separatrices. The thin lines are the symmetry lines. The dashed lines are some level curves of the first integral $\lambda(\theta, \rho)$.

The fundamental geometric property of elliptic billiards is that *any segment (or its prolongation) of a billiard trajectory inside the ellipse $C$ is tangent*[2] *to one fixed confocal conic $C(\lambda)$*. Thus, the confocal conics are *caustics*. The notion of tangency in the degenerate cases is the following. A line is *tangent to* $C(b)$ when it passes through one of the foci $(\pm c, 0)$, where $c = \sqrt{a^2 - b^2}$. A line is *tangent to* $C(a)$ when it coincides with the y-axis. As a by-product, we obtain that the elliptic billiard map is integrable, since the function $\lambda^2 : \mathbb{A} \to \mathbb{R}$ is a first integral. Using the symplectic coordinates $(\theta, \rho)$ defined in the previous subsection, this first integral becomes $\lambda^2(\theta, \rho) = b^2 + c^2 \sin^2 \theta - \rho^2$.

The phase portrait of the elliptic billiard map $f$ is displayed in figure 1. It shows that the $\infty$-shaped level set $\lambda^{-1}(b) = \{(\rho, \theta) : \rho = \pm c \sin \theta\}$ contains the two-periodic points and the four separatrices connecting them.

The ellipse is symmetric with regard to both axes of coordinates. Therefore, the elliptic billiard map $f$ is reversible. Let $R$ be the reversor introduced in §3.1, whose associated symmetry lines $l_+ := \mathrm{Fix}\{R\}$ and $l_- := \mathrm{Fix}\{f \circ R\}$ are also shown in figure 1. The $\infty$-shaped level set $\lambda^{-1}(b)$ intersects transversely each symmetry line at four points. Consequently, inside the ellipse there are four $x$-axial and four $y$-axial homoclinic billiard trajectories (see figure 2), which persist under symmetric perturbations.

[2] In a projective sense, that is, the points of tangency can be proper or improper.

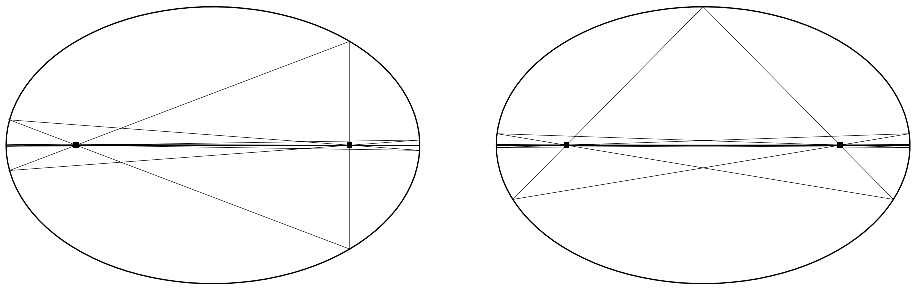

Fig. 2. The two kinds of axial homoclinic trajectories inside an ellipse: x-axial (left) and y-axial (right). The foci are marked with squares.

### *3.3 Perturbed elliptic billiard tables*

We restrict our study to the polynomial perturbations that can be written as

$$C_\epsilon = \left\{(x,y) \in \mathbb{R}^2 : x^2/a^2 + y^2/b^2 + \epsilon P(y^2/\gamma^2) = 1\right\} \tag{3}$$

for some $P(s) \in \mathbb{R}[s]$ such that $P(0) = P'(0) = 0$ and $n := \deg[P(s)] \geq 2$. Here $\epsilon$ is the *perturbative parameter*, $2n$ is the *degree of the perturbation*, and $\gamma = b/e$, where $e = (1 - b^2/a^2)^{1/2}$ is the eccentricity of the ellipse.

This perturbation preserves the axial symmetries of the unperturbed ellipse, the two-periodic hyperbolic trajectory from $c_- = (-a, 0)$ to $c_+ = (a, 0)$, and the characteristic exponent $h$ of this trajectory. The last claim follows from the condition $P(0) = P'(0) = 0$. It suffices to realize that neither the length $\ell = |c_+ - c_-|$ nor the curvature $\kappa_\pm$ of $C_\epsilon$ at $c_\pm$ depend on $\epsilon$.

On the other hand, this kind of perturbation always splits the separatrices [4]. That is, the invariant curves no longer coincide, although they still have some common points: the *homoclinic* points, which creates a complicate *homoclinic tangle.* In figure 3 we show a small part of this homoclinic tangle for a quartic perturbation. In the figure there are just eight primary homoclinic orbits: the axial ones, and all the lobes have the same geometric area: $|A|$.

To end, we recall that symmetric perturbations preserves the axial homoclinic trajectories, which become the key objects for the study developed in this paper. Henceforth, let $\mathcal{T}_\epsilon^+$ (respectively, $\mathcal{T}_\epsilon^-$) be any of the four $y$-axial (respectively, $x$-axial) persistent homoclinic trajectories inside the perturbed ellipse. Let $\omega^\pm$ be the Lazutkin's homoclinic invariant of $\mathcal{T}_\epsilon^\pm$. Finally, let $A = \Delta_{\text{Length}} = \text{Length}\mathcal{T}_\epsilon^- - \text{Length}\mathcal{T}_\epsilon^+$ and $\Omega = \omega^+ + \omega^-$.

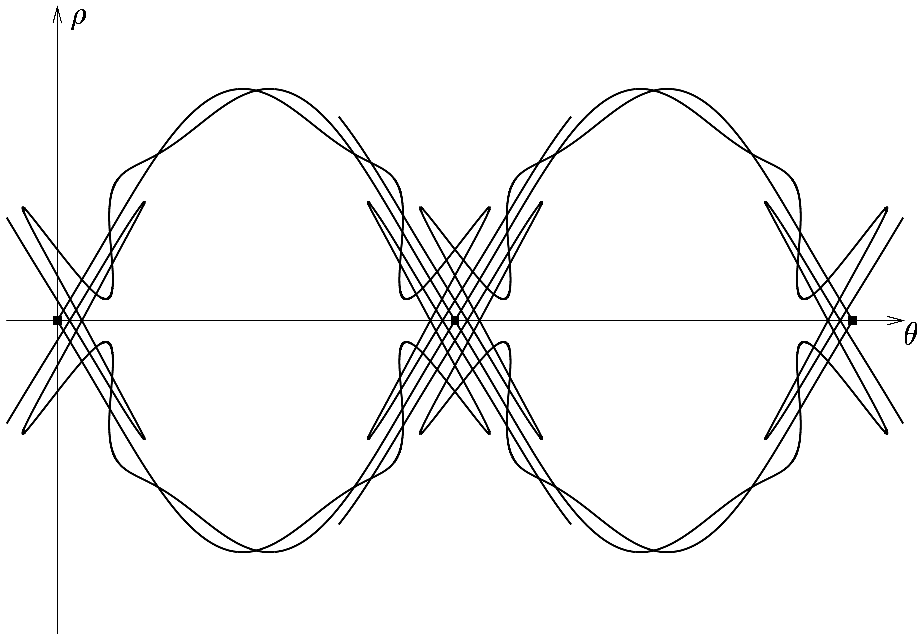


Fig. 3. The homoclinic tangle for $a = 1$, $b = \frac{4}{5}$, $\epsilon = 2$, and $P(s) = s^2$.

## 4 Exponentially small separatrix splittings in some billiard tables

Before to present our numerical results on the exponentially small asymptotic behavior of the splitting quantities $A$ and $\Omega$, we shall derive some analytical predictions about it obtained by using a discrete Melnikov method.

Melnikov methods are very well-known techniques for studying the splitting of separatrices under small perturbations of dynamical systems with homoclinic connections. If the perturbation is of order $\epsilon$, then a standard computation provides the first order term in $\epsilon$ of almost any splitting object. In the frame of area-preserving maps, this has to do with the fact that the splitting potential $\Theta(t)$ introduced in §2.5 has the form $\Theta(t) = \epsilon L(t) + \mathrm{O}(\epsilon^2)$, for some function $L(t)$, called the *Melnikov potential*.

Coming back to our billiard problem, and for the sake of brevity, we shall restrict our digressions to the perturbations (3) of the form $P(s) = s^n$ with $n \geq 2$. They correspond to the monomial perturbations introduced in (1). In [7] it is shown that the Melnikov potential associated to these perturbations is

$$L_n(t) = a e^{2n} \sum_{k \in \mathbb{Z}} \ell_n(t + kh), \tag{4}$$

where $\ell_n(t) = -\nu(t) \operatorname{sech}^{2n-2} t$ and $\nu(t) = \operatorname{sech}(t - h/2) \operatorname{sech}(t + h/2)$. Then if

$$A = A(h, \epsilon) = \epsilon A_1(h) + \mathrm{O}(\epsilon^2), \qquad \Omega = \Omega(h, \epsilon) = \epsilon \Omega_1(h) + \mathrm{O}(\epsilon^2)$$

are the splitting quantities we want to study, their Melnikov terms are

$$A_1(h) = L_n(h/2) - L_n(0), \qquad \Omega_1(h) = L''_n(h/2) + L''_n(0). \tag{5}$$

These Melnikov terms turn out to be exponentially small in $h$. Namely,

$$\left.\begin{aligned} A_1(h) &= a\,\mathrm{e}^{-\pi^2/h}\,\alpha^0(h) + \mathrm{O}(\mathrm{e}^{-3\pi^2/h}), \\ \Omega_1(h) &= 16\pi^2 a h^{-2}\,\mathrm{e}^{-2\pi^2/h}\,\omega^0(h) + \mathrm{O}(h^{-2}\,\mathrm{e}^{-4\pi^2/h}) \end{aligned}\right\} \qquad (h \to 0^+), \tag{6}$$

where $\alpha^0(h)$ and $\omega^0(h)$ are some even analytic functions such that

$$\left.\begin{aligned} \alpha_0^0 &:= \alpha^0(0) = (-1)^n 4\pi^2 \textstyle\sum_{j=0}^{n-2} \frac{(-1)^j \pi^{2j}}{(2j+1)!} \\ \omega_0^0 &:= \omega^0(0) = (-1)^n 8\pi^2 \textstyle\sum_{j=0}^{n-2} \frac{(-1)^j (2\pi)^{2j}}{(2j+1)!} \end{aligned}\right\} \qquad \forall n \geq 2. \tag{7}$$

The functions $\alpha^0(h)$ and $\omega^0(h)$ can be explicitly computed for any $n \geq 2$. As a sample, when the perturbation is quartic: $n = 2$, they are given by

$$\alpha^0(h) = \frac{\omega^0(h)}{2} = \frac{16\pi^2 \sinh^2(h/2)}{h^2 \cosh^4(h/2)} = 4\pi^2 \left(1 - \frac{5h^2}{12} + \frac{77h^4}{720} + \cdots\right). \tag{8}$$

The proof of these claims is rather technical and it has been deferred to §A.1. The key point is the elliptic character of the Melnikov potential.

The *splitting constants* $\alpha_0^0$ and $\omega_0^0$ defined in (7) do not vanish for any $n \geq 2$, because $\pi$ is transcendental. (Nevertheless, $\lim_{n\to+\infty} \alpha_0^0 = \lim_{n\to+\infty} \omega_0^0 = 0$.) In particular, we have the following exponentially small Melnikov predictions

$$A \sim a\epsilon\,\mathrm{e}^{-\pi^2/h}\,\alpha_0^0, \qquad \Omega \sim 16\pi^2 a\epsilon h^{-2}\,\mathrm{e}^{-2\pi^2/h}\,\omega_0^0 \tag{9}$$

for small enough values of $h$ and $\epsilon$. However, the interpretation of the sentence "small enough" is crucial for understanding the different questions to be asked. Of course, the *regular case* —fixed $h > 0$ and $\epsilon \to 0$— falls outside the scope of this paper. We restrict our attention to the *singular* situations:

- The *perturbative singular case*: $\epsilon \to 0$ and $h \to 0^+$.
- The *nonperturbative singular case*: $\epsilon \neq 0$ fixed and $h \to 0^+$.

To begin with, we formulate below a very precise and refined quantitative conjecture on the exponentially small asymptotic behavior of some splitting quantities related to the monomial perturbations (1). (There are no reasons to believe that there are substantial differences between monomial and general perturbations, but the study of monomial perturbations is much simpler.)

Among many other things, if this conjecture was true, then the Melnikov prediction (9) would give the exact measure of the splitting size in the perturbative singular case, whereas in the nonperturbative singular one the only mistake would stem from the splitting constants $\alpha_0^0$ and $\omega_0^0$, which would have to be substituted by new splitting constants $\alpha_0^\epsilon = \alpha_0^0 + \mathrm{O}(\epsilon)$ and $\omega_0^\epsilon = \omega_0^0 + \mathrm{O}(\epsilon)$.

**Conjecture 1** *For any integer $n \geq 2$ and for any small enough $\epsilon \neq 0$, there exist two series $\sum_{j\geq 0} \alpha_j^\epsilon h^{2j}$ and $\sum_{j\geq 0} \omega_j^\epsilon h^{2j}$ such that the lobe area $A = \Delta_{\mathrm{Length}}$, the Lazutkin's homoclinic invariants $\omega^\pm$, and the sum $\Omega = \omega^+ + \omega^-$ associated to the monomial perturbations (1) have the asymptotic expansions*

$$\left.\begin{aligned} A &\stackrel{\mathrm{as}}{=} a\epsilon\, \mathrm{e}^{-\pi^2/h} \textstyle\sum_{j\geq 0} \alpha_j^\epsilon h^{2j} \\ \omega^\pm &\stackrel{\mathrm{as}}{=} \pm 2\pi^2 a\epsilon h^{-2}\, \mathrm{e}^{-\pi^2/h} \textstyle\sum_{j\geq 0} \alpha_j^\epsilon h^{2j} \\ \Omega &\stackrel{\mathrm{as}}{=} 16\pi^2 a\epsilon h^{-2}\, \mathrm{e}^{-2\pi^2/h} \textstyle\sum_{j\geq 0} \omega_j^\epsilon h^{2j} \end{aligned}\right\} \qquad (h \to 0^+, \epsilon \text{ fixed}). \tag{10}$$

*In addition, the following properties hold:*

*(1) If $\alpha^0(h) = \sum_{j\geq 0} \alpha_j^0 h^{2j}$ and $\omega^0(h) = \sum_{j\geq 0} \omega_j^0 h^{2j}$ are the Taylor expansions of the even analytic functions that appear in the Melnikov terms (6), then $\alpha_j^\epsilon = \alpha_j^0 + \mathrm{O}(\epsilon)$ and $\omega_j^\epsilon = \omega_j^0 + \mathrm{O}(\epsilon)$ for all $j \geq 0$.*

*(2) The series $\sum_{j\geq 0} \alpha_j^\epsilon h^{2j}$ and $\sum_{j\geq 0} \omega_j^\epsilon h^{2j}$ are Gevrey-1 of type $\rho = 1/2\pi^2$.*

*(3) The sequence $(\bar{\alpha}_j^\epsilon)_{j\geq 0}$ defined by*

$$\bar{\alpha}_j^\epsilon := \frac{(2\pi^2)^{2j} \alpha_j^\epsilon}{(2j+2)!\epsilon} \tag{11}$$

*has some finite limit $\bar{\alpha}_\infty^\epsilon \neq 0$ when $j \to +\infty$. In fact, there exist some asymptotic coefficients $\bar{\beta}_l^\epsilon$ such that $\bar{\alpha}_j^\epsilon \stackrel{\mathrm{as}}{=} \bar{\alpha}_\infty^\epsilon + \sum_{l\geq 2} \bar{\beta}_l^\epsilon j^{-l}$ as $j \to +\infty$.*

*(4) The sequences $(\bar{\omega}_j^0)_{j\geq 0}$, $(\bar{\omega}_j^1)_{j\geq 0}$ and $(\hat{\omega}_j^\epsilon)_{j\geq 0}$ defined by*

$$\bar{\omega}_j^0 + \epsilon\bar{\omega}_j^1 + \epsilon^2 j^6 \hat{\omega}_j^\epsilon := \frac{(2\pi^2)^{2j}}{(2j+2)!} \frac{\omega_j^\epsilon - \omega_j^0}{\epsilon} \tag{12}$$

*have some finite limits $\bar{\omega}_\infty^0, \bar{\omega}_\infty^1, \hat{\omega}_\infty^\epsilon \neq 0$ when $j \to +\infty$. In fact, there exist some asymptotic coefficients $\bar{\eta}_l^0$, $\bar{\eta}_l^1$ and $\hat{\eta}_l^\epsilon$ such that $\bar{\omega}_j^0 \stackrel{\mathrm{as}}{=} \bar{\omega}_\infty^0 + \sum_{l\geq 2} \bar{\eta}_l^0 j^{-l}$, $\bar{\omega}_j^1 \stackrel{\mathrm{as}}{=} \bar{\omega}_\infty^1 + \sum_{l\geq 2} \bar{\eta}_l^1 j^{-l}$, and $\hat{\omega}_j^\epsilon \stackrel{\mathrm{as}}{=} \hat{\omega}_\infty^\epsilon + \sum_{l\geq 1} \hat{\eta}_l^\epsilon j^{-l}$ as $j \to +\infty$.*

*(5) There exist some values $\bar{\alpha}_\infty^0, \hat{\omega}_\infty^0, \bar{\beta}_l^0, \hat{\eta}_l^0 \in \mathbb{R}$ such that $\bar{\alpha}_\infty^\epsilon = \bar{\alpha}_\infty^0 + \mathrm{O}(\epsilon)$, $\hat{\omega}_\infty^\epsilon = \hat{\omega}_\infty^0 + \mathrm{O}(\epsilon)$, $\bar{\beta}_l^\epsilon = \bar{\beta}_l^0 + \mathrm{O}(\epsilon)$, and $\hat{\eta}_l^\epsilon = \hat{\eta}_l^0 + \mathrm{O}(\epsilon)$.*

*(6) Under the quartic monomial perturbation (that is, when $n = 2$), it turns out that $\bar{\omega}_\infty^0 = 2\bar{\alpha}_\infty^0 = -16$, $\bar{\beta}_2^0 = 14\pi^4/3$, $\bar{\beta}_3^0 = -2\pi^4$, $\bar{\beta}_4^0 = 2\pi^4/3 - 8\pi^8/5$, $\bar{\beta}_5^0 = -9\pi^8/5$, and $\bar{\eta}_l^0 = 2\bar{\beta}_l^0$, for all $l \geq 2$.*

**Remark 2** *In fact, we believe that the quantities introduced in the conjecture that depend on the perturbative parameter $\epsilon$ (the Gevrey coefficients $\alpha_j^\epsilon$ and*

*$\omega_j^\epsilon$, the limit values $\bar{\alpha}_\infty^\epsilon$ and $\hat{\omega}_\infty^\epsilon$, the asymptotic coefficients $\bar{\beta}_l^\epsilon$ and $\hat{\eta}_l^\epsilon$) are analytic at $\epsilon = 0$. If this was true, then it would be very clear why all these quantities are $O(\epsilon)$-close to their finite limits as $\epsilon \to 0$.*

**Remark 3** *We see from equation (8) that the function $\alpha^0(h)$ is analytic in the open disk $|h| < \pi$ if $n = 2$, and this also holds for any $n \geq 3$. Thus, given any $r \in (0, \pi)$, its Taylor coefficients verify the Cauchy inequalities $|\alpha_j^0| \leq M r^{-2j}$ for some $M = M(r) > 0$. So $\lim_{j\to+\infty} (2\pi^2)^{2j} \alpha_j^0/(2j+2)! = 0$. This motivates the presence of $\epsilon$ in the denominator of the normalized sequence (11). On the contrary, the limit $\bar{\alpha}_\infty^\epsilon$ would be $\mathrm{O}(\epsilon)$ and then $\bar{\alpha}_\infty^0 = \lim_{\epsilon\to 0} \bar{\alpha}_\infty^\epsilon = 0$.*

**Remark 4** *Since $\omega^0(h) = 2\alpha^0(h)$ for the quartic monomial perturbation (this relation does not hold for other monomial perturbations), one could think that the relations $\bar{\omega}_\infty^0 = 2\bar{\alpha}_\infty^0$ and $\bar{\eta}_l^0 = 2\bar{\beta}_l^0$ when $n = 2$ are evident, but this is a completely wrong argument, because although the Taylor coefficients verify the relations $\omega_j^0 = 2\alpha_j^0$, they have nothing to do with the limits $\bar{\omega}_\infty^0$ and $\bar{\alpha}_\infty^0$, as has been stressed in the previous remark.*

**Remark 5** *We have discovered that the asymptotic series $\sum_{l\geq 2} \bar{\beta}_l^\epsilon j^{-l}$ begins with a $\mathrm{O}(j^{-2})$-term. In other words, $\bar{\beta}_1^\epsilon = 0$. At the present time we do not know of an ultimate reason for this, but it allows us to control in a very precise way the size of the error in the numerical experiments, since the relative error in our final results must be of the same order that the computed value of $\bar{\beta}_1^\epsilon$. In our computations, $|\bar{\beta}_1^\epsilon| \leq 10^{-40}$. The same trick can be applied to the series $\sum_{l\geq 2} \bar{\eta}_l^0 j^{-l}$ and $\sum_{l\geq 2} \bar{\eta}_l^1 j^{-l}$, but not to $\sum_{l\geq 1} \hat{\eta}_l^\epsilon j^{-l}$.*

**Remark 6** *As regards the fact that $\bar{\beta}_l^0$ and $\bar{\eta}_l^0$ are rational combinations of powers of $\pi^4$ for $2 \leq l \leq 5$ and $n = 2$, it is worth noting that we have not been able to find similar expressions neither for other indexes (that is, for $l \geq 6$) nor for other monomial perturbations (that is, for $n \geq 3$).*

Next, we shall present some numerical evidences supporting this conjecture. The algorithms used to perform the computations are briefly explained in §B.

The use of a very expensive multiple-precision arithmetic encourages us to consider perturbations as simple as possible. Accordingly, we have restricted the computations to the monomial perturbations of degree four, six, and eight. That is, we have considered only the cases $n = 2, 3, 4$.

In a first step, we analyze the differences between the Gevrey coefficients and the Taylor coefficients. We have plotted in figure 4 the graphs of the relative errors of the splitting constants in the range $0 < \epsilon \leq 1$ when $n = 2, 3, 4$. We see that $\alpha_0^\epsilon = \alpha_0^0 + \mathrm{O}(\epsilon)$ and $\omega_0^\epsilon = \omega_0^0 + \mathrm{O}(\epsilon)$. We have also plotted in figure 5 the graphs of the first five Gevrey coefficients $\alpha_j^\epsilon$ and $\omega_j^\epsilon$ in the range $0 < \epsilon \leq 1$ when $n = 2$. Besides, we have marked the points $(0, \alpha_j^0)$ and $(0, \omega_j^0)$ corresponding to the first five Taylor coefficients of the functions (8). It turns

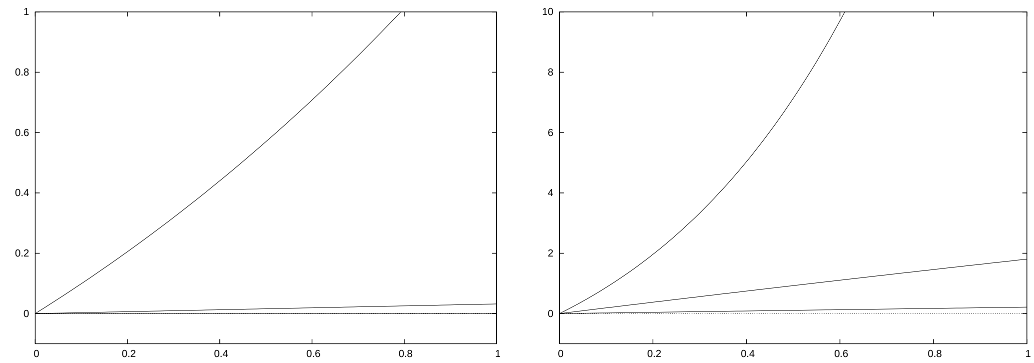


Fig. 4. $|\alpha_0^\epsilon/\alpha_0^0 - 1|$ (left) and $|\omega_0^\epsilon/\omega_0^0 - 1|$ (right) versus $\epsilon$, for $n = 2, 3, 4$.

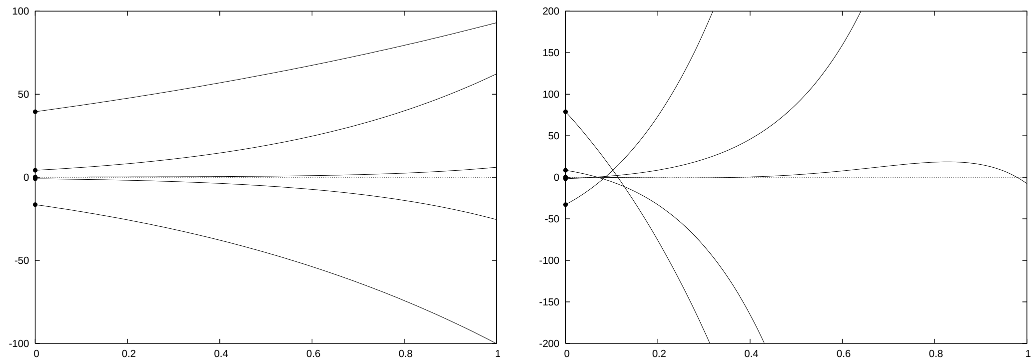


Fig. 5. The Gevrey coefficients $\alpha_j^\epsilon$ (left) and $\omega_j^\epsilon$ (right) versus $\epsilon$, for $0 \leq j \leq 4$ and $n = 2$. The marked points correspond to the Taylor coefficients $\alpha_j^0$ and $\omega_j^0$.

out that $\lim_{\epsilon\to 0} \alpha_j^\epsilon = \alpha_j^0$ and $\lim_{\epsilon\to 0} \omega_j^\epsilon = \omega_j^0$ for $j \leq 4$. The same behavior is observed for bigger values of $j$ and for the monomial perturbations of degree six and eight. We have skipped the corresponding graphs and figures for the sake of clarity and brevity.

In a second step, we study the Gevrey character of the asymptotic series $\sum_j \alpha_j^\epsilon h^{2j}$ and $\sum_j \omega_j^\epsilon h^{2j}$. By definition, a series $\sum_j f_j x^j$ is Gevrey-1 of type $\rho$ if and only if the radius of convergence of its Borel transform $\sum_j f_j s^{j-1}/(j-1)!$ is equal to $1/\rho$. In particular, if the limit $\lim_{j\to+\infty}(|f_j|/j!)^{1/j}$ exists and it is equal to $\rho$, then $\sum_j f_j x^j$ is Gevrey-1 of type $\rho$. Therefore, to prove that the series $\sum_j \alpha_j^\epsilon h^{2j}$ and $\sum_j \omega_j^\epsilon h^{2j}$ are Gevrey-1 of type $\rho = 1/2\pi^2$, it suffices to check that the sequences

$$\tilde{\alpha}_j^\epsilon := (2\pi^2)\,\sqrt[2j]{|\alpha_j^\epsilon|/(2j)!}, \qquad \tilde{\omega}_j^\epsilon := (2\pi^2)\,\sqrt[2j]{|\omega_j^\epsilon|/(2j)!} \tag{13}$$

tend to one as $j \to +\infty$. To carry out this idea, we have computed the first three-hundred Gevrey coefficients $\alpha_j^\epsilon$ and $\omega_j^\epsilon$ for $n = 2, 3, 4$ and several values of $\epsilon$. To be precise, we have performed the computations for $\epsilon = 10^{-k}$ with $k = 1, 2, \ldots, 9, 10, 50$. All the considered values of $\epsilon$ have the same behavior. The results for $\epsilon = \frac{1}{10}$ are shown in figure 6. They strongly suggest that the

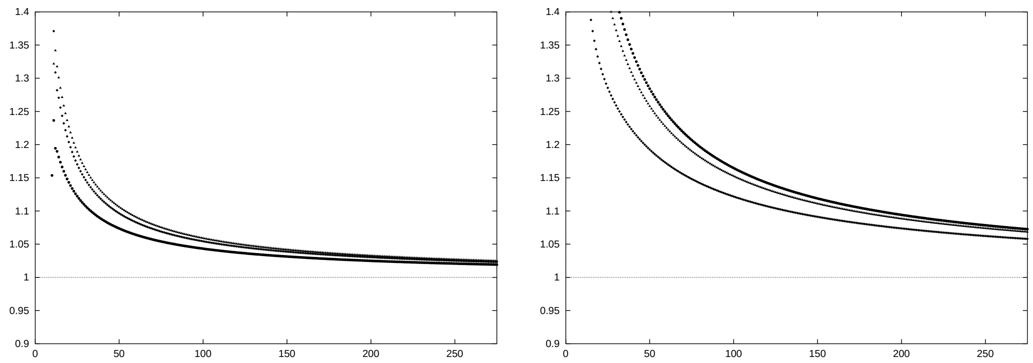


Fig. 6. The coefficients $\tilde{\alpha}_j^\epsilon$ (left) and $\tilde{\omega}_j^\epsilon$ (right) versus $j$, for $\epsilon = \frac{1}{10}$ and $n = 2, 3, 4$.

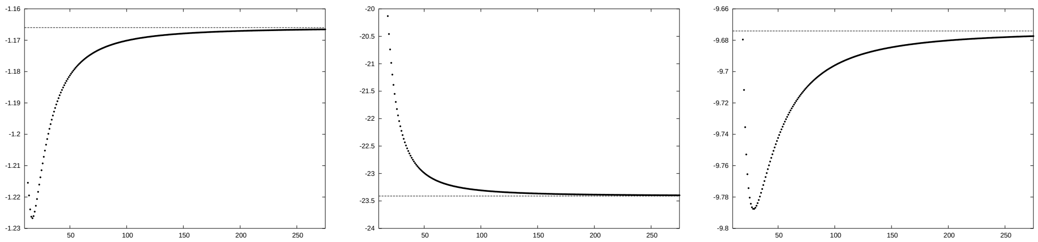


Fig. 7. The coefficients $\bar{\alpha}_j^\epsilon$ versus $j$, for $\epsilon = \frac{1}{10}$. Here, $n = 2$ (left), $n = 3$ (center), and $n = 4$ (right). The dashed lines correspond to the limits $\bar{\alpha}_\infty^\epsilon := \lim_{j\to+\infty} \bar{\alpha}_j^\epsilon$.

sequences (13) tend to one as $j \to +\infty$ in a very regular way.

We describe now with more detail the behavior of the Gevrey coefficients $\alpha_j^\epsilon$ as $j \to +\infty$. We know that the series $\sum_{j\geq 0} \alpha_j^\epsilon h^{2j}$ is Gevrey-1 of type $\rho = 1/2\pi^2$ when there exist constants $C, \ell > 0$ such that $|\alpha_j^\epsilon| \leq C(2\pi^2)^{-2j}\Gamma(2j+\ell)$, where $\Gamma(z)$ stands for the Gamma function. If the third item of our conjecture was true, then one could take $\ell = 2$ and $C = \mathrm{O}(\epsilon)$. In order to show that this part of the conjecture holds, we have plotted the normalized coefficients (11) in figure 7 for $\epsilon = \frac{1}{10}$ and $n = 2, 3, 4$. Apparently, the limits $\bar{\alpha}_\infty^\epsilon = \lim_{j\to+\infty} \bar{\alpha}_j^\epsilon$ exist, are finite, and do not vanish for that value of $\epsilon$. The computations for other values of $\epsilon$ give rise to similar pictures. The limits $\bar{\alpha}_\infty^\epsilon$ have been obtained by using an extrapolation algorithm based on the hypothesis that there exist some asymptotic coefficients $\bar{\beta}_l^\epsilon$ such that

$$\bar{\alpha}_j^\epsilon \stackrel{\text{as}}{=} \bar{\alpha}_\infty^\epsilon + \sum_{l\geq 2} \bar{\beta}_l^\epsilon j^{-l} \qquad (j \to +\infty).$$

The extrapolation gives very accurate results, with relative errors below $10^{-40}$, which is a strong evidence in favor of this hypothesis. As a by-product, we have also obtained the first asymptotic coefficients $\bar{\beta}_l^\epsilon$ for several values of $\epsilon$. We have listed them in table 1 for $\epsilon = \frac{1}{10}$, $l \leq 5$ and $n = 2, 3, 4$. For lack of space only fifteen decimal digits are displayed.

Next, we have studied the limits as $\epsilon \to 0$ of the values $\bar{\alpha}_\infty^\epsilon$ and the asymptotic

Table 1
The limit values $\bar{\alpha}_\infty^\epsilon$ and the asymptotic coefficients $\bar{\beta}_l^\epsilon$ for $\epsilon = \frac{1}{10}$ and $2 \le l \le 5$.

| | $n=2$ | $n=3$ | $n=4$ |
|---|---|---|---|
| $\bar{\alpha}_\infty^\epsilon$ | $-1.16597562512247\ldots$ | $-23.4104844051732\ldots$ | $-9.67410358657990\ldots$ |
| $\bar{\beta}_2^\epsilon$ | $-43.6306192301996\ldots$ | $1010.18554836853\ldots$ | $-262.997174628752\ldots$ |
| $\bar{\beta}_3^\epsilon$ | $148.054861925489\ldots$ | $2395.77105626966\ldots$ | $4088.05738204272\ldots$ |
| $\bar{\beta}_4^\epsilon$ | $6774.23466092879\ldots$ | $-8446.72532630471\ldots$ | $31947.0851156197\ldots$ |
| $\bar{\beta}_5^\epsilon$ | $-9940.25385554344\ldots$ | $-308104.356004984\ldots$ | $-238298.206234643\ldots$ |

Table 2
The limit values $\bar{\alpha}_\infty^\epsilon$ for $\epsilon = 10^{-k}$ and $n = 2$.

| $k$ | $\bar{\alpha}_\infty^\epsilon$ |
|---|---|
| 1 | $-1.16597562512246837352904013544346035854 2\ldots$ |
| 2 | $-7.44909253137167391142027154556207354347 5\ldots$ |
| 3 | $-7.94609838326256700513771793376378509718 9\ldots$ |
| 4 | $-7.99462160290839916697721348573772304673 9\ldots$ |
| 5 | $-7.99946227781080870290429937719644804594 4\ldots$ |
| 6 | $-7.99994622895615479174132885006975552607 1\ldots$ |
| 7 | $-7.99999462290736609263084925148440966284 2\ldots$ |
| 8 | $-7.99999946229085411527189515800315767707 9\ldots$ |
| 9 | $-7.99999994622908658658715186128423371770 7\ldots$ |
| 10 | $-7.99999999462290867040931468382642654013 3\ldots$ |
| 50 | $-8.00000000000000000000000000000000000000 0\ldots$ |

coefficients $\bar{\beta}_l^\epsilon$. It follows from our experiments that there exists some real numbers $\bar{\alpha}_\infty^0$ and $\bar{\beta}_l^0$ such that $\bar{\alpha}_\infty^\epsilon = \bar{\alpha}_\infty^0 + \mathrm{O}(\epsilon)$ and $\bar{\beta}_l^\epsilon = \bar{\beta}_l^0 + \mathrm{O}(\epsilon)$, for all $l \ge 2$. For instance, it suffices a fast look to table 2 to conclude that $\bar{\alpha}_\infty^0 = -8$ when $n = 2$. (All the decimal digits displayed in that table are correct.)

In the same way, we have computed the limits $\bar{\alpha}_\infty^0$ and the first asymptotic coefficients $\bar{\beta}_l^0$ for $n = 2, 3, 4$, see table 3. Again only fifteen decimal digits are displayed for lack of space. Once we obtained those limits, we performed a very simple analysis on them to check if they can be written in terms of celebrated constants like e or $\pi$. The amazing result is that

$$\bar{\beta}_2^0 = 14\pi^4/3, \quad \bar{\beta}_3^0 = -2\pi^4, \quad \bar{\beta}_4^0 = 2\pi^4/3 - 8\pi^8/5, \quad \bar{\beta}_5^0 = -9\pi^8/5$$

for the quartic perturbation: $n = 2$. We have not found so nice results for the

Table 3
The constants $\bar{\alpha}^0_\infty$ and $\bar{\beta}^0_l$ for $2 \le l \le 5$.

| | $n=2$ | $n=3$ | $n=4$ |
|---|---|---|---|
| $\bar{\alpha}^0_\infty$ | $-8.00000000000000\ldots$ | $-28.7884960229620\ldots$ | $-9.88592076539110\ldots$ |
| $\bar{\beta}^0_2$ | $454.575758158678\ldots$ | $1357.89383552383\ldots$ | $-279.105562715324\ldots$ |
| $\bar{\beta}^0_3$ | $-194.818182068004\ldots$ | $2716.60085733978\ldots$ | $4225.89824119769\ldots$ |
| $\bar{\beta}^0_4$ | $-15116.7102316902\ldots$ | $-27480.9190533494\ldots$ | $32532.1232487986\ldots$ |
| $\bar{\beta}^0_5$ | $-17079.3558289270\ldots$ | $-355090.495717297\ldots$ | $-239903.150880230\ldots$ |

perturbations of degree six and eight.

We study now the asymptotic behavior of the Gevrey coefficients $\omega^\epsilon_j$. To be more precise, we present some evidences on the asymptotic behavior of the sequences $(\bar{\omega}^0_j)_{j\ge 0}$, $(\bar{\omega}^1_j)_{j\ge 0}$, and $(\hat{\omega}^\epsilon_j)_{j\ge 0}$ defined in (12). We have plotted the first terms of these sequences in figures 8, 9, and 10, respectively, for $\epsilon = \frac{1}{10}$ and $n = 2,3,4$. For each index $j \ge 0$, the terms $\bar{\omega}^0_j$, $\bar{\omega}^1_j$, and $\hat{\omega}^\epsilon_j$ have been obtained by means of an extrapolation in the perturbative parameter $\epsilon$, from the computed values of $\omega^\epsilon_j$ in the net $\epsilon = 10^{-1}, 10^{-2}, \ldots, 10^{-10}$. The figures strongly suggest that the limits

$$\bar{\omega}^0_\infty = \lim_{j\to+\infty} \bar{\omega}^0_j, \qquad \bar{\omega}^1_\infty = \lim_{j\to+\infty} \bar{\omega}^1_j, \qquad \hat{\omega}^\epsilon_\infty = \lim_{j\to+\infty} \hat{\omega}^\epsilon_j$$

exist, are finite, and do not vanish. As before, these limits have been obtained by means of another extrapolation algorithm based on the hypothesis that there exist some asymptotic coefficients $\bar{\eta}^0_l$, $\bar{\eta}^1_l$ and $\hat{\eta}^\epsilon_l$ such that

$$\bar{\omega}^0_j \overset{\text{as}}{=} \bar{\omega}^0_\infty + \sum_{l\ge 2} \bar{\eta}^0_l j^{-l}, \qquad \bar{\omega}^1_j \overset{\text{as}}{=} \bar{\omega}^1_\infty + \sum_{l\ge 2} \bar{\eta}^1_l j^{-l}, \qquad \hat{\omega}^\epsilon_j \overset{\text{as}}{=} \hat{\omega}^\epsilon_\infty + \sum_{l\ge 1} \hat{\eta}^\epsilon_l j^{-l}$$

as $j \to +\infty$. For instance, we have listed in table 4 the limits $\bar{\omega}^0_\infty$ and the first asymptotic coefficients $\bar{\eta}^0_l$ for $n = 2,3,4$. The tables with the limits $\bar{\omega}^1_\infty$ and $\hat{\omega}^\epsilon_\infty$ jointly with the first asymptotic coefficients $\bar{\eta}^1_l$ and $\hat{\eta}^\epsilon_l$ have been skipped for the sake of brevity. When the perturbation is quartic (that is, when $n = 2$), we see that $\bar{\omega}^0_\infty = 2\bar{\alpha}^0_\infty$ and $\bar{\eta}^0_l = 2\bar{\beta}^0_l$, for all $l \ge 2$, compare tables 3 and 4.

## 5 Almost invisible homoclinic bifurcations in a billiard table

The previous section dealed with some splitting quantities associated to the eight axial homoclinic trajectories under the monomial perturbations (1). For more general perturbations there exist other primary homoclinic trajectories

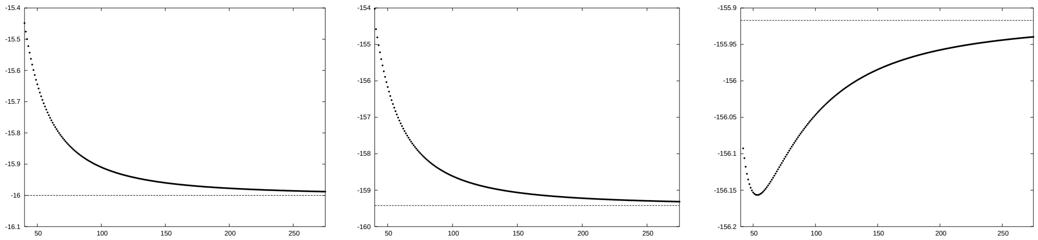


Fig. 8. The coefficients $\bar{\omega}_j^0$ versus $j$, for $n = 2$ (left), $n = 3$ (center), and $n = 4$ (right). The dashed lines correspond to the limits $\bar{\omega}_\infty^0 := \lim_{j\to+\infty} \bar{\omega}_j^0$.

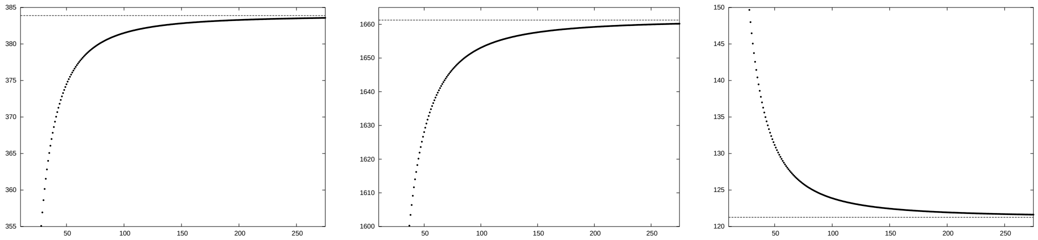


Fig. 9. The coefficients $\bar{\omega}_j^1$ versus $j$, for $n = 2$ (left), $n = 3$ (center), and $n = 4$ (right). The dashed lines correspond to the limits $\bar{\omega}_\infty^1 := \lim_{j\to+\infty} \bar{\omega}_j^1$.

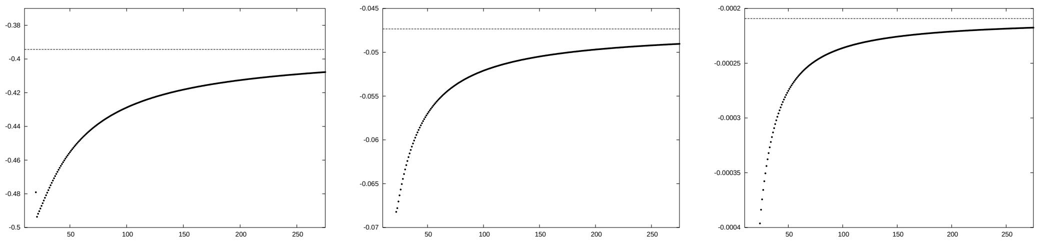


Fig. 10. The coefficients $\hat{\omega}_j^\epsilon$ versus $j$, for $\epsilon = \frac{1}{10}$. Here, $n = 2$ (left), $n = 3$ (center), and $n = 4$ (right). The dashed lines correspond to the limits $\hat{\omega}_\infty^\epsilon := \lim_{j\to+\infty} \hat{\omega}_j^\epsilon$.

Table 4
The constants $\bar{\omega}_\infty^0$ and $\bar{\eta}_l^0$ for $2 \leq l \leq 5$.

| | $n = 2$ | $n = 3$ | $n = 4$ |
|---|---|---|---|
| $\bar{\omega}_\infty^0$ | $-16.0000000000000\ldots$ | $-159.420897719167\ldots$ | $-155.917027114383\ldots$ |
| $\bar{\eta}_2^0$ | $909.151516317356\ldots$ | $7964.04298581449\ldots$ | $-1927.69790108031\ldots$ |
| $\bar{\eta}_3^0$ | $-389.636364136009\ldots$ | $12485.6464528904\ldots$ | $58142.9600713378\ldots$ |
| $\bar{\eta}_4^0$ | $-30233.4204633805\ldots$ | $-172192.452195518\ldots$ | $541110.879347534\ldots$ |
| $\bar{\eta}_5^0$ | $-34158.7116578540\ldots$ | $-1896348.13662865\ldots$ | $-4664320.82576045\ldots$ |

and several primary homoclinic bifurcations take place. Since this work is devoted to singular problems, we restrict our attention to the bifurcations that appear for small values of the characteristic exponent. We shall detect (and describe) some almost invisible homoclinic bifurcations under a binomial perturbation. They take place in an exponentially small range of the parameter space in the singular limit $h \to 0^+$. V.F. Lazutkin [16] studied analytically a similar problem in the framework of generalized standard maps.

To begin with, let us present some surprising bounds on the number $\rho$ of primary homoclinic trajectories inside the polynomial perturbation (3). These bounds depend only on the degree of the perturbation. Their proof has been relegated to §A.2.

**Lemma 7** *Given any polynomial $P(s)$ such that $P(0) = P'(0) = 0$ and $n := \deg[P(s)] \geq 2$, there exists $\epsilon_0 = \epsilon_0(P) > 0$ such that*

$$\rho \in \{8, 16, 24, \ldots, 8(n-1)\} \qquad \forall \epsilon \in (0, \epsilon_0).$$

As a corollary of this lemma, we get that the only primary homoclinic orbits inside small enough quartic perturbations (that is, when $n = 2$) are the eight axial homoclinic orbits, because $\rho \in \{8\}$. Thus, the case $n = 3$ is the first scenario in which is possible to find homoclinic bifurcations. Consequently, let us consider the perturbation of degree six

$$C_d = \left\{(x, y) \in \mathbb{R}^2 : x^2/a^2 + y^2/b^2 + \epsilon(y^2/\gamma^2 + d)y^4/\gamma^4 = 1\right\} \tag{14}$$

associated to the monic polynomial $P(s) = s^3 + ds^2$, where $d \in \mathbb{R}$ is an additional parameter. Under this perturbation, we know from lemma 7 that $\rho = \rho(d) \in \{8, 16\}$ for small enough $\epsilon$, and so the following questions are quite natural. Are both values ($\rho = 8$ and $\rho = 16$) realized? What homoclinic bifurcations take place when $\rho$ changes? The first rough numerical explorations in the space of parameters ($h$, $\epsilon$ and $d$) misleadingly suggest that $\rho \equiv 8$ for small values of $h$. Let us explain in detail what bifurcations really occur and why they are almost invisible in the singular case.

We have explained in §2.5 that there exists a real analytic $h$-periodic function $\Psi(t)$, called the *splitting function*, whose roots are in 1-to-1 correspondence with the primary homoclinic points. We also recall that our billiard problem has four separatrices, and so we are confronted to four splitting functions. But, due to the symmetries, these four functions coincide and it suffices to find the number of roots modulo $h$ of one of them and then to quadruplicate it to get the total number of primary homoclinic orbits. Moreover, since our problem is reversible, we can normalize $\Psi(t)$ in such a way that it becomes odd and the axial homoclinic points are located at the points $t_+ \equiv 0 \pmod h$ and $t_- \equiv h/2 \pmod h$. Then, the transition from $\rho = 8 = 4 \times 2$ to $\rho = 16 = 4 \times 4$ can only occur as displayed in the qualitative pictures of figure 11. That is, at the bifurcations $\Psi(t)$ has a triple root at $t = t_+$ or else at $t = t_-$. This means, in a more geometric language, that the bifurcations always take place through cubic tangencies of the invariant curves along the four $y$-axial homoclinic orbits

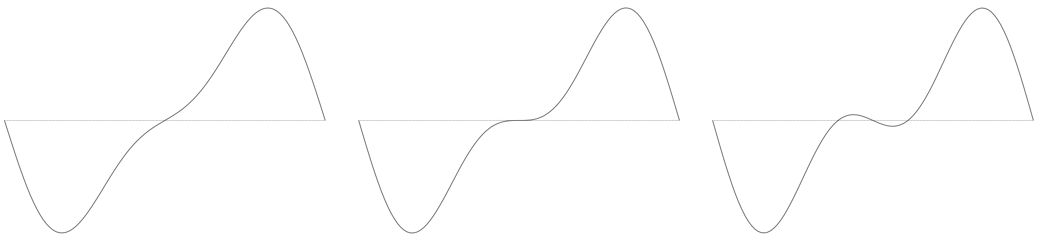

Fig. 11. The graph of the splitting function $\Psi(t)$ in a fundamental domain before bifurcation (left), just at bifurcation (center), and after bifurcation (right).

or else along the four $x$-axial homoclinic ones.

Consequently, we can reduce the search of primary homoclinic bifurcations inside the symmetric curve (14) to the study of the vanishing of the Lazutkin's invariants $\omega^\pm$ of the axial homoclinic trajectories. We also shall study the vanishing of the area $A$ of the lobe enclosed between them for completeness. Of course, these quantities depend on $h$, $\epsilon$ and $d$. Once fixed $\epsilon$ and $h$, we look for changes as the additional parameter $d$ varies.

In order to accomplish it, we consider the normalized functions [3]

$$\bar{A}(d) := \epsilon^{-1}\, \mathrm{e}^{\pi^2/h}\, A(h,\epsilon,d), \qquad \bar{\omega}^\pm(d) := \epsilon^{-1} h^2\, \mathrm{e}^{\pi^2/h}\, \omega^\pm(h,\epsilon,d)/2\pi^2.$$

We have plotted in figure 12 their graphs in the range $d \in [-2,0]$, for $\epsilon = \frac{1}{2}$ and $h = 1,2$. Let $d_* = d_*^\epsilon(h)$ and $d_\pm = d_\pm^\epsilon(h)$ be the only roots of these functions in the interval $[-2,0]$. In turns out that $d_+ < d_* < d_-$. Accordingly to the above digression, the roots $d_\pm$ are *values of bifurcation*. In fact, after a careful computation and visualization of the invariant curves, we obtain the following results. At $d = d_+$ (respectively, $d = d_-$) the invariant curves have a cubic tangency at the four $y$-axial (respectively, $x$-axial) homoclinic orbits. Besides, $\rho = 8$ for $-2 \le d \le d_+$ and $d_- \le d \le 0$, whereas $\rho = 16$ for $d \in (d_+, d_-)$.

We note that if $h$ is relatively small, the three roots $d_*$, $d_+$ and $d_-$ become almost indistinguishable. See, for instance, the case $h = 1$ in figure 12. In fact, the functions $\bar{A}(d)$ and $\bar{\omega}^+(d)$ match absolutely at the scale of that picture. To see this more clearly, we have plotted in figure 13 the graphs of the functions $d_-^\epsilon(h)$, $d_*^\epsilon(h)$, and $d_+^\epsilon(h)$, for $\epsilon = \frac{1}{10}$ in several ranges. In the last picture of that figure, when $0 < h < 1$, it is very difficult to distinguish the three graphs. (Apparently, the limits $\lim_{h\to+\infty} d_\pm^\epsilon(h)$ and $\lim_{h\to+\infty} d_*^\epsilon(h)$ there exist and are finite, but they fall out of the scope of this work.)

We shall check that the bifurcation values $d_+^\epsilon(h)$ and $d_-^\epsilon(h)$ are exponentially close in $h$, and so it is difficult to detect these bifurcations using only brute

[3] The factors $\epsilon^{-1}\, \mathrm{e}^{\pi^2/h}$ and $\epsilon^{-1} h^2\, \mathrm{e}^{\pi^2/h}/2\pi^2$ regularize the singular behavior of the quantities $A$ and $\omega^\pm$ as $\epsilon \to 0$ and $h \to 0^+$. Besides, $\bar{A} \sim \pm\bar{\omega}^\pm$ as $\epsilon \to 0$ and $h \to 0^+$.

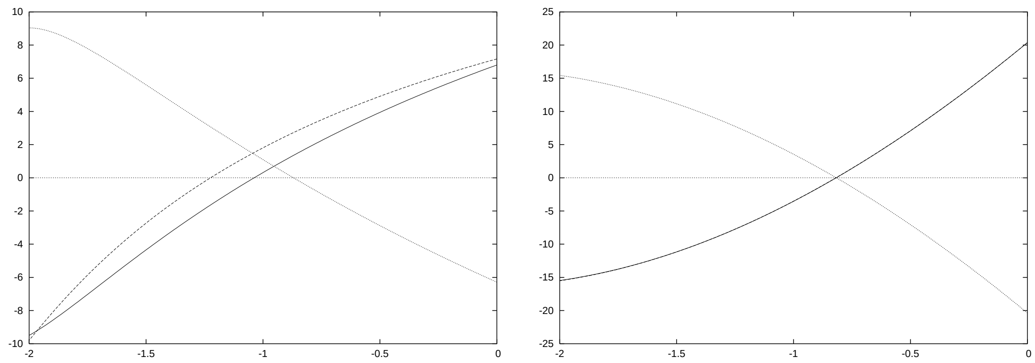


Fig. 12. The graphs of the functions $\bar{A}(d)$ (continuous lines), $\bar{\omega}^-(d)$ (dotted lines), and $\bar{\omega}^+(d)$ (dashed lines), for $h = 2$ (left) and $h = 1$ (right). Here, $\epsilon = \frac{1}{2}$.

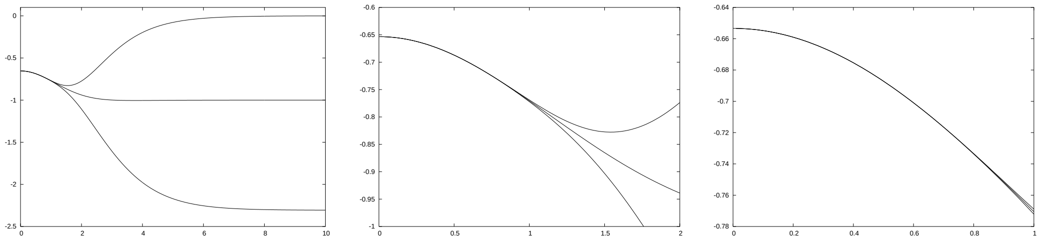

Fig. 13. The graphs of the functions $d_-^\epsilon(h)$, $d_*^\epsilon(h)$, and $d_+^\epsilon(h)$, for $\epsilon = \frac{1}{10}$ in the ranges $0 < h < 10$ (left), $0 < h < 2$ (center), and $0 < h < 1$ (right).

force methods. Due to this, as a first step in our study, we shall derive some analytical predictions of the bifurcation values by using the Melnikov method described in §4. The Melnikov potential associated to the binomial perturbation (14) is linear in the parameter $d$. Concretely, it is the elliptic function

$$L(t; d) = L_3(t) + dL_2(t),$$

where $L_n(t)$ is defined in (4). The function $L_n(t)$ is the Melnikov potential associated to the monomial perturbation (1). Thus, if $d_\pm^\epsilon(h) = d_\pm^0(h) + \mathrm{O}(\epsilon)$ are the bifurcation values, their Melnikov approximations $d_\pm^0(h)$ are given by

$$d_\pm^0(h) = -L_3''(t_\pm)/L_2''(t_\pm), \qquad \text{where } t_+ = 0 \text{ and } t_- = h/2, \tag{15}$$

because they are the roots of the functions $L''(t_\pm; \cdot)$. In the same way, if $d_*^\epsilon(h) = d_*^0(h) + \mathrm{O}(\epsilon)$, then

$$d_*^0(h) = \frac{L_3(h/2) - L_3(0)}{L_2(0) - L_2(h/2)},$$

since $d_*^0(h)$ is the root of the function $L(h/2; \cdot) - L(0; \cdot)$. We have computed these Melnikov approximations in §A.3. The result is:

$$\left. \begin{aligned} d^0_+(h), d^0_*(h), d^0_-(h) &= d^0(h) + \mathrm{O}(\mathrm{e}^{-\pi^2/h}), \\ d^0_-(h) - d^0_+(h) &= \mathrm{e}^{-\pi^2/h}\, \delta^0(h) + \mathrm{O}(\mathrm{e}^{-2\pi^2/h}) \end{aligned} \right\} \qquad (h \to 0^+), \tag{16}$$

where $d^0(h)$ and $\delta^0(h)$ are the even analytic functions

$$\begin{aligned} d^0(h) &= \frac{1}{\cosh^2(h/2)} - \frac{2(\pi^2 + h^2)\tanh^2(h/2)}{3h^2} = 1 - \frac{\pi^2}{6} + \frac{\pi^2 - 15}{36} h^2 + \cdots \\ \delta^0(h) &= 32\pi^2 h^{-2} \tanh^2(h/2) = 8\pi^2 \left(1 - \frac{h^2}{6} + \frac{17h^4}{720} - \frac{31h^6}{10080} + \cdots\right). \end{aligned}$$

Hence, we guess that the bifurcation values $d^\epsilon_\pm(h)$ are close to $d^0_0 := d^0(0) = 1 - \pi^2/6 \simeq -0.645$, for small enough values of $h$ and $\epsilon$. See, for instance, figure 13. We also have the following exponentially small Melnikov prediction for the range in which the primary homoclinic bifurcations take place:

$$\Delta_{\mathrm{Bifur}} := d^\epsilon_-(h) - d^\epsilon_+(h) \sim \mathrm{e}^{-\pi^2/h}\, \delta^0_0 \tag{17}$$

for small enough values of $h$ and $\epsilon$. Here, $\delta^0_0 := \delta^0(0) = 8\pi^2$.

We state below a conjecture on the asymptotic behavior of the bifurcation range $\Delta_{\mathrm{Bifur}}$. If it was true, the prediction (17) would give the right answer in the singular perturbative case, whereas in the singular nonperturbative case the asymptotic constant $\delta^0_0 = 8\pi^2$ would have to be substituted by a new asymptotic constant $\delta^\epsilon_0 = 8\pi^2 + \mathrm{O}(\epsilon)$.

**Conjecture 8** *For any small enough $\epsilon \neq 0$, there exists a series $\sum_{j\geq 0} \delta^\epsilon_j h^{2j}$ such that the following asymptotic expansion holds:*

$$\Delta_{\mathrm{Bifur}} \stackrel{\mathrm{as}}{=} \mathrm{e}^{-\pi^2/h} \sum_{j\geq 0} \delta^\epsilon_j h^{2j} \qquad (h \to 0^+, \epsilon \textit{ fixed}).$$

*Besides, if $\sum_{j\geq 0} \delta^0_j h^{2j}$ is the Taylor expansion of $\delta^0(h) = 32\pi^2 h^{-2}\tanh^2(h/2)$, then $\delta^\epsilon_j = \delta^0_j + \mathrm{O}(\epsilon)$ for all $j \geq 0$. In particular, $\delta^\epsilon_0 = 8\pi^2 + \mathrm{O}(\epsilon)$.*

We have limited the exposition of the numerical evidences supporting this conjecture, for the sake of brevity, to one table. Namely, table 5, in which it can be realized very clearly that $\delta^\epsilon_0 = 8\pi^2 + \mathrm{O}(\epsilon)$. The computation of other asymptotic coefficients gives rise to similar tables.

Table 5
The constant $\delta_0^\epsilon$ for $\epsilon = 10^{-k}$. The last row contains the limit $\delta_0^0 = 8\pi^2$.

| $k$ | $\delta_0^\epsilon$ |
|---|---|
| 10 | 78.9568351978330244918250679636819034908091432514355285038189 ... |
| 20 | 78.9568352087148689495877435531181542121109228680918690955383 ... |
| 30 | 78.9568352087148689506759279989003906379204897708863980836546 ... |
| 40 | 78.9568352087148689506759279990092090824987134134674144626028 ... |
| 50 | 78.9568352087148689506759279990092090825095952579252368268609 ... |
| $\infty$ | 78.9568352087148689506759279990092090825095952579263250113068 ... |

## 6 Conclusion

We have formulated several conjectures related to the singular phenomena that appear in some billiards tables. We have also presented some very accurate numerical experiments, which support strongly both conjectures. The next step must be to prove these conjectures. We hope that this problem would be a stimulating challenge for some readers, whether in this framework with billiard maps or in the much more celebrated about generalized standard maps.


## Acknowledgements

The author has been partially supported by the Grant CIRIT 2001SGR-70 (Catalonia) and the MCyT-FEDER Grant BFM2003-9504 (Spain). I would like to thank to Leopold Palomo, Pablo Sánchez, and Jaume Timoneda for their efforts to show me the possibilities of Beowulf clusters. I am also very grateful to R. de la Llave for its valuable advises on the PARI system. Besides, I am indebted to A. Delshams, E. Fontich, and C. Simó for very useful remarks and comments. Finally, it is a pleasant obligation to express my gratitude to the PARI developers that make possible the existence of an excellent tool.


## A Some details of the Melnikov computations

We have adopted a very compact style along this appendix, avoiding to give the more cumbersome and less crucial details in the computations, because similar ones can be found in the literature. See, for instance, [4,5,7].

## A.1 Melnikov potential for monomial perturbations

Here, we address some computations related to the elliptic Melnikov potential $L_n(t) = ae^{2n} \sum_{k\in\mathbb{Z}} \ell_n(t+kh)$ introduced in (4). All the computed objects are expressed in terms of the even derivatives of the function

$$\psi(t) = (2K/h)^2 \operatorname{dn}^2(2Kt/h, k),$$

where $\operatorname{dn}(u) = \operatorname{dn}(u,k)$ is one of the twelve Jacobian elliptic functions. Here, $k$ is the *modulus*, $K = \int_0^{\pi/2}(1-k^2\sin u)^{-1/2}\,\mathrm{d}u$ is the *complete elliptic integral of the first kind*, and $K' = \int_0^{\pi/2}(1-k'^2\sin u)^{-1/2}\,\mathrm{d}u$ where $k'$ is the *complementary modulus*: $k^2+k'^2=1$. Finally, the quantity $q = \mathrm{e}^{-\pi K'/K}$ is called the *nome*. We refer to [24] for a general background on elliptic functions.

We recall that *elliptic functions are characterized (modulo additive constants) by their periods, poles and principal parts*: the difference of elliptic functions with the same periods, poles and principal parts is a bounded entire function, and hence constant by Liouville's theorem.

Thus, we are naturally led to the search for the poles (and their principal parts) of $L_n(t) = ae^{2n} \sum_{k\in\mathbb{Z}} \ell_n(t+kh)$.

The function $\ell_n(t)$ is meromorphic and $\pi\mathrm{i}$-periodic. Its poles are the points in the sets $\pi\mathrm{i}/2 + \pi\mathrm{i}\mathbb{Z}$ and $\pi\mathrm{i}/2 \pm h/2 + \pi\mathrm{i}\mathbb{Z}$. From now on, $a_j(f,\tau)$ stands for the coefficient of the term $(t-\tau)^j$ in the Laurent expansion of a meromorphic function $f(t)$ around $t=\tau$. The poles $t_0^\pm \in \pi\mathrm{i}/2 \pm h/2 + \pi\mathrm{i}\mathbb{Z}$ are simple and, due to the symmetry, $a_{-1}(\ell_n, t_0^+) + a_{-1}(\ell_n, t_0^-) = 0$. The poles $t_0 \in \pi\mathrm{i}/2 + \pi\mathrm{i}\mathbb{Z}$ have order $2n-2$ and $a_{-j}(\ell_n, t_0) = 0$ for all odd integers $j \geq 1$.

Therefore, $L_n(t) = ae^{2n} \sum_{k\in\mathbb{Z}} \ell_n(t+kh)$ is an elliptic function characterized (modulo an additive constant) by the following properties: (1) Its periods are $h$ and $\pi\mathrm{i}$; (2) Its poles are the points in the set $\pi\mathrm{i}/2 + h\mathbb{Z} + \pi\mathrm{i}\mathbb{Z}$; and (3) Its principal part around a pole $t_0$ is $ae^{2n}\sum_{j=1}^{n-1} a_{-2j}(\ell_n, \pi\mathrm{i}/2)(t-t_0)^{-2j}$.

On the other hand, the square of the Jacobian elliptic function $\operatorname{dn}(u) = \operatorname{dn}(u,k)$ is characterized (modulo an additive constant) by the properties: (1') Its periods are $2K$ and $2K'\mathrm{i}$; (2') Its poles are the points in the set $K'\mathrm{i} + 2K\mathbb{Z} + 2K'\mathrm{i}\mathbb{Z}$; and (3') The principal part around any pole $u_0$ is $-(u-u_0)^{-2}$, see [24, §22].

Hence, if we take $q = \mathrm{e}^{-\pi^2/h}$, then $K' = K\pi/h$ and

$$L_n(t) = \text{constant} - ae^{2n} \sum_{j=1}^{n-1} \frac{\xi_{n,j}(h)}{(2j-1)!}\psi^{(2j-2)}(t). \tag{A.1}$$

where $\xi_{n,j}(h) = a_{-2j}(\ell_n, \pi\mathrm{i}/2)$. In particular, the Melnikov terms (5) are

$$A_1(h) = ae^{2n}\sum_{j=1}^{n-1}\frac{\xi_{n,j}(h)\delta_j(h)}{(2j-1)!}, \qquad \Omega_1(h) = -ae^{2n}\sum_{j=1}^{n-1}\frac{\xi_{n,j}(h)\sigma_j(h)}{(2j-1)!}$$

where $\delta_j(h) := \psi^{(2j-2)}(0) - \psi^{(2j-2)}(h/2)$ and $\sigma_j(h) := \psi^{(2j)}(0) + \psi^{(2j)}(h/2)$. We need two lemmas to study the asymptotic behavior of the above quantities.

**Lemma 9** *For any $1 \le j \le n-1$, there exists an even analytic function $\bar{\xi}_{n,j}(h)$ such that $\bar{\xi}_{n,j}(0) = (-1)^n 4^{n-j}$ and $\xi_{n,j}(h) = h^{2(j-n)}\bar{\xi}_{n,j}(h)$.*

The above lemma is obtained using a trick contained in [4, §4.4.4].

**Lemma 10** *If $T = 2\pi/h$, then*

*a)* $\psi(t) = 2T^2[1/8 + q^2 + q\cos Tt + 2q^2\cos 2Tt + \mathrm{O}(q^3)]$,
*b)* $\psi^{(4)}(t_\pm)/\psi''(t_\pm) = -T^2(1 \pm 24q + \mathrm{O}(q^2))$, *where $t_+ = 0$ and $t_- = h/2$.*
*c)* $\delta_j(h) := \psi^{(2j-2)}(0) - \psi^{(2j-2)}(h/2) = 4T^{2j}q[(-1)^{j-1} + \mathrm{O}(q)]$, *for all $j \ge 1$.*
*d)* $\sigma_j(h) := \psi^{(2j)}(0) + \psi^{(2j)}(h/2) = 2(2T)^{2j+2}q^2[(-1)^j + \mathrm{O}(q)]$, *for all $j \ge 1$.*

**PROOF.** The first item follows directly from the definition of $\psi(t)$ and the Fourier expansion

$$\mathrm{dn}(2Kt/h, k) = \frac{\pi}{2K} + \frac{2\pi}{K}\sum_{n\ge 1}\frac{q^n}{1+q^{2n}}\cos(nTt),$$

which can be found in [24, page 511]. The others follow from the first one. □

Now we are ready to prove the formulae for the constants $\alpha_0^0 = \alpha^0(0)$ and $\omega_0^0 = \omega^0(0)$ given in (7). For instance, using the couple of lemmas we see that

$$\alpha^0(h) = 4e^{2n}\sum_{j=1}^{n-1}\frac{(-1)^{j-1}T^{2j}}{(2j-1)!}\xi_{n,j}(h) = 4(e/h)^{2n}\sum_{j=1}^{n-1}\frac{(-1)^{j-1}(2\pi)^{2j}}{(2j-1)!}\bar{\xi}_{n,j}(h)$$

and $\alpha_0^0 = \alpha^0(0) = 4^{1-n}\sum_{j=1}^{n-1}\frac{(-1)^{j-1}(2\pi)^{2j}}{(2j-1)!}(-1)^n 4^{n-j} = (-1)^n 4\pi^2\sum_{j=0}^{n-2}\frac{(-1)^j\pi^{2j}}{(2j+1)!}$. We recall that $e = \tanh(h/2) \sim h/2$ and $\bar{\xi}_{n,j}(0) = (-1)^n 4^{n-j}$. The constant $\omega_0^0$ is obtained in the same way.

Finally, we are going to check that when the monomial perturbation is quartic the functions $\alpha^0(h)$ and $\omega^0(h)$ have the form given in (8). If $n = 2$, then

$$\alpha^0(h) = 4e^4T^2\xi_{2,1}(h) = 16\pi^2h^{-2}\tanh^4(h/2)a_{-2}(\ell_2, \pi\mathrm{i}/2),$$

and a straightforward computation shows that

$$a_{-2}(\ell_2, \pi\mathrm{i}/2) = \frac{1}{\sinh^2(h/2)}.$$

The proof for $\omega^0(h)$ follows the same lines.

### *A.2 Proof of the bounds on the number of primary homoclinic orbits*

Here we shall prove the three claims contained in lemma 7.

The first claim is the lower bound $\rho \geq 8$, which is trivial. It does not require additional comments, because we already know that the eight axial homoclinic trajectories persist under small enough symmetric perturbations.

The second claim is the upper bound $\rho \leq 8(n-1)$. We know that in our billiard problem there exists a real analytic $h$-periodic odd function $\Psi(t)$, called the splitting function, whose roots modulo $h\mathbb{Z}$ are in one-to-four correspondence with the primary homoclinic trajectories. (The factor four has been explained at the beginning of §5.) We also recall that $\Psi(t) = \Theta'(t) = \epsilon L'(t) + \mathrm{O}(\epsilon)$, where $L'(t)$ is the derivative of the Melnikov potential associated to the polynomial perturbation (3). Therefore, to prove the upper bound, it suffices to see that $L'(t)$ has at the most $2n-2$ roots in $\mathbb{R}/h\mathbb{Z}$, counted with multiplicity.

If $P(s) = \sum_{j=2}^{n} p_j s^j$, then $L(t) = \sum_{j=2}^{n} p_j L_j(t)$, where the functions $L_j(t)$ are defined in (4). From the properties of the functions $L_j(t)$ listed in §A.1, we deduce that the only poles of the elliptic function $L(t)$ are the points in the set $\pi\mathrm{i}/2 + \mathbb{Z}h + \mathbb{Z}\pi\mathrm{i}$, and all of them have order $2n-2$. Its derivative $L'(t)$ has the same poles, but of order $2n-1$. Thus, $L'(t)$ has exactly $2n-1$ roots in any complex cell. (Non-constant elliptic functions have the same number of poles and roots in a cell, counted with multiplicity.) Besides, using the symmetry $L'(-t) = -L'(t)$ and the periodicity $L'(t + h + 2\pi\mathrm{i}) = L'(t)$, we obtain that $L'(h/2 + \pi\mathrm{i}) = -L'(h/2 + \pi\mathrm{i})$. So, $h/2 + \pi\mathrm{i}$ is a complex root of $L'(t)$ and the number of real roots (modulo $h\mathbb{Z}$) of $L'(t)$ is less or equal than $2n-2$.

The last claim is that the number of primary homoclinic trajectories changes in eights. This property follows from two facts. First, due to the symmetries, the four separatrices have the same homoclinic bifurcations, so that they appear in fours. Second, due to the reversibility, the $h$-periodic splitting function $\Psi(t)$ is odd and the axial homoclinic points are located at the points $t_+ \equiv 0 \pmod{h}$ and $t_- \equiv h/2 \pmod{h}$. Then, it turns out that any change in the interval $(0, h/2)$ generates a twin change in $(h/2, h)$ —see, for instance, figure 11—, and the number of roots of $\Psi(t)$ changes in twos. Finally, $8 = 4 \times 2$.

We recall the notation $\xi_{n,j}(h) = a_{-2j}(\ell_n, \pi\mathrm{i}/2)$ introduced in §A.1. If we set $\eta^{-1} = \sinh(h/2)$, it is easy to check that $\xi_{2,1}(h) = \eta^2$, $\xi_{3,1}(h) = \eta^2(2/3 - \eta^2)$, and $\xi_{3,2}(h) = -\eta^2$. Thus, using formula (A.1) we find the expressions

$$L_2''(t) = -ae^4\eta^2\psi''(t), \qquad L_3''(t) = ae^6\eta^2\left((\eta^2 - 2/3)\psi''(t) + \psi^{(4)}(t)/6\right).$$

Finally, from the second item contained in lemma 10 we see that the Melnikov approximations of the bifurcation values $d^\epsilon_\pm(h) = d^0_\pm(h) + \mathrm{O}(\epsilon)$ defined in (15) verify the asymptotic estimate

$$d^0_\pm(h) = -\frac{L_3''(t_\pm)}{L_2''(t_\pm)} = e^2\left(\eta^2 - \frac{2}{3} + \frac{\psi^{(4)}(t_\pm)}{6\psi''(t_\pm)}\right) = \frac{e^2(6\eta^2 - 4 - T^2)}{6} + \mathrm{O}(q),$$

whereas their difference is exponentially small:

$$d^0_-(h) - d^0_+(h) = \frac{e^2}{6}\left(\frac{\psi^{(4)}(t_-)}{\psi''(t_-)} - \frac{\psi^{(4)}(t_+)}{\psi''(t_+)}\right) = 8T^2e^2q + \mathrm{O}(q^2).$$

This implies that the functions $d^0(h)$ and $\delta^0(h)$ defined implicitly in (16) are $d^0(h) = e^2(\eta^2 - 2/3 - T^2/6)$ and $\delta^0(h) = 8T^2e^2$. To obtain their final forms, it suffices to recall that $e = \tanh(h/2)$, $T = 2\pi/h$, and $\eta^{-1} = \sinh(h/2)$. The computation of $d^0_*(h)$ follow the same lines. We skip it.

## B Some details of the numerical computations

In this appendix we shall explain the ideas behind the computations and we shall describe some of the algorithms used. We shall emphasize the points that depend strongly on the peculiarities of billiard maps, since the others have already been described in [6,19].

### *B.1 Multiple-precision arithmetic: main problems and basic principles*

To begin with, let us explain why the use of a multiple-precision arithmetic is necessary. We recall that the lobe area is computed as the difference of actions $A = W[O^-] - W[O^+]$, see §2.1. This difference causes an important loss of significant digits, even for moderate values of $h$. It can only be overcome by computing the actions with more correct digits than the lost ones. For sample, let us set $a = 1$, $\epsilon = \frac{1}{10}$, and $n = 2$. Under this setup, the actions are

$$W[O^+] \simeq -0.09992503852271571273156135027523196060135775467,$$
$$W[O^-] \simeq -0.09992503852271571273156135027523196060135716286,$$

for $h = 10^{-1}$. Then $A \simeq 5.9181 \cdot 10^{-43}$ and so the cancellation causes the loss of more than 40 digits. Smaller values of $h$ cause stronger cancellations: $A \simeq 1.0137 \cdot 10^{-428}$ for $h = 10^{-2}$, and $A \simeq 2.1022 \cdot 10^{-4286}$ for $h = 10^{-3}$. The cancellations are even worse in the computation of $\Omega = \omega^+ + \omega^-$. And to top it all, in order to get hundreds of coefficients in the expansions (10) we use an extrapolation method, which shall require the computation of $A$ and $\Omega$ with a very high precision for hundreds of small values of $h$.

Our programs have been written using the PARI system [1]. The PARI system is a package capable of doing formal computation on recursive typees at high speed. Although it is possible to use PARI as a C library, we have used it as a sophisticated programmable calculator, which contain most of the control instructions of a standard laguage like C.

The main numerical difficulties that appear during the study of the singular splitting of separatrices of our billiard maps are the computation of:

- The billiard maps and their differentials with an arbitrary precision $P$;
- The Taylor expansions of the invariant curves up to an arbitrary order $K$;
- The homoclinic invariants $A$ and $\Omega$ with an arbitrary precision $Q$; and
- The Gevrey expansions (10) up to an arbitrary order $J$.

The quantities $Q$ and $J$ must be inputs of the algorithm, because they set some properties of the objects we are looking for. On the contrary, $P$ and $K$ are determined in an automatic way when the computation begins. For instance, in the computation of the lobe area $A = W[O^-] - W[O^+]$ the number of digits lost by cancellations is approximately equal to $R + S$, where $R = R(\epsilon) = |\log_{10} \epsilon|$ and $S = S(h) = \pi^2 h^{-1} \log_{10} \mathrm{e}$. Hence, to compute $A$ with precision $Q$ we must take $P \approx Q + R + S$. Analogously, we set $P \approx Q + R + 2S$ to compute the sum $\Omega = \omega^+ + \omega^-$ with the same precision. On the other hand, the order $K$ is determined under the following optimization criterion. If $K$ is too big, the Taylor expansions become too expensive, but a too low $K$ is also expensive, because then the number of iterates to reach the homoclinic points (from a local fundamental domain in which the Taylor expansion gives an enough accurate approximation of the invariant curves) grows too much. Therefore, there exists some optimal order for which the computations become the fastest ones. This optimal value can be estimated, see [6, §5D]. To acquaint the reader with the magnitude of our computations, we note that we have reached the values $Q = 1500$ and $J = 300$, with $P = 7000$ and $K = 1100$. In the previous literature there are not so extreme computations, being [6] the only one that reaches a comparable level.

The problem of the computation of the map and its differential is trivial for the standard map, the Hénon map, the perturbed McMillan maps, and others generalized standard maps. These maps have explicit expressions in terms of polynomial or trigonometric functions. The billiard map is slightly harder, since we have to solve a nonlinear equation to find the intersection of the reflected ray with the convex curve, see §B.2.

The local invariant curves of weakly hyperbolic objects must be developed up to high orders, see [19] for general comments. Then the initial iterates can be taken far enough from the hyperbolic object and so the homoclinic points can be attained in a few iterations. Here, few means thousands, instead of millions. In this way, undesirable accumulation errors due to the large amount of operations is avoided and computing time is reduced. The Taylor coefficients of the invariant curves of many analytic area-preserving maps can be obtained recursively. The recursive algorithm for our billiard maps is described in §B.3. Its derivation is less direct than for maps given by closed explicit formulae.

The method to compute the symmetric primary homoclinic points and their homoclinic invariants does not depend very much on the form of the map. The symmetric homoclinic points are found as the intersections between the unstable invariant curve and the corresponding symmetry lines. The global unstable curve is obtained from the local one by forward iteration of the map. The extrapolation method to obtain the first coefficients of the asymptotic expansion of the homoclinic invariants is very standard. We refer to [6] for a general background on these methods.

The main principle to design valid algorithms for the above computations is that, since the use of a multiple-precision arithmetic is unavoidable, we have to mitigate its cost in all the possible ways. To mention just the most obvious way, we shall solve any nonlinear equation by using the quadratically convergent Newton's method. Of course, we will begin the method in single precision and later we will refine the result by doubling the number of digits after each Newton iteration. This methodology causes a cascade of changes in the number of digits used along a concrete computation, because sometimes each evaluation of the initial nonlinear terms requires the solution of another nonlinear equation and so forth, but the increase in speed is spectacular.

### *B.2 Billiard maps and their differentials*

In §3.1 we have modeled billiards inside a closed convex curve $C$ by means of diffeomorphisms on an annulus, but from a numerical point of view it is better to model them by means of diffeomorphisms defined on the phase space

$$M = \{m = (q, p) \in C \times \mathbb{S} : p \text{ is directed outward } C \text{ at } q\}$$

consisting of points $q = (x, y) \in C$ and velocities $p = (u, v) \in \mathbb{S}$. That is, we use the four coordinates $x$, $y$, $u$, $v$, restricted to the conditions $(x, y) \in C$ and $u^2 + v^2 = 1$. Then the billiard map $f(q, p) = (q', p')$ is defined as follows. The new velocity $p'$ is the reflection of $p$ with respect to the tangent line $T_qC$. The new point $q'$ is determined by imposing that $q' = q - \tau p' \in C$ for some $\tau < 0$. The existence and uniqueness of $q'$ follows from the convexity of the curve $C$.

For brevity, hereafter we restrict the study to the monomial perturbations (1), which are convex for all $\epsilon \geq 0$ and for all integer $n \geq 2$. For further reference, we write their implicit equations as

$$x^2 = \mu_0 + \mu_1 y^2 + \mu_n y^{2n}, \tag{B.1}$$

where $\mu_0 = a^2$, $\mu_1 = -a^2/b^2$, and $\mu_n = -\epsilon a^2/\gamma^{2n}$. We look for an algorithm to compute the billiard map $f$ jointly with its differential $\mathrm{d}f$ as fast as possible with arbitrary (but fixed) accuracy, for relatively small values of $h$ and not very big values of $\epsilon$. Typically, $10^{-3} \leq h \leq 10^{-1}$ and $0 < \epsilon \leq 1$.

Given $q = (x, y) \in C$, $p = (u, v) \in \mathbb{S}$, $\dot{q} = (\dot{x}, \dot{y}) \in T_qC$, and $\dot{p} = (\dot{u}, \dot{v}) \in T_p\mathbb{S}$, we want to compute $(q', p') = f(q, p)$ and $[\dot{q}', \dot{p}'] = \mathrm{d}f(q, p)[\dot{q}, \dot{p}]$. We write $q' = (x', y') \in C$, $p' = (u', v') \in \mathbb{S}$, $\dot{q}' = (\dot{x}', \dot{y}') \in T_{q'}C$, and $\dot{p}' = (\dot{u}', \dot{v}') \in T_{p'}\mathbb{S}$. Using these notations, we perform the computation in two steps.

- *Computation of the new velocity:* We set $r = (\alpha, \beta)$ and $\dot{r} = (\dot{\alpha}, \dot{\beta})$, where $\alpha = x$, $\dot{\alpha} = \dot{x}$, $\beta = -(\mu_1 + n\mu_n y^{2n-2})y$ and $\dot{\beta} = -(\mu_1 + (2n-1)n\mu_n y^{2n-2})\dot{y}$. Then the vector $r$ is normal to the curve (B.1) at the point $q$. Therefore, $p' = p - \nu \cdot r$ and $\dot{p}' = \dot{p} - \nu \cdot \dot{r} - \dot{\nu} \cdot r$, where the quantities

$$\nu = 2\frac{\langle p, r\rangle}{\langle r, r\rangle}, \qquad \dot{\nu} = \frac{\langle p', \dot{p} - \nu \cdot \dot{r}\rangle}{\langle p', r\rangle}$$

  have been determined by imposing that $p' \in \mathbb{S}$ and $\dot{p}' \in T_{p'}\mathbb{S}$, respectively. We note that $p' \in \mathbb{S} \Leftrightarrow \langle p', p'\rangle = 1$ and $\dot{p}' \in T_{p'}\mathbb{S} \Leftrightarrow \langle p', \dot{p}'\rangle = 0$.
- *Computation of the new point:* $q' = q - \tau \cdot p'$ and $\dot{q}' = \dot{q} - \tau \cdot \dot{p}' - \dot{\tau} \cdot p'$, where $\tau$ is the only real root of the polynomial $\Xi_n(t) = t^{2n-1} + \sum_{j=0}^{2n-2} \xi_j t^j$ given by the coefficients

$$\xi_0 = 2\frac{\langle p', r\rangle}{\mu_n v'^{2n}}, \qquad \xi_1 = \frac{(\mu_1 + (2n-1)n\mu_n y^{2n-2})v'^2 - u'^2}{\mu_n v'^{2n}},$$

and $\xi_j = \binom{2n}{j+1}(-y/v')^{2n-j-1}$ for $j = 2, \ldots, 2n-2$, whereas the quantity

$$\dot{\tau} = \frac{\langle \dot{q} - \tau \cdot \dot{p}', r' \rangle}{\langle p', r' \rangle}$$

is determined by imposing that $\dot{q}' \in T_{q'}C$: If $r'$ is any normal vector to $C$ at $q'$, then $\dot{q}' \in T_{q'}C \Leftrightarrow \langle \dot{q}', r' \rangle = 0$.

The more expensive part is to solve the polynomial equation $\Xi_n(\tau) = 0$, which is obtained by imposing that the new impact point $q' = q - \tau p'$ verifies (B.1). The root $\tau$ is simple and negative. It is computed by Newton's method taking as initial approximation its value for $\epsilon = 0$, namely

$$\tau_0 = 2\frac{xu'/a^2 + yv'/b^2}{(u'/a)^2 + (v'/b)^2} = 2\frac{xu' - \mu_1 yv'}{u'^2 - \mu_1 v'^2}.$$

### *B.3 Taylor expansions of the invariant curves*

The dynamics on the unstable invariant curve can be linearized. There exists some analytic maps $q = (x, y) : \mathbb{R} \to C$ and $p = (u, v) : \mathbb{R} \to \mathbb{S}$ such that

$$q(0) = (a, 0), \qquad p(0) = (1, 0), \qquad f(q(r), p(r)) = -(q(\lambda r), p(\lambda r)),$$

where $\lambda = \mathrm{e}^h$ is the characteristic multiplier of the hyperbolic periodic orbit.

Due to the axial symmetries of the monomial perturbations (1), the functions $x(r)$ and $u(r)$ are even, whereas $y(r)$ and $v(r)$ are odd. Our goal is to develop a recursive algorithm to compute the Taylor expansions $x(r) = \sum_{k\geq 0} x_k r^{2k}$, $y(r) = \sum_{k\geq 1} y_k r^{2k-1}$, $u(r) = \sum_{k\geq 0} u_k r^{2k}$, and $v(r) = \sum_{k\geq 1} v_k r^{2k-1}$ up to any order. The key idea is to realize that these four expansions can be determined by using the following four functional equations:

(a) $u(r)^2 + v(r)^2 = 1$,
(b) $x(r)^2 = \mu_0 + \mu_1 y(r)^2 + \mu_n y(r)^{2n}$,
(c) $\langle p(\lambda r) + p(r), \dot{q}(r) \rangle = 0$, and
(d) $\det[q(\lambda r) + q(r), p(\lambda r)] = 0$.

Equation (a) means that the velocities $p = (u, v)$ have unit norm: $p \in \mathbb{S}$. Equation (b) follows from the fact that the points $q = (x, y)$ are on the curve defined by (B.1). Equation (c) holds because the difference of consecutive velocities is normal to the curve at the old impact point. Finally, we have stated in equation (d) that the difference of consecutive impact points is parallel to the new velocity. Only equation (b) depends on the form of the curve. If $\star \in \{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}\}$ and $l \in \mathbb{Z}_+$, we denote by $(\star)_l$ the equation obtained by equating the $\mathrm{O}(r^l)$-terms in both sides of the functional equation $(\star)$. For instance, $(\mathrm{a})_{2k}$ reads as $\sum_{s=0}^{k} u_s u_{k-s} + \sum_{s=1}^{k} v_s v_{k+1-s} = 0$.

We know that $x_0 = a$ and $u_0 = 1$. Then equations $(\mathrm{b})_2$, $(\mathrm{c})_1$, and $(\mathrm{d})_1$ give rise to a system whose one-parametric family of non-trivial solutions is: $y_1 \neq 0$, $v_1 = \lambda^{-1/2} y_1/b$, and $x_1 = -a y_1/b$. We take $y_1 = 2b$, $v_1 = 2\lambda^{-1/2}$ and $x_1 = -2a$. Besides, the coefficient $u_1 = -2/\lambda$ is found using equation $(\mathrm{a})_2$.

Now, let us suppose that $x_0, \ldots, x_k, y_1, \ldots, y_k, u_0, \ldots, u_k, v_1, \ldots, v_k$ have been computed for some integer $k \geq 1$. Then using the equations $(\mathrm{b})_{2k+2}$, $(\mathrm{c})_{2k+1}$, and $(\mathrm{d})_{2k+1}$, we get the linear system

$$\begin{pmatrix} 2a & 4a^2/b & 0 \\ 4(k+1) & (4k+2)(\lambda+1)\lambda^{-1/2} & 2(\lambda^{2k+1}+1)b \\ 0 & (\lambda^{2k+1}+1) & -2\lambda^{2k+1}a \end{pmatrix} \begin{pmatrix} x_{k+1} \\ y_{k+1} \\ v_{k+1} \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

whose independent term depends only on previously computed coefficients. The determinant of this system is $4ab(\lambda^{2k}-1)(1-\lambda^{2k+2}) \neq 0$, for any $k \geq 1$. Thus, the coefficients $x_{k+1}$, $y_{k+1}$ and $v_{k+1}$ can be computed. Next, we compute the coefficient $u_{k+1}$ from equation $(\mathrm{a})_{2k+2}$. Therefore, this algorithm can be applied recursively to obtain the Taylor expansions up to any order.

## References


[1] C. Batut, K. Belabas, D. Bernardi, H. Cohen and M. Olivier, *User's Guide to PARI/GP* (Freely available from `http://www.parigp-home.de/`).

[2] G.D. Birkhoff, *Dynamical Systems* (AMS, Providence, 1927).

[3] V.L. Chernov, On separatrix splitting of some quadratic area-preserving maps of the plane, *Regul. Chaotic Dyn.* **3** (1998) 49–65.

[4] A. Delshams and R. Ramírez-Ros, Poincaré-Melnikov-Arnold method for analytic planar maps, *Nonlinearity* **9** (1996) 1–26.

[5] A. Delshams and R. Ramírez-Ros, Exponentially small splitting of separatrices for perturbed integrable standard-like maps, *J. Nonlinear Sci.* **8** (1998) 317–352.

[6] A. Delshams and R. Ramírez-Ros, Singular separatrix splitting and the Melnikov method: An experimental study, *Exp. Math.* **8** (1999) 29–48.

[7] A. Delshams, Yu. Fedorov and R. Ramírez-Ros, Homoclinic billiard orbits inside symmetrically perturbed ellipsoids, *Nonlinearity* **14** (2001) 1141–1195.

[8] E. Fontich and C. Simó, The splitting of separatrices for analytic diffeomorphisms, *Ergodic Theory Dynam. Systems* **10** (1990) 295–318.

[9] V.G. Gelfreich, V.F. Lazutkin and N.V. Svanidze, A refined formula for the separatrix splitting for the standard map, *Phys. D* **71** (1994) 82–101.

[10] V.G. Gelfreich, A proof of the exponentially small transversality of the separatrices for the standard map, *Comm. Math. Phys.* **201** (1999) 155–216.

[11] V.G. Gelfreich and D. Sauzin, Borel summation and splitting of separatrices for the Hénon map, *Ann. Inst. Fourier (Grenoble)* **51** (2001) 513–567.

[12] V.G. Gelfreich and V.F. Lazutkin, Splitting of separatrices: perturbation theory and exponential smallness, *Russian Math. Surveys* **56** (2001) 499–558.

[13] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems* (Cambridge Univ. Press, Cambridge, 1995).

[14] V.V. Kozlov and D.V. Treshchëv, *Billiards: A Genetic Introduction to the Dynamics of Systems with Impacts*, Trans. Math. Monographs **89** (AMS, Providence, 1991).

[15] V.F. Lazutkin, Splitting of separatrices for the standard map. Preprint VINITI (1984).

[16] V.F. Lazutkin, The splitting of separatrices for a standard family of area-preserving transformations, *Transl. Amer. Math. Soc. Ser. 2* **157** (1993) 27–35.

[17] R.S. MacKay, J.D. Meiss and I.C. Percival, Transport in Hamiltonian systems, *Phys. D* **13** (1984) 55-81.

[18] E.M. McMillan, A problem in the stability of periodic systems, in: E. Brittin and H. Odabasi, eds., *Topics in Modern Physics* (Colorado Assoc. Univ. Press, Boulder, 1971) 219–244.

[19] C. Simó, Analytical and numerical computation of invariant manifolds, in: D. Benest and C. Froeschlé, eds., *Modern Methods in Celestial Mechanics* (Editions Frontières, Gif-sur-Yvette, 1990) 285–330.

[20] C. Simó, Analytical and numerical detection of exponentially small phenomena, in: B. Fiedler, K. Gröger and J. Sprekels, eds., *Proceedings of the International Conference on Differential Equations* (World Sci. Publishing, River Edge, 2000) 967–976.

[21] C. Simó. Private communication.

[22] S. Tabachnikov, *Billiards*, Panor. Synth **1** (SMF, Paris, 1995).

[23] M.B. Tabanov, Separatrices splitting for Birkhoff's billiard in a symmetric convex domain, close to an ellipse, *Chaos* **4** (1994) 595–606.

[24] E.T. Whittaker and G.N. Watson, *A Course of Modern Analysis* (Cambridge Univ. Press, Cambridge, 1927).