Melnikov potential for exact symplectic maps

Amadeu Delshams and Rafael Ramírez-Ros
Departament de Matemàtica Aplicada I
Universitat Politècnica de Catalunya
Diagonal 647, 08028 Barcelona, Spain

Revised version, March 1997

Abstract

The splitting of separatrices of hyperbolic fixed points for exact symplectic maps of n degrees of freedom is considered. The non-degenerate critical points of a real-valued function (called the Melnikov potential) are associated to transverse homoclinic orbits and an asymptotic expression for the symplectic area between homoclinic orbits is given. Moreover, if the unperturbed invariant manifolds are completely doubled, it is shown that there exist, in general, at least 4 primary homoclinic orbits ($4n$ in antisymmetric maps). Both lower bounds are optimal.

Two examples are presented: a $2n$-dimensional central standard-like map and the Hamiltonian map associated to a magnetized spherical pendulum. Several topics are studied about these examples: existence of splitting, explicit computations of Melnikov potentials, transverse homoclinic orbits, exponentially small splitting, etc.

AMS Subject Classification (1991): 33E05, 34C37, 34E10, 57R19, 57R70, 58F05

1 Introduction

In a previous work [DR96], the authors were able to develop a general theory for perturbations of an integrable planar map with a separatrix to a hyperbolic fixed point. The splitting of the perturbed invariant curves was measured, in first order with respect to the parameter of perturbation, by means of a periodic *Melnikov function* M defined on the unperturbed separatrix. In case of area preserving perturbations, M has zero mean and therefore there exists a periodic function L (called the *Melnikov potential*) such that $M = L'$. Consequently, if L is not identically constant (respectively, has non-degenerate critical points), the separatrix splits (respectively, the perturbed curves cross transversely). Moreover, under some hypothesis of meromorphicity, the Melnikov potential is elliptic and there exists a *Summation Formula* (see Appendix) to compute it explicitly.

The aim of this paper is to develop a similar theory for more dimensions. The natural frame is to consider exact symplectic perturbations of a $2n$-dimensional exact map with a n-dimensional separatrix associated to a hyperbolic fixed point.

Exact symplectic maps $F : \mathcal{P} \to \mathcal{P}$ are defined on exact manifolds, i.e., $2n$-dimensional manifolds $\mathcal{P}$ endowed with a symplectic form ω which is exact: $\omega = -\,\mathrm{d}\phi$; and they are characterized by the equation $F^*\phi - \phi = \mathrm{d}S$ for some function $S : \mathcal{P} \to \mathbf{R}$, called *generating function* of F.

The typical example of an exact symplectic manifold is provided by a cotangent bundle $T^*\mathcal{M}$, together with the canonical forms ϕ_0, ω_0, which in cotangent coordinates (x, y) read as $\phi_0 = y\,\mathrm{d}x$, $\omega_0 = \mathrm{d}x \wedge \mathrm{d}y$. Typical exact symplectic maps are the so-called *twist* maps, which satisfy $F^*(y\,\mathrm{d}x) - y\,\mathrm{d}x = Y\,\mathrm{d}X - y\,\mathrm{d}x = \mathrm{d}\mathcal{L}(x, X)$, where $(X, Y) = F(x, y)$. The fact that the generating function S can be written in terms of old and new coordinates: $S(x, y) = \mathcal{L}(x, X)$, is the *twist condition* that gives the name to these maps. The function $\mathcal{L}$ is called twist generating function. As in [Eas91], we will not restrict ourselves to this typical case, since the results to be presented in this paper are valid on arbitrary exact symplectic manifolds and the twist condition is not needed.

The exact symplectic structure plays a fundamental role in our construction, since it allows us to work neatly with geometric objects. For example, it is used to introduce two homoclinic invariants: the action of a homoclinic orbit and the symplectic area between two homoclinic orbits, called simply homoclinic area.

Namely, let $p_\infty \in \mathcal{P}$ be a hyperbolic fixed point of F, which lies in the intersection of the n-dimensional invariant manifolds $\mathcal{W}^{\mathrm{u,s}}$. Given a *homoclinic orbit* $\mathcal{O} = (p_k)_{k\in\mathbf{Z}}$ of F, i.e., $\mathcal{O} \subset (\mathcal{W}^{\mathrm{u}} \cap \mathcal{W}^{\mathrm{s}}) \setminus \{p_\infty\}$ and $F(p_k) = p_{k+1}$, we define the *homoclinic action* of the orbit $\mathcal{O}$ as

$$W[\mathcal{O}] := \sum_{k\in\mathbf{Z}} S(p_k),$$

where, in order to get an absolutely convergent series, the generating function S has been determined by imposing $S(p_\infty) = 0$. Given another homoclinic orbit $\mathcal{O}'$ of F, the *homoclinic area* between the two homoclinic orbits $\mathcal{O}$, $\mathcal{O}'$ is defined as the difference of homoclinic actions $\Delta W[\mathcal{O}, \mathcal{O}'] := W[\mathcal{O}] - W[\mathcal{O}']$. These two objects are *symplectic invariants*, i.e., they neither depend on the symplectic coordinates used, nor on the choice of the one-form ϕ. It is worth noting that in the planar case, the homoclinic area is the standard (algebraic) area of the lobes between the invariant curves [MMP84,Mat86,Eas91] and also measures the flux along the homoclinic tangle, which is related to the study of transport [MMP84,RW88,Mei92].

The unperturbed role will be played by an exact symplectic diffeomorphism $F_0 : \mathcal{P} \to \mathcal{P}$, defined on a $2n$-dimensional exact manifold $\mathcal{P}$, which possesses a hyperbolic fixed point p_∞ and a n-dimensional *separatrix* $\Lambda \subset \mathcal{W}^{\mathrm{u}}_0 \cap \mathcal{W}^{\mathrm{s}}_0$, where $\mathcal{W}^{\mathrm{u,s}}_0$ denote the invariant manifolds associated to p_∞.

Consider now a family of exact symplectic diffeomorphisms $\{F_\varepsilon\}$, as a general perturbation of the situation above, and let $S_\varepsilon = S_0+\varepsilon S_1+\mathrm{O}(\varepsilon^2)$ be the generating function of F_ε.

The main analytical results of this paper are stated and proved in section 2. There, the Melnikov potential is introduced as the real-valued smooth function $L:\Lambda\to\mathbf{R}$ given by

$$L(p) := \sum_{k\in\mathbf{Z}} \widehat{S}_1(p_k), \qquad p_k = F_0^k(p),$$

where $\widehat{S}_1:\mathcal{P}\to\mathbf{R}$ is defined as $\widehat{S}_1(p) = S_1(p) - \phi(F_0(p))[F_1(p)]$, and F_1 is the first order variation in ε of the family $\{F_\varepsilon\}$, that is, $F_1(p) = [\partial F_\varepsilon(p)/\partial\varepsilon]|_{\varepsilon=0}$. Obviously, S_1 is determined by imposing $\widehat{S}_1(p_\infty)=0$, in order to get an absolutely convergent series.

In theorem 2.1 it is established that

(i) the Melnikov potential L is F_0-invariant: $L\circ F_0 = L$,

(ii) if $L\not\equiv$ constant, the perturbed invariant manifolds $\mathcal{W}_\varepsilon^{\mathrm{u,s}}$ split for $0<|\varepsilon|\ll 1$,

(iii) the non-degenerate critical points of L are associated to transverse intersections of the perturbed invariant manifolds,

(iv) the above-mentioned homoclinic invariants are given in first order by L.

As a matter of fact, the perturbed homoclinic orbits detected by the Melnikov potential are all of them *primary homoclinic orbits* $\mathcal{O}_\varepsilon$ of F_ε, i.e., they are smooth in ε for $|\varepsilon|$ small enough.

The Melnikov potential admits several reformulations. For example, if F_ε is a twist map on a cotangent bundle $T^*\mathcal{M}$, with twist generating function $\mathcal{L}_\varepsilon = \mathcal{L}_0+\varepsilon\mathcal{L}_1+\mathrm{O}(\varepsilon^2)$, $\widehat{S}_1$ has the simple form $\widehat{S}_1(p) = \mathcal{L}_1(\pi(p),\pi(F_0(p)))$, where $\pi:T^*\mathcal{M}\to\mathcal{M}$ is the natural projection. Consequently, the Melnikov potential reads as [DRS97]

$$L(p) = \sum_{k\in\mathbf{Z}} \mathcal{L}_1(x_k,x_{k+1}), \qquad x_k = \pi(p_k),$$

where $\mathcal{L}_1$ is determined by imposing $\mathcal{L}_1(x_\infty,x_\infty)=0$, and $x_\infty=\pi(p_\infty)$. Another interesting situation, that allows us to compare the continuous and discrete frames, is to consider *Hamiltonian maps*. Let $H_\varepsilon:\mathcal{P}\times\mathbf{R}\to\mathbf{R}$ be a time-periodic Hamiltonian of period T, and $F_\varepsilon=\Psi_\varepsilon^T$, where $\Psi_\varepsilon^t(p)$ is the solution of the associated Hamiltonian equations with initial condition p at $t=0$. If $H_\varepsilon = H_0+\varepsilon H_1+\mathrm{O}(\varepsilon^2)$, then $\widehat{S}_1(p) = -\int_0^T H_1(\Psi_0^t(p),t)\,\mathrm{d}t$, so the Melnikov potential takes the form (already known to Poincaré)

$$L(p) = -\int_{\mathbf{R}} H_1(\Psi_0^t(p),t)\,\mathrm{d}t,$$

where H_1 is determined by imposing $H_1(\Psi_0^t(p_\infty),t)\equiv 0$, or simply $H_1(p_\infty,t)\equiv 0$, if H_0 is autonomous.

An essential ingredient for the proof of theorem 2.1 is the fact that the invariant manifolds $\mathcal{W}_\varepsilon^{\mathrm{u,s}}$ are exact Lagrangian immersed submanifolds of $\mathcal{P}$ and therefore can be expressed in terms of generating functions $L_\varepsilon^{\mathrm{u,s}}$. The Lagrangian property of the invariant manifolds was already noticed by Poincaré [Poi99] for flows, although we learned it for maps from E. Tabacman [Tab95], as well as the expression for the invariant manifolds given in proposition 2.1, in the twist frame. The relationship between $L_1^{\mathrm{u,s}}$ and S_1, the first order variations in ε of the generating functions $L_\varepsilon^{\mathrm{u,s}}$ and S_ε, gives then the formula for the Melnikov potential. The tools utilized are very similar to those of D. Treschev [Tre94]. However, D. Treschev considers autonomous Hamiltonian flows, and the conservation of energy makes easier the deduction of the continuous version of equation (2.5).

In that frame (Hamiltonian-Lagrangian flows), it is worth noting that a variational approach to the Melnikov method was carried out by S. Angenent [Ang93] for Hamiltonian systems with $1\frac{1}{2}$ degrees of freedom, and that a mechanism for finding homoclinic orbits in positively definite symplectic diffeomorphisms is due to S. Bolotin [Bol94], based on interpolating them by Hamiltonian flows.

Section 2 contains also some remarks on the non-symplectic case: a vector-valued Melnikov function M is then defined, whose non-degenerated zeros are associated to transverse homoclinic orbits.

The last part of section 2 is devoted to gain information on the number of primary homoclinic orbits after perturbation. Since the Melnikov potential L is F_0-invariant, it can be defined on the *reduced separatrix* $\Lambda^* := \Lambda/F_0$, which is the quotient of the separatrix by the unperturbed map. The reduced separatrix is a compact manifold without boundary, provided that the unperturbed invariant manifolds are *completely doubled*, i.e., $\mathcal{W}_0^{\mathrm{u}} = \mathcal{W}_0^{\mathrm{s}}$ and $\mathcal{W}_0^{\mathrm{u,s}} \setminus \{p_\infty\}$ is a submanifold of $\mathcal{P}$ and not only an immersed submanifold of $\mathcal{P}$. This is equivalent to require that the separatrix is $\Lambda = \mathcal{W}_0^{\mathrm{u,s}} \setminus \{p_\infty\}$. Several dynamical consequences of this fact can be pointed out using topological tools. In particular, Morse theory gives lower bounds on the number of primary transverse homoclinic orbits, under conditions of generic position: in theorem 2.2 it is stated that the number of primary homoclinic orbits is at least 4.

Moreover, if the maps F_ε have a common symmetry $I : \mathcal{P} \to \mathcal{P}$ ($F_\varepsilon \circ I = I \circ F_\varepsilon$, and $F_\varepsilon(p_\infty) = I(p_\infty) = p_\infty$) such that the one-form ϕ is preserved by I: $I^*\phi = \phi$, then the Melnikov potential is I-invariant (see lemma 2.6). Consequently, it can be considered as a function over the quotient manifold $\Lambda_I^* := \Lambda/\{F_0, I\}$. If, in addition, I is an involution ($I^2 = \mathrm{Id}$) such that $DI(p_\infty) = -\mathrm{Id}$, the family $\{F_\varepsilon\}$ will be called *antisymmetric*. In this case, in theorem 2.2 it is stated that the number of primary homoclinic orbits is at least $4n$ and that they appear coupled in (anti)symmetric pairs: $\mathcal{O}_\varepsilon$ is a primary homoclinic orbit if and only if $I(\mathcal{O}_\varepsilon)$ also is.

It is worth mentioning that any family of odd maps $F_\varepsilon : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ (with the standard symplectic structure) is antisymmetric.

To prove theorem 2.2, it is enough to check that the sum of the $\mathbf{Z}_2$-Betti numbers of Λ^* and Λ_I^* are 4 and $2n$, respectively. This is accomplished by computing the $\mathbf{Z}_2$-homology of Λ^* and Λ_I^*.

Both lower bounds are optimal, as it is shown in several perturbations of maps with a central symmetry, so that the unperturbed invariant manifolds are completely doubled.

It is important to notice that the invariant manifolds of a product of uncoupled planar maps with double loops are not completely doubled, see remark 2.3, and hence, the topological results do not hold in this case. Indeed, the number of primary homoclinic orbits may be rather different under perturbation; for instance, it is possible to construct explicitly perturbations with an infinite number of primary homoclinic orbits, all of them being transverse. The study of this kind of phenomena is currently being researched.

In section 3, as a first example, we consider the family of twist maps on $\mathbf{R}^{2n}$:

$$F_\varepsilon(x,y) = \left(y, -x + \frac{2\mu y}{1+|y|^2} + \varepsilon \nabla V(y) \right), \quad \mu > 1, \, \varepsilon \in \mathbf{R},$$

with $V : \mathbf{R}^n \to \mathbf{R}$ determined by imposing $V(0) = 0$. The map above is a perturbation of the McLachlan map [McL94], which is a multi-dimensional generalization of the McMillan map [McM71], which in its turn is a particular case of the standard-like Suris integrable maps [Sur89]. The McLachlan map has a central symmetry that makes the dynamics over the separatrix essentially one-dimensional. This is the key fact that allow us to perform a complete analysis, since the *natural parametrizations* (3.2) can be introduced.

If the potential V is entire and not identically zero, in theorem 3.1 it is proved that the manifolds $\mathcal{W}_\varepsilon^{\mathrm{u,s}}$ of the map F_ε split, for $0 < |\varepsilon| \ll 1$. This result is obtained simply by checking that the Melnikov potential is not constant. Moreover, if V is a polynomial, the Melnikov potential can be computed explicitly.

In particular, if V is a quadratic form: $V(y) = y^\top B y$ for some symmetric $n \times n$ matrix B, in proposition 3.1 it is stated that under generic conditions on B ($\det(B) \neq 0$ and B does not have multiple eigenvalues), the perturbed invariant manifolds are transverse along exactly $4n$ primary homoclinic orbits.

If V is linear: $V(y) = b^\top y$ for some vector $b \in \mathbf{R}^n \setminus \{0\}$, in proposition 3.2 it is stated that the perturbed invariant manifolds are transverse along exactly 4 primary homoclinic orbits.

The difference between both kinds of perturbations is that quadratic potentials V give rise to odd maps, whereas linear ones do not. Moreover, propositions 3.1 and 3.2 give the unperturbed homoclinic orbits that survive and the first order (in ε) of the homoclinic areas between the different primary homoclinic orbits.

The weakly hyperbolic case $0 < h \ll 1$, $\cosh(h) := \mu$, is also studied for the case of a quadratic potential V, and asymptotic expressions for the homoclinic areas are given at the end of section 3. It turns out that, for some distinguished pairs, interlaced in the same way as in the case of 1 degree of freedom, the homoclinic area predicted by the Melnikov potential is exponentially small with respect to the hyperbolicity parameter h. Of course, this does not prove that the splitting size is exponentially small in singular cases, i.e., when ε and h tend simultaneously to zero.

The last section is devoted to the study of the Hamiltonian maps arising from time-periodic perturbations of an (undamped) magnetized spherical pendulum. This model was introduced by J. Gruendler [Gru85] as a first example of application of the Melnikov method for high-dimensional (continuous) systems. The Hamiltonians considered have the form [Gru85]

$$H_\varepsilon : \mathbf{R}^{2n} \times \mathbf{R} \to \mathbf{R}, \quad H_\varepsilon(x,y,t) = v^2/2 + (r^4 - r^2)/2 + \varepsilon V(x, t/h), \quad h > 0, \varepsilon \in \mathbf{R},$$

where $v = |y|$, $r = |x|$, and $V = V(x,\varphi)$ is 1-periodic in φ. We determine V by imposing $V(0,\varphi) \equiv 0$. Note that small values of h correspond to a quick forcing.

General perturbations, and not only symplectic ones, are considered in [Gru85]. As a consequence, the homoclinic orbits are given in the general case by non-degenerate zeros of a vector-valued Melnikov function, instead of non-degenerate critical points of the real-valued Melnikov potential. We have computed the Melnikov potential for the Hamiltonian perturbations studied in [Gru85], and have verified that his Melnikov function is the gradient of our Melnikov potential.

Most of the results stated above for the McLachlan map also hold for this Hamiltonian map. There is, however, a significant difference. One cannot deduce a priori that the Melnikov potential is not identically constant without computing it. This has to do with the fact that the Melnikov potential is simply periodic and regular for the polynomial perturbations considered, in contrast with the complex period and singularities that the Melnikov potential has for the entire perturbations of the McLachlan map.

To finish the account of results, let us point out that a similar Melnikov analysis for perturbed ellipsoidal billiards has not been included for the sake of brevity and will appear elsewhere. Such billiards are a high-dimensional version of perturbed elliptic billiard *tables*, which have already been studied in several papers [LT93,Tab94,DR96, Lom96a].

After this research was complete, we became aware of some recent papers [Lom93, Lom96b] of H. Lomelí for twist maps on the annulus $\mathbf{A}^n = T^*\mathbf{T}^n = \mathbf{T}^n \times \mathbf{R}^n$ that resemble our method. However, they do not contain explicit computations (i.e., in terms of known functions) of the Melnikov potential, since complex variable methods are not used. Besides, in those papers it is assumed that the separatrix is globally horizontal, a condition that does not hold for homoclinics in $\mathbf{R}^{2n}$, since the separatrix must fold to go back to the fixed point.

Another related papers are [Sun96,BGK95], but their approach is rather different, since they deal, like [Gru85], with the general case, with no symplectic structure, and therefore a vector-valued Melnikov function is needed. This makes an important difference not only from a computational point of view (there are not explicit (analytic) computations in these works), but also from a theoretical point of view, since *Morse theory* cannot be applied in the general situation. We also want to mention the work [BF96], where perturbations of n-dimensional maps having homo-heteroclinic connections to compact normally hyperbolic invariant manifolds are considered.

2 Main results

For the sake of simplicity, we will assume that the objects here considered are smooth. For a general background on symplectic geometry we refer to [Arn76,GS77,AM78]. The basic properties of immersed submanifolds can be found in [pages 6–11]

2.1 Exact objects

A $2n$-dimensional manifold $\mathcal{P}$ together with an exact non-degenerate two-form ω over it, is called an *exact symplectic manifold.* Then, $\omega = -\,\mathrm{d}\phi$ for some one-form ϕ, usually called *Liouville form*, *symplectic potential* or *action form*.

A map $F : \mathcal{P} \to \mathcal{P}$ is called *exact symplectic* (or simply, exact) if $\oint_\gamma \phi = \oint_{F\gamma} \phi$ for all closed path $\gamma \subset \mathcal{P}$ or, equivalently, if $F^*\phi - \phi = \mathrm{d}S$ for some function $S : \mathcal{P} \to \mathbf{R}$, called *generating function* of F.

A n-dimensional submanifold $\Lambda \subset \mathcal{P}$ is called an *exact Lagrangian submanifold* (or simply, an exact submanifold) if $\oint_\gamma \phi = 0$ for all closed path $\gamma \subset \Lambda$ or, equivalently, if $\imath_\Lambda^*\phi = \mathrm{d}L$ for some function $L : \Lambda \to \mathbf{R}$, called *generating function* of Λ. Here $\imath_\Lambda : \Lambda \hookrightarrow \mathcal{P}$ stands for the inclusion map.

Unfortunately, the invariant manifolds that we will deal with are not submanifolds, but just *immersed* submanifolds. Thus, the introduction of some technicalities seems unavoidable in order to give a rigorous exposition of the subject, and more precisely, to introduce the notion of separatrix, where the distance between the perturbed invariant manifolds will be measured.

Given a manifold $\mathcal{N}$, we recall that a map $g : \mathcal{N} \to \mathcal{P}$ is called an *immersion* when its differential $\mathrm{d}g(z)$ has maximal rank at any point $z \in \mathcal{N}$. If g is one-to-one onto its image $\mathcal{W} = g(\mathcal{N})$, there is a natural way to make $\mathcal{W}$ a smooth manifold: the topology on $\mathcal{W}$ is the one which makes g a homeomorphism and the charts on $\mathcal{W}$ are the pull-backs via g^{-1} of the charts on $\mathcal{N}$. The manifold $\mathcal{W}$ constructed in this way is called an *immersed submanifold* of $\mathcal{P}$ and its dimension is equal to the dimension of $\mathcal{N}$. It is important to notice that the topology of the immersed manifold need not be the same as the induced one via the inclusion $\mathcal{W} \subset \mathcal{P}$ or, in other words, that $\mathcal{W}$ need not be a submanifold of $\mathcal{P}$ in the usual sense.

Figure 1 shows an example of a double loop $\mathcal{W} = g(\mathbf{R})$ to $p_\infty = \lim_{z\to\pm\infty} g(z)$ for an immersion $g : \mathbf{R} \to \mathbf{R}^2$. At p_∞, the induced topology on $\mathcal{W}$ via the inclusion $\mathcal{W} \subset \mathbf{R}^2$ is not the same as the induced one via g. Both $g(B)$, for all open bounded interval $B \subset \mathbf{R}$, and $\mathcal{W} \setminus \{p_\infty\}$ are submanifolds, but not $\mathcal{W}$. This situation is a particular case of the following elementary result [page 11]

Lemma 2.1 *Let $g : \mathcal{N} \to \mathcal{P}$ be a one-to-one immersion and set $\mathcal{W} = g(\mathcal{N})$.*

(i) Let B be an open subset of $\mathcal{N}$ with compact closure. Then, $g_{|B} : B \to \mathcal{P}$ is an embedding, that is, a homeomorphism onto its image $g(B)$. Thus, $g(B)$ is a submanifold of $\mathcal{P}$, which will be called an embedded disk *in $\mathcal{W}$.*

(ii) Let $\Sigma \subset \mathcal{W}$ be the set of points where the two topologies on $\mathcal{W}$ (the one induced by the inclusion $\mathcal{W} \subset \mathcal{P}$ and the one that makes g a homeomorphism) differ. Then, $\Lambda = \mathcal{W} \setminus \Sigma$ is a submanifold of $\mathcal{P}$. Indeed, $\mathcal{W}$ is not a submanifold of $\mathcal{P}$ just at the points of Σ.

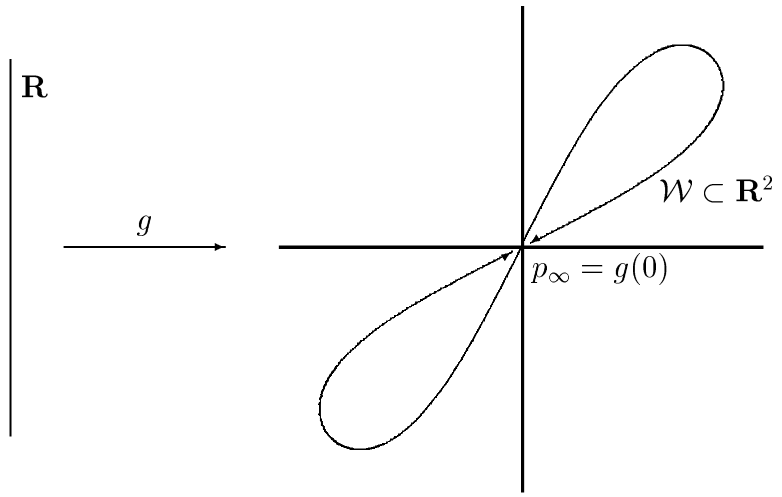

Figure 1: $g = (g_1, g_2) : \mathbf{R} \to \mathbf{R}^2$, where $g_1(z) = \frac{3}{2} z/(1+z^2)$, $g_2(z) = g_1(2z)$.

For the sake of clearness, submanifolds and immersed submanifolds will be denoted by different letters, namely Λ and $\mathcal{W}$, respectively. For immersed submanifolds $\mathcal{W}$, the map $\imath_{\mathcal{W}} : \mathcal{W} \to \mathcal{P}$ stands for the inclusion map, as before. It should be noted that $\imath_{\mathcal{W}}$ is smooth, even when $\mathcal{W}$ is not a submanifold of $\mathcal{P}$, because of the differential structure given to $\mathcal{W}$. Moreover, if $\gamma \subset \mathcal{P}$ is a (closed) path, we will say that γ is a (closed) path *in* the immersed submanifold $\mathcal{W}$ if and only if γ is contained in $\mathcal{W}$ and it is *continuous in the topology of* $\mathcal{W}$. For example, if γ is one loop of figure 1, it is a closed path in $\mathbf{R}^2$ but not in $\mathcal{W}$.

With these notations and definitions, we are naturally led to define exact immersed submanifolds in the same way as exact submanifolds. A n-dimensional immersed submanifold $\mathcal{W} \subset \mathcal{P}$ is called *exact* if $\oint_\gamma \phi = 0$ for all closed path γ in $\mathcal{W}$ or, equivalently, if $\imath_{\mathcal{W}}^* \phi = \mathrm{d}L$ for some function $L : \mathcal{W} \to \mathbf{R}$, called *generating function* of $\mathcal{W}$.

The symplectic potential ϕ is determined except for the addition of a closed zero-form, and the generating functions of maps or (immersed) submanifolds are determined except for an additive constant. Henceforth, the symbol ${}_{\mathcal{W}}\!\int_p^q \phi$, with $p, q \in \mathcal{W}$, will denote the integral of ϕ along an arbitrary path from p to q in $\mathcal{W}$. It only makes sense for an exact immersed submanifold $\mathcal{W}$, since then the integral does not depend on the path. The difference of values of L can be expressed as an integral of this kind:

$$L(q) - L(p) = \int_p^q \mathrm{d}L = {}_{\mathcal{W}}\!\int_p^q \phi, \qquad \forall p, q \in \mathcal{W}. \tag{2.1}$$

Lemma 2.2 *Let $\mathcal{W}$ be a connected exact immersed submanifold of $\mathcal{P}$, invariant under an exact map F. Let L and S be their respective generating functions. Then,*

$$S(p) + \text{constant} = L(F(p)) - L(p), \qquad \forall p \in \mathcal{W}. \tag{2.2}$$

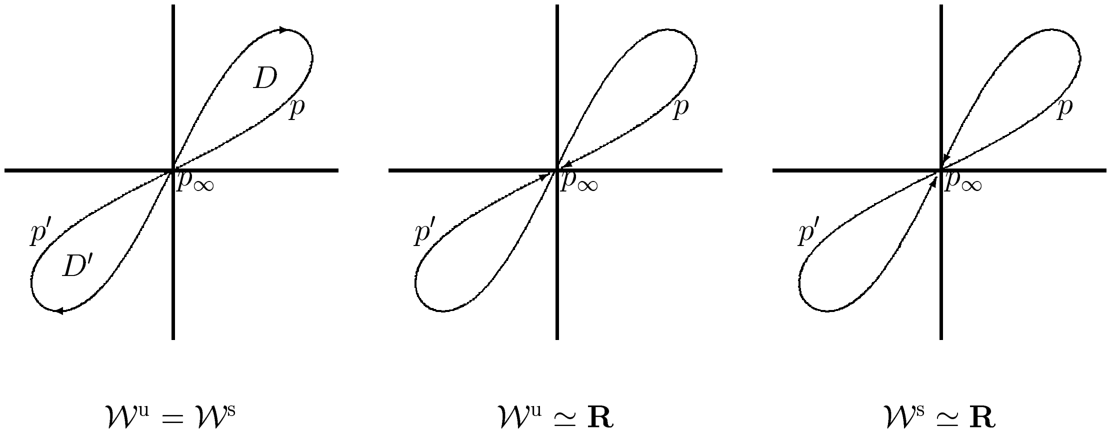

Figure 2: The invariant manifolds $\mathcal{W}^{\mathrm{u}}$ and $\mathcal{W}^{\mathrm{s}}$ are different as smooth manifolds, and are not submanifolds of $\mathbf{R}^2$. There exist no paths $\gamma^{\mathrm{u,s}}$ in $\mathcal{W}^{\mathrm{u,s}}$ from p to p' such that $\gamma^{\mathrm{u}} = \gamma^{\mathrm{s}}$.

Moreover, if $p_\infty \in \mathcal{W}$ is a fixed point of F, the constant is $-S(p_\infty)$.

Proof. From $\mathrm{d}S = F^*\phi - \phi$ and $\mathrm{d}L = \imath_{\mathcal{W}}^*\phi$ we get

$$\mathrm{d}\left(S_{|\mathcal{W}}\right) = \imath_{\mathcal{W}}^*\,\mathrm{d}S = \left(F_{|\mathcal{W}}\right)^*\,\mathrm{d}L - \,\mathrm{d}L = \,\mathrm{d}\left(L \circ F_{|\mathcal{W}} - L\right),$$

where $S_{|\mathcal{W}} = S \circ \imath_{\mathcal{W}}$ and $F_{|\mathcal{W}} = \left(\imath_{\mathcal{W}}\right)^{-1} \circ F \circ \imath_{\mathcal{W}}$ are the restrictions of S and F to $\mathcal{W}$. Thus, $S - L \circ F + L$ is constant over $\mathcal{W}$ by connectedness and (2.2) is proved. To end the proof we only need to evaluate equation (2.2) at $p = p_\infty$. □

Let $p_\infty \in \mathcal{P}$ be a hyperbolic fixed point of F. The point p_∞ lies in the intersection of the n-dimensional *unstable* and *stable* invariant manifolds of the map F associated to p_∞:

$$\mathcal{W}^{\mathrm{u}} := \left\{p \in \mathcal{P} : \lim_{k \to -\infty} F^k(p) = p_\infty\right\}, \qquad \mathcal{W}^{\mathrm{s}} := \left\{p \in \mathcal{P} : \lim_{k \to +\infty} F^k(p) = p_\infty\right\}.$$

The manifolds $\mathcal{W}^{\mathrm{u,s}}$ need not be submanifolds of $\mathcal{P}$, but just connected immersed submanifolds, see figure 2. In fact, $\mathcal{W}^{\mathrm{u,s}} = g^{\mathrm{u,s}}(\mathbf{R}^n)$ for some one-to-one immersions $g^{\mathrm{u,s}} : \mathbf{R}^n \to \mathcal{P}$, such that $g^{\mathrm{u,s}}(0) = p_\infty$ and $dg^{\mathrm{u,s}}(0)[\mathbf{R}^n]$ is the tangent space to $\mathcal{W}^{\mathrm{u,s}}$ at p_∞ [II §6] Since F is exact, they are exact immersed submanifolds: if γ is a closed path in $\mathcal{W}^{\mathrm{u}}$ ($\mathcal{W}^{\mathrm{s}}$), then $\oint_\gamma \phi = \oint_{F^k\gamma} \phi \longrightarrow \oint_{p_\infty} \phi = 0$, when $k \to -\infty$ ($k \to +\infty$). It should be noted that if $\gamma \subset \mathcal{P}$ is closed and contained in $\mathcal{W}^{\mathrm{u}}$ (resp. $\mathcal{W}^{\mathrm{s}}$), but it is *not* a path in $\mathcal{W}^{\mathrm{u}}$ (resp. $\mathcal{W}^{\mathrm{s}}$), the above argument fails. (For instance, if γ is one loop of figure 2.)

We denote by $L^{\mathrm{u,s}}$ the generating functions of $\mathcal{W}^{\mathrm{u,s}}$ and we determine the generating functions S, $L^{\mathrm{u,s}}$ by imposing $S(p_\infty) = L^{\mathrm{u,s}}(p_\infty) = 0$. The next proposition gives a nice interpretation of the generating functions of the stable and unstable invariant manifolds in terms of the generating function of the map.

Proposition 2.1 *Given $p^{\mathrm{u,s}} \in \mathcal{W}^{\mathrm{u,s}}$, let us denote $p_k^{\mathrm{u,s}} = F^k(p^{\mathrm{u,s}})$, for $k \in \mathbf{Z}$. Then,*

$$L^{\mathrm{u}}(p^{\mathrm{u}}) = \sum_{k<0} S(p_k^{\mathrm{u}}), \qquad L^{\mathrm{s}}(p^{\mathrm{s}}) = -\sum_{k\geq 0} S(p_k^{\mathrm{s}}).$$

Proof. From lemma 2.2, one has $S(p_k^{\mathrm{u,s}}) = L^{\mathrm{u,s}}(p_{k+1}^{\mathrm{u,s}}) - L^{\mathrm{u,s}}(p_k^{\mathrm{u,s}})$, for all $k \in \mathbf{Z}$. To get the formulae above, we simply consider the telescopic sums

$$\begin{aligned} L^{\mathrm{u}}(p^{\mathrm{u}}) &= \sum_{k<0} [L^{\mathrm{u}}(p_{k+1}^{\mathrm{u}}) - L^{\mathrm{u}}(p_k^{\mathrm{u}})] = \sum_{k<0} S(p_k^{\mathrm{u}}), \\ L^{\mathrm{s}}(p^{\mathrm{s}}) &= \sum_{k\geq 0} [L^{\mathrm{s}}(p_k^{\mathrm{s}}) - L^{\mathrm{s}}(p_{k+1}^{\mathrm{s}})] = -\sum_{k\geq 0} S(p_k^{\mathrm{s}}). \end{aligned}$$

These series are absolutely convergent, since $S(p_\infty) = 0$ and p_k^{u} (p_k^{s}) tends to p_∞ at an exponential rate as k tends to $-\infty$ ($+\infty$). □

Let $\mathcal{O} = (p_k)_{k\in\mathbf{Z}}$ be a *homoclinic orbit* of F, i.e., $\mathcal{O} \subset (\mathcal{W}^{\mathrm{u}} \cap \mathcal{W}^{\mathrm{s}}) \setminus \{p_\infty\}$ and $F(p_k) = p_{k+1}$. We define the *homoclinic action* of the orbit $\mathcal{O}$ as $W[\mathcal{O}] := L^{\mathrm{u}}(p_k) - L^{\mathrm{s}}(p_k)$. This definition does not depend on k, since a direct application of proposition 2.1 with $p_k^{\mathrm{u,s}} = p_k$ yields an equivalent k-independent definition

$$W[\mathcal{O}] := \sum_{k\in\mathbf{Z}} S(p_k). \tag{2.3}$$

Let $\mathcal{O}'$ be another homoclinic orbit of F. The *homoclinic area* between the two homoclinic orbits $\mathcal{O}$, $\mathcal{O}'$ is defined as the difference of homoclinic actions $\Delta W[\mathcal{O}, \mathcal{O}'] := W[\mathcal{O}] - W[\mathcal{O}']$. For a motivation of this name, consider $p \in \mathcal{O}$, $p' \in \mathcal{O}'$, $\gamma^{\mathrm{u,s}}$ a path from p to p' *in* $\mathcal{W}^{\mathrm{u,s}}$, $\gamma = \gamma^{\mathrm{u}} - \gamma^{\mathrm{s}}$, and suppose that D is an oriented 2-chain such that $\partial D = \gamma$. Then, by equation (2.1) and Stokes' formula, we have

$$\Delta W[\mathcal{O}, \mathcal{O}'] = \oint_\gamma \phi = -\iint_D \omega. \tag{2.4}$$

This formula shows clearly that the homoclinic area is a *symplectic invariant*, i.e., it neither depends on the symplectic coordinates used, nor on the choice of the symplectic potential ϕ. The homoclinic action can be considered as the homoclinic area between the homoclinic orbit at hand and the "orbit" of the fixed point p_∞. Thus, it is a symplectic invariant, too.

In particular, if $\mathcal{P} = \mathbf{R}^2$ with the standard area as the symplectic structure, and $p \in \mathcal{O}$, $p' \in \mathcal{O}'$ are consecutive intersections of the invariant manifolds, then the homoclinic area $\Delta W[\mathcal{O}, \mathcal{O}']$ is simply the (algebraic) area of the associated lobe.

Remark 2.1 Set $\mathcal{W} = \mathcal{W}^{\mathrm{u}} \cap \mathcal{W}^{\mathrm{s}}$ and let p, p' be two points of the same connected component of $\mathcal{W}$. When it is possible to choose the paths $\gamma^{\mathrm{u,s}}$ in $\mathcal{W}^{\mathrm{u,s}}$ from p to p' such that $\gamma = \gamma^{\mathrm{u}} - \gamma^{\mathrm{s}} = 0$, $\Delta W[\mathcal{O}, \mathcal{O}'] = \int_\gamma \phi = 0$, i.e., the actions coincide. They can be different if p and p' are not in the same component of $\mathcal{W}$. For instance, if p, p', D and D' are as in figure 2, $W[\mathcal{O}] = -\iint_D \omega > 0$ and $W[\mathcal{O}'] = \iint_{D'} \omega < 0$.

2.2 Families of exact objects

Now, we carry out the generalization of lemma 2.2 and proposition 2.1 for families of exact immersed submanifolds and maps, depending (in a smooth way) on a small parameter ε. First, let us recall the following standard fact from symplectic geometry [Wei73,GS77].

Lemma 2.3 *In any point p of any Lagrangian submanifold Λ of $\mathcal{P}$ there exists a neighbourhood $p \in U \subset \mathcal{P}$ and local coordinates (x, y) over U such that $\phi = y\,\mathrm{d}x$ (i.e., $\omega = \mathrm{d}x \wedge \mathrm{d}y$) and the set $\Lambda \cap U$ is given by the equation $y = 0$.*

We recall that a n-dimensional submanifold Λ is Lagrangian if $\imath_\Lambda^* \omega = 0$. In particular, exact submanifolds are Lagrangian. The coordinates above are called *cotangent coordinates* since they give a symplectic change of variables from the neighbourhood U onto a neighbourhood V of p in the cotangent space $T^*\Lambda$.

Let $g_\varepsilon : \mathcal{N} \to \mathcal{P}$ be one-to-one immersions and set $\mathcal{W}_\varepsilon = g_\varepsilon(\mathcal{N})$. We will say that the family of immersed submanifolds $\{\mathcal{W}_\varepsilon\}$ is *smooth* (at $\varepsilon = 0$) when for any embedded disk $\Lambda \subset \mathcal{W}_0$ there exists a smooth family of embedded disks $\{\Lambda_\varepsilon\}$ such that $\Lambda_\varepsilon \subset \mathcal{W}_\varepsilon$ and $\Lambda_0 = \Lambda$. We remember that embedded disks are submanifolds of $\mathcal{P}$, and they are exact if so are the immersed submanifolds.

Lemma 2.4 *Let $\{\mathcal{W}_\varepsilon\}$ be a smooth (at $\varepsilon = 0$) family of connected exact immersed submanifolds.*

(i) Let $p \in \mathcal{W}_0$ and $\Lambda \subset \mathcal{W}_0$ be an embedded disk containing p. Let $\{\Lambda_\varepsilon\}$ be a smooth family of embedded disks such that $\Lambda_\varepsilon \subset \mathcal{W}_\varepsilon$ and $\Lambda_0 = \Lambda$. Let U be a neighbourhood of p in $\mathcal{P}$, where cotangent coordinates (x, y) exist for Λ. Thus, the set $\Lambda_\varepsilon \cap U$ has the form $y = \varepsilon \partial L_\varepsilon(x)/\partial x$, for some function L_ε, since Λ_ε is an exact submanifold. We can write $L_\varepsilon = L_1 + \mathrm{O}(\varepsilon)$. Then, the function $L_1 : \mathcal{W}_0 \to \mathbf{R}$ is well-defined, that is, it neither depends on the family $\{\Lambda_\varepsilon\}$, nor on the cotangent coordinates. (Of course, L_1 is determined except for an additive constant.)

(ii) Assume that $\mathcal{W}_\varepsilon$ is invariant under some exact map F_ε. Let $S_\varepsilon = S_0 + \varepsilon S_1 + \mathrm{O}(\varepsilon^2)$ be the generating function of F_ε, and $F_1(p) = [\partial F_\varepsilon(p)/\partial \varepsilon]|_{\varepsilon=0}$ be the first order variation in ε of the family $\{F_\varepsilon\}$. Then,

$$S_1(p) - \phi(F_0(p))[F_1(p)] + \text{constant} = L_1(F_0(p)) - L_1(p), \quad \forall p \in \mathcal{W}_0. \tag{2.5}$$

Besides, the constant is $\phi(p_\infty)[F_1(p_\infty)] - S_1(p_\infty)$, if $p_\infty \in \mathcal{W}_0$ is a fixed point of F_0.

Remark 2.2 It is clear that $\Lambda_\varepsilon \cap U$ has the equation $y = \varepsilon \partial L_1(x)/\partial x + \mathrm{O}(\varepsilon^2)$. From (i), the function $L_1 : \mathcal{W}_0 \to \mathbf{R}$ is a geometrical object associated to the family $\{\mathcal{W}_\varepsilon\}$, and therefore its differential gives the first order variation at $\varepsilon = 0$ of the family along the coordinate y in *any* cotangent coordinates (x, y). We will call L_1 the *infinitesimal generating function* of the family $\{\mathcal{W}_\varepsilon\}$.

Proof. (i) On the one hand, any two families $\{\Lambda_\varepsilon\}$, $\{\Lambda'_\varepsilon\}$ coincide on a small neighbourhood of the point p. This proves the independence on the family. On the other hand, the independence on the cotangent coordinates for a fixed family is proved in [Tre94], using coordinates.

A geometric interpretation of L_1, useful in order to prove below (ii) (and consequently another proof of the fact that $L_1 : \mathcal{W}_0 \to \mathbf{R}$ is well-defined), is given now. It is inspired in a similar construction that can be found in [page 238]

Let $\mathcal{E} \subset \mathbf{R}$ be the small neighbourhood of 0 where ε runs. Given $p \in \mathcal{W}_0$, we denote by $\widehat{p} : \mathcal{E} \to \mathcal{P}$ any smooth curve such that $\widehat{p}(\varepsilon) \in \mathcal{W}_\varepsilon$ which has a non-tangent contact with $\mathcal{W}_0$ at p for $\varepsilon = 0$. Moreover, $\sigma(p,\varepsilon)$ will denote the path $\widehat{p}(\tau)$, $0 \le \tau \le \varepsilon$. Given $p, q \in \mathcal{W}_0$, let $D(p,q,\varepsilon)$ be any oriented 2-chain of $\mathcal{P}$ such that

$$\partial D(p,q,\varepsilon) = \gamma(p,q,\varepsilon) - \gamma(p,q,0) + \sigma(p,\varepsilon) - \sigma(q,\varepsilon),$$

where $\gamma(p,q,\varepsilon)$ is any path from $\widehat{p}(\varepsilon)$ to $\widehat{q}(\varepsilon)$ *in* $\mathcal{W}_\varepsilon$. Such a construction is possible, provided that $|\varepsilon|$ is small enough. Let us set

$$\Delta(p,q,\varepsilon) := -\iint_{D(p,q,\varepsilon)} \omega = \varepsilon \Delta_1(p,q) + \mathrm{O}(\varepsilon^2).$$

This integral neither depends on the symplectic coordinates, nor on the choice of the paths $\gamma(p,q,\varepsilon)$. In addition, its first order term $\Delta_1(p,q)$ does not depend on the choice of the curves $\widehat{p}$ and $\widehat{q}$, since such different choices only affect second order terms of $\Delta(p,q,\varepsilon)$.

Now, it will be shown that $L_1(q) - L_1(p) = \Delta_1(p,q)$, if $p, q \in \mathcal{W}_0$ are close enough over $\mathcal{W}_0$, that is, if there exist an embedded disk $\Lambda_0 \subset \mathcal{W}_0$ and an open $U \subset \mathcal{P}$, where cotangent coordinates (x,y) are defined, such that $p, q \in \Lambda_0 = \{y = 0\} \cap U$. We denote by $\pi : U \to \Lambda_0$ the projection $\pi(z) = p$, if $z = (x,y)$ and $p = (x,0)$ are the cotangent coordinates of z and p, respectively. We determine the curves $\widehat{p}, \widehat{q}$ by imposing $\pi \circ \widehat{p} \equiv p$, $\pi \circ \widehat{q} \equiv q$, and we choose $\gamma(p,q,\varepsilon)$ in such a way that they are contained in U. Then,

$$\begin{aligned} \Delta(p,q,\varepsilon) &= \int_{\gamma(p,q,\varepsilon)-\gamma(p,q,0)+\sigma(p,\varepsilon)-\sigma(q,\varepsilon)} y \, \mathrm{d}x = \int_{\gamma(p,q,\varepsilon)} y \, \mathrm{d}x = \varepsilon \int_{\pi(\gamma(p,q,\varepsilon))} \frac{\partial L_\varepsilon}{\partial x}(x) \, \mathrm{d}x \\ &= \varepsilon[L_\varepsilon(\pi(\widehat{q}(\varepsilon))) - L_\varepsilon(\pi(\widehat{p}(\varepsilon)))] = \varepsilon[L_1(q) - L_1(q)] + \mathrm{O}(\varepsilon^2). \end{aligned}$$

Finally, if $p, q \in \mathcal{W}_0$ are arbitrary, we consider a chain of points $(r_j)_{0 \le j \le J}$ such that $r_0 = p$, $r_J = q$, and two consecutive points of the chain are close enough so that $L_1(r_j) - L_1(r_{j-1}) = \Delta_1(r_{j-1}, r_j)$ holds. Then, a trivial argument with telescopic sums shows that $L_1(q) - L_1(p) = \Delta_1(p,q)$, since $\Delta_1(r,s) + \Delta_1(s,t) = \Delta_1(r,t)$ holds for all $r, s, t \in \mathcal{W}_0$.

(ii) Given $p \in \mathcal{W}_0$, we set $q = F_0(p)$. For any curve $\widehat{p}$ like the previous ones, let $\widehat{q}(\varepsilon) = F_\varepsilon(\widehat{p}(\varepsilon))$. If $v = (\mathrm{d}\widehat{p}/\mathrm{d}\varepsilon)(0)$, then $w = (\mathrm{d}\widehat{q}/\mathrm{d}\varepsilon)(0) = \mathrm{d}F_0(p)[v] + F_1(p)$, so v (i.e., $\widehat{p}$) can be chosen in such a way that $\widehat{q}$ is not tangent to $\mathcal{W}_0$ at q, due to the fact that the map $v \mapsto w$ is bijective. Using (i), we get

$$L_1(q) - L_1(p) = -\lim_{\varepsilon \to 0} \varepsilon^{-1} \iint_{D(p,q,\varepsilon)} \omega = \lim_{\varepsilon \to 0} \varepsilon^{-1} \int_{\gamma(p,q,\varepsilon)-\gamma(p,q,0)+\sigma(p,\varepsilon)-\sigma(q,\varepsilon)} \phi.$$

Now, by equations (2.1) and (2.2), there exist constants $c(\varepsilon)$ (independent of the point p) such that

$$\begin{aligned}
\lim_{\varepsilon\to 0}\varepsilon^{-1}\int_{\gamma(p,q,\varepsilon)-\gamma(p,q,0)}\phi &= \lim_{\varepsilon\to 0}\varepsilon^{-1}\left\{ {}_{\mathcal{W}_\varepsilon}\!\int_{\hat p(\varepsilon)}^{\hat q(\varepsilon)}\phi - {}_{\mathcal{W}_0}\!\int_p^q \phi \right\} \\
&= \lim_{\varepsilon\to 0}\varepsilon^{-1}[L_\varepsilon(\widehat{q}(\varepsilon)) - L_\varepsilon(\widehat{p}(\varepsilon)) - L_0(q) + L_0(p)] \\
&= \lim_{\varepsilon\to 0}\varepsilon^{-1}[S_\varepsilon(\widehat{p}(\varepsilon)) + c(\varepsilon) - S_0(p) - c(0)] \\
&= S_1(p) + \mathrm{d}S_0(p)[v] + (\mathrm{d}c/\mathrm{d}\varepsilon)(0).
\end{aligned}$$

Finally, we use that $F_0^*\phi - \phi = \mathrm{d}S_0$ and consequently,

$$\begin{aligned}
\lim_{\varepsilon\to 0}\varepsilon^{-1}\int_{\sigma(p,\varepsilon)-\sigma(q,\varepsilon)}\phi &= \lim_{\varepsilon\to 0}\varepsilon^{-1}\int_0^\varepsilon (\phi(\widehat{p}(\tau))\,[(\mathrm{d}\widehat{p}/\mathrm{d}\varepsilon)(\tau)] - \phi(\widehat{q}(\tau))\,[(\mathrm{d}\widehat{q}/\mathrm{d}\varepsilon)(\tau)])\,\mathrm{d}\tau \\
&= \phi(p)[v] - \phi(q)[w] = -\mathrm{d}S_0(p)[v] - \phi(F_0(p)[F_1(p))],
\end{aligned}$$

and the proof follows. □

Let $F_0 : \mathcal{P} \to \mathcal{P}$ be an exact symplectic diffeomorphism with a hyperbolic fixed point p_∞ and invariant manifolds $\mathcal{W}_0^{\mathrm{u,s}}$. Let us consider a family of exact symplectic diffeomorphisms $\{F_\varepsilon\}$, as a general perturbation of the situation above, and let $S_\varepsilon = S_0 + \varepsilon S_1 + \mathrm{O}(\varepsilon^2)$, be the generating function of F_ε. In order to simplify some formulae later, we introduce the function

$$\widehat{S}_1 : \mathcal{P} \to \mathbf{R}, \qquad \widehat{S}_1(p) = S_1(p) - \phi(F_0(p))[F_1(p)], \tag{2.6}$$

where $F_1(p) = [\partial F_\varepsilon(p)/\partial\varepsilon]|_{\varepsilon=0}$.

From the invariant manifold theory for maps [II §6] it follows that for small $|\varepsilon|$ there exists a hyperbolic fixed point $p_\infty(\varepsilon)$ of the perturbed map F_ε near p_∞. Moreover, $p_\infty(\varepsilon)$ lies in the intersection of two (connected) exact immersed submanifolds $\mathcal{W}_\varepsilon^{\mathrm{u,s}}$, and the families $\{\mathcal{W}_\varepsilon^{\mathrm{u,s}}\}$ are smooth (at $\varepsilon = 0$). We denote by $L_1^{\mathrm{u,s}}$ their infinitesimal generating functions and, as usual, we determine S_1, $L_1^{\mathrm{u,s}}$ by $\widehat{S}_1(p_\infty) = L_1^{\mathrm{u,s}}(p_\infty) = 0$.

Proposition 2.2 *Given $p^{\mathrm{u,s}} \in \mathcal{W}_0^{\mathrm{u,s}}$, let $p_k^{\mathrm{u,s}} = F_0^k(p^{\mathrm{u,s}})$, for $k \in \mathbf{Z}$. Then,*

$$L_1^{\mathrm{u}}(p^{\mathrm{u}}) = \sum_{k<0}\widehat{S}_1(p_k^{\mathrm{u}}), \qquad L_1^{\mathrm{s}}(p^{\mathrm{s}}) = -\sum_{k\geq 0}\widehat{S}_1(p_k^{\mathrm{s}}).$$

Proof. Identical to the proof of proposition 2.1, but using equation (2.5) instead of equation (2.2). □

2.3 Melnikov potential

Assume now that the invariant manifolds $\mathcal{W}_0^{\mathrm{u,s}}$ are *doubled*, that is, $\mathcal{W} := \mathcal{W}_0^{\mathrm{u}} = \mathcal{W}_0^{\mathrm{s}}$.

Then, we can consider three topologies on $\mathcal{W}$: the one induced by the inclusion $\mathcal{W} \subset \mathcal{P}$, and the two ones induced by the inclusions $\mathcal{W} \subset \mathcal{W}_0^{\mathrm{u,s}}$. We define the *bifurcation set* Σ and the *separatrix* Λ of this problem as the subset of $\mathcal{W}$ of points where the three topologies do not coincide, and $\Lambda := \mathcal{W} \setminus \Sigma$, respectively.

Lemma 2.5 *The bifurcation set and the separatrix have the following properties:*

(i) Λ *is an exact submanifold of* $\mathcal{P}$ *and* $p_\infty \in \Sigma$.

(ii) Λ *and* Σ *are* F_0*-invariant.*

(iii) Let p, p' *be points on the same connected component of* Λ. *Then, the unperturbed homoclinic orbits* $\mathcal{O}$ *and* $\mathcal{O}'$ *generated by* p *and* p', *have the same action.*

Proof.
(i) On the one hand, using (ii) of lemma 2.1, Λ is a submanifold. It must be exact, since it is contained in the exact immersed submanifolds $\mathcal{W}_0^{\mathrm{u}}$, $\mathcal{W}_0^{\mathrm{s}}$.

On the other hand, $\mathcal{W}_0^{\mathrm{u}}$ and $\mathcal{W}_0^{\mathrm{s}}$ have a transverse intersection at p_∞, so their topology at p_∞ as immersed submanifolds can not coincide and $p_\infty \in \Sigma$. (Indeed, Σ is just formed by the points of $\mathcal{W}$ where this set has self-intersections, considered as a subset of $\mathcal{P}$.)

(ii) Since $\mathcal{W}$ is F_0-invariant, it is enough to see that Σ is invariant, and this follows from the fact that F_0 is a diffeomorphism.

(iii) This is clear from remark 2.1. We can connect p and p' by a path in Λ, and so in $\mathcal{W}_0^{\mathrm{u,s}}$, since their topologies coincide on Λ. □

Remark 2.3 In the planar case with a double loop (∞), the bifurcation set is just the hyperbolic fixed point. In general, for more dimensions the situation is not so simple. For example, let $F_0 : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ be the product of n planar maps $f_j : \mathbf{R}^2 \to \mathbf{R}^2$, each one with a double loop $\Gamma_j = \{p^j_\infty\} \cup \Lambda_j$ where $p^j_\infty \in \mathbf{R}^2$ stands for the fixed point of f_j and Λ_j are the two components of $\Gamma_j \setminus \{p^j_\infty\}$, for $j = 1, \ldots, n$. Then, $\Lambda = \Lambda_1 \times \cdots \times \Lambda_n$ has 2^n connected components and $\Sigma = (\Gamma_1 \times \cdots \times \Gamma_n) \setminus \Lambda$ contains strictly the hyperbolic fixed point $p_\infty = (p^1_\infty, \ldots, p^n_\infty) \in \mathbf{R}^{2n}$. In particular, $\Lambda \neq \mathcal{W}_0^{\mathrm{u,s}} \setminus \{p_\infty\}$.

Remark 2.4 As the case of a planar map with a single loop ($\propto$) shows, the situation $\mathcal{W}_0^{\mathrm{u}} \neq \mathcal{W}_0^{\mathrm{s}}$ does not exclude that $\mathcal{W}_0^{\mathrm{u}} \cap \mathcal{W}_0^{\mathrm{s}}$ can contain n-dimensional submanifolds. For the sake of simplicity, we have defined the notion of separatrix only if the invariant manifolds are doubled and then, from the arguments above, the separatrix Λ satisfies: (a) Λ is a doubly asymptotic exact submanifold, invariant by F_0, and (b) the three topologies on Λ coincide (the ones induced by the inclusions $\Lambda \subset \mathcal{P}$, $\Lambda \subset \mathcal{W}_0^{\mathrm{u}}$, and $\Lambda \subset \mathcal{W}_0^{\mathrm{s}}$). Since these properties are the only ones needed in this section, they can be taken as a definition for a separatrix when $\mathcal{W}_0^{\mathrm{u,s}}$ are partially doubled: $\mathcal{W}_0^{\mathrm{u}} \neq \mathcal{W}_0^{\mathrm{s}}$. Thus, with this definition, the analytical results of this paper also apply to this case.

By remark 2.2, the differential of $L_1^{\mathrm{u,s}}$ gives the first order variation of $\mathcal{W}_\varepsilon^{\mathrm{u,s}}$ at $\varepsilon = 0$. Besides, since $L_1^{\mathrm{u,s}}$ is defined over $\mathcal{W}_0^{\mathrm{u,s}}$ and $\Lambda \subset \mathcal{W}_0^{\mathrm{u,s}}$, the perturbed invariant manifolds $\mathcal{W}_\varepsilon^{\mathrm{u,s}}$ can be *compared* over the separatrix Λ. For this purpose, we introduce the real-valued function

$$L : \Lambda \to \mathbf{R}, \qquad L(p) := L_1^{\mathrm{u}}(p) - L_1^{\mathrm{s}}(p) = \sum_{k \in \mathbf{Z}} \widehat{S}_1(p_k), \qquad p_k = F_0^k(p), \tag{2.7}$$

called the *Melnikov potential* of the problem. The series above is absolutely convergent since any orbit in the manifold Λ tends to p_∞ at an exponential rate as $|k| \to \infty$ and $\widehat{S}_1(p_\infty) = 0$. Thus, L is well-defined, and its differential gives the first order distance, along the coordinate y in any cotangent coordinates (x, y), between the perturbed invariant manifolds. This geometric interpretation is the fundamental point to find conditions for the splitting of the separatrices.

It still remains to check the smoothness of L on Λ. It is clear that $L_1^{\mathrm{u,s}}$ are smooth over $\mathcal{W}_0^{\mathrm{u,s}}$, but since the smooth structures on $\mathcal{W}_0^{\mathrm{u}}$, $\mathcal{W}_0^{\mathrm{s}}$ do not coincide, $L_1^{\mathrm{u}} - L_1^{\mathrm{s}}$ could be defined over the whole intersection $\mathcal{W}$ but need not be smooth on the bifurcation set Σ. Thus, it is necessary to restrict ourselves to a subset of $\mathcal{W}$ where the two smooth structures coincide, and because of this, we have defined the separatrix Λ as the set $\mathcal{W} \setminus \Sigma$ to get a smooth L on Λ.

Before stating our main analytical result, we must introduce the kind of perturbed homoclinic orbits that can be detected by "Melnikov methods". A *primary homoclinic orbit* of the perturbed problem is a perturbed homoclinic orbit $\mathcal{O}_\varepsilon$ of F_ε, defined for $|\varepsilon|$ small enough and depending in a smooth way on ε. This is a perturbative definition, since in the multi-dimensional case (contrary to the planar case, see [Wig91]), it seems difficult to give a geometric definition. Non-primary homoclinic orbits are invisible for the standard Melnikov techniques. (However, a new Melnikov-like theory has been recently developed in [Rom95], to study secondary homoclinic orbits for time-periodic perturbations of integrable planar differential equations.)

Theorem 2.1 *Under the above notations and hypothesis:*

(i) *L is F_0-invariant (i.e., $L \circ F_0 = L$).*

(ii) *If L is not locally constant, the manifolds $\mathcal{W}_\varepsilon^{\mathrm{u,s}}$ split for $0 < |\varepsilon| \ll 1$, i.e., the separatrix Λ is not preserved by the perturbation.*

(iii) *If $p \in \Lambda$ is a non-degenerate critical point of L, the manifolds $\mathcal{W}_\varepsilon^{\mathrm{u,s}}$ are transverse along a primary homoclinic orbit $\mathcal{O}_\varepsilon$ of F_ε for $0 < |\varepsilon| \ll 1$, with $\mathcal{O}_0 = (F_0^k(p))_{k\in\mathbf{Z}}$. Moreover, when all the critical points of L are non-degenerate, all the primary homoclinic orbits arising from Λ are found in this way.*

(iv) *Let $\mathcal{O}_\varepsilon$ be a primary homoclinic orbit such that $\mathcal{O}_0 = (F_0^k(p))_{k\in\mathbf{Z}}$ for some $p \in \Lambda$. Then, the homoclinic action admits the asymptotic expression $W[\mathcal{O}_\varepsilon] = W[\mathcal{O}_0] + \varepsilon L(p) + \mathrm{O}(\varepsilon^2)$. Given another orbit $\mathcal{O}'_\varepsilon$ such that $\mathcal{O}'_0 = (F_0^k(p'))_{k\in\mathbf{Z}}$ for some p' in the same connected component of Λ as p, the homoclinic area is given by*

$$\Delta W[\mathcal{O}_\varepsilon, \mathcal{O}'_\varepsilon] = \varepsilon[L(p) - L(p')] + \mathrm{O}(\varepsilon^2).$$

Proof.

(i) A shift in the index of the sum does not change its value, so L is F_0-invariant.

(ii) If $\mathrm{d}L$ is not zero, the perturbed invariant manifolds do not coincide at first order, so they split.

(iii) This result follows directly from the geometric interpretation of the Melnikov potential and the Implicit Function Theorem.

(iv) Let $\mathcal{O}_\varepsilon = (\widehat{p}_k(\varepsilon))_{k\in\mathbf{Z}}$, $p_k = \widehat{p}_k(0) = F_0^k(p)$, and $v_k = (\mathrm{d}\widehat{p}_k/\mathrm{d}\varepsilon)(0)$. From equation (2.3), $\mathrm{d}S_0 = F_0^*\phi - \phi$, and $\mathrm{d}F_0(p_k)[v_k] = v_{k+1} - F_1(p_k)$, we obtain:

$$\begin{aligned}
W[\mathcal{O}_\varepsilon] &= \sum_{k\in\mathbf{Z}} S_\varepsilon(\widehat{p}_k(\varepsilon)) \\
&= \sum_{k\in\mathbf{Z}} \{S_0(p_k) + \varepsilon(S_1(p_k) + \mathrm{d}S_0(p_k)[v_k]) + \mathrm{O}(\varepsilon^2)\} \\
&= \sum_{k\in\mathbf{Z}} S_0(p_k) + \varepsilon \sum_{k\in\mathbf{Z}} \{S_1(p_k) + \phi(p_{k+1})\,[\mathrm{d}F_0(p_k)[v_k]] - \phi(p_k)[v_k]\} + \mathrm{O}(\varepsilon^2) \\
&= W[\mathcal{O}_0] + \varepsilon \sum_{k\in\mathbf{Z}} \left\{\widehat{S}_1(p_k) + \phi(p_{k+1})[v_{k+1}] - \phi(p_k)[v_k]\right\} + \mathrm{O}(\varepsilon^2) \\
&= W[\mathcal{O}_0] + \varepsilon L(p) + \mathrm{O}(\varepsilon^2).
\end{aligned}$$

Finally, the asymptotic formula for the homoclinic area follows from its definition, using (iii) of lemma 2.5. □

Remark 2.5 The actions of homoclinic orbits arising from different connected components of the separatrix need not be equal at $\varepsilon = 0$, see remark 2.1, whereas the splitting size is of order $\mathrm{O}(\varepsilon)$. Thus, it seems inappropriate to measure the splitting comparing the action of homoclinic orbits arising from different components of Λ. For instance, in the planar case with a double loop, the geometric sense of the area between primary homoclinic orbits arising from different loops is very unclear.

Remark 2.6 If L has some non-degenerate critical point, the perturbed invariant manifolds of F_ε have a transverse intersection and, in particular, a topological crossing. Thus, using some recent results contained in [BW95], the perturbed maps have positive topological entropy, for $0 < |\varepsilon| \ll 1$.

Let us see now that the Melnikov potential is invariant under additional diffeomorphisms, if the family $\{F_\varepsilon\}$ has suitable symmetries. We recall that given a diffeomorphism $I : \mathcal{P} \to \mathcal{P}$ the family $\{F_\varepsilon\}$ is called *I-symmetric* if $F_\varepsilon \circ I = I \circ F_\varepsilon$ and $F_\varepsilon(p_\infty) = I(p_\infty) = p_\infty$, for all ε.

Lemma 2.6 *Assume that the family $\{F_\varepsilon\}$ is I-symmetric, and that the symplectic potential is preserved by the symmetry: $I^*\phi = \phi$. Then, the Melnikov potential L is I-invariant: $L \circ I = L$.*

Proof. Let $p \in \mathcal{W} = \mathcal{W}_0^{\mathrm{u,s}}$ and $q = I(p)$. Using that $F_0^k \circ I = I \circ F_0^k$ for all $k \in \mathbf{Z}$, we get

$$\lim_{k\to\infty} F_0^k(q) = \lim_{k\to\infty} I(F_0^k(p)) = I\left(\lim_{k\to\infty} F_0^k(p)\right) = I(p_\infty) = p_\infty.$$

This proves that $\mathcal{W}$ is I-invariant. Thus, the separatrix Λ also is, by the same argument as in (ii) of lemma 2.5, and the expression $L \circ I$ makes sense on Λ.

From $F_\varepsilon^*\phi - \phi = \mathrm{d}S_\varepsilon$, $I^*\phi = \phi$, and $F_\varepsilon \circ I = I \circ F_\varepsilon$ we have

$$\mathrm{d}(S_\varepsilon \circ I) = I^*(\mathrm{d}S_\varepsilon) = I^* F_\varepsilon^*\phi - I^*\phi = F_\varepsilon^* I^*\phi - \phi = F_\varepsilon^*\phi - \phi = \mathrm{d}S_\varepsilon.$$

Hence, $S_\varepsilon \circ I - S_\varepsilon$ is a constant function that evaluated at p_∞ vanishes, so S_ε (and in particular S_1) are I-invariant.

The first order terms of $F_\varepsilon \circ I = I \circ F_\varepsilon$ give $F_1 \circ I = DI(F_0)[F_1]$. Using this equality, we see that the function $\phi(F_0)[F_1]$ is also I-invariant:

$$\phi(F_0 \circ I)[F_1 \circ I] = \phi(I \circ F_0)[DI(F_0)[F_1]] = I^*\phi(F_0)[F_1] = \phi(F_0)[F_1].$$

Thus, the difference $\widehat{S}_1 = S_1 - \phi(F_0)[F_1]$ is I-invariant, too.

Finally, $L \circ I = \sum_{k\in\mathbf{Z}}(\widehat{S}_1 \circ F_0^k \circ I) = \sum_{k\in\mathbf{Z}}(\widehat{S}_1 \circ I \circ F_0^k) = \sum_{k\in\mathbf{Z}}(\widehat{S}_1 \circ F_0^k) = L.$ □

As we have seen, the differential of L measures the distance between invariant manifolds and thus $M = \mathrm{d}L$ is called the *Melnikov function* of the problem. It can be also constructed in the non-symplectic case, although it is not longer the differential of a function. We recall now this construction, but we will not go further in this direction, since the non-symplectic framework is out of the spirit of this paper. For the sake of simplicity, we only consider $\mathcal{P} = \mathbf{R}^{2n}$.

Assume that a diffeomorphism $F_0 : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ has a separatrix Λ and n first integrals $H_1, \ldots, H_n$, independent over the separatrix (but not necessarily in involution, since this concept requires a symplectic structure), and let $F_\varepsilon = F_0 + \varepsilon F_1 + \mathrm{O}(\varepsilon^2)$ be a general perturbation of F_0.

Given $p \in \Lambda$, let Π_p be the n-dimensional linear variety spanned by the point p and the vectors $\nabla H_j(p)$ ($1 \le j \le n$). Since Π_p is transverse to Λ at p, there exist $p^{\mathrm{u,s}}(\varepsilon) \in \mathcal{W}_\varepsilon^{\mathrm{u,s}} \cap \Pi_p$, depending in a smooth way on ε, such that $p^{\mathrm{u,s}}(0) = p$. A natural measure of the distance between the invariant manifolds is given by the difference of first integrals ("energies")

$$\Delta(p,\varepsilon) = H(p^{\mathrm{u}}(\varepsilon)) - H(p^{\mathrm{s}}(\varepsilon)) = \varepsilon M(p) + \mathrm{O}(\varepsilon^2), \qquad H = (H_1, \ldots, H_n)^\top,$$

where $M : \Lambda \to \mathbf{R}^n$ is the vector-valued Melnikov function of the problem. It is easy to generalize (actually, rewrite) the proof given in [DR96] for the planar case to see that

$$M(p) = \sum_{k\in\mathbf{Z}} DH(p_{k+1})[F_1(p_k)], \qquad p_k = F_0^k(p). \tag{2.8}$$

Remark 2.7 Some similar results can be found in [BGK95], although with a less geometrical (and more functional) setting. They only can prove that a *necessary* condition for the existence of primary homoclinic orbits is the existence of zeros for M. Our geometrical construction shows that the existence of non-degenerate zeros for M is a *sufficient* condition for the existence of transverse primary homoclinic orbits, even in the non-symplectic case. However, it should be noted that [BGK95] deals with a broader range of maps; for example, the existence of first integrals is not needed.

2.4 Twist maps

Now, we present another formulation of the method that is useful for the physical problems that verify the twist condition, since the formula for the Melnikov potential is simpler. For more details on twist maps, the reader is referred to [Gol94a,Gol94b,BG96]. We follow closely the notations and definitions of the later reference.

An *exact symplectic twist map* (or simply, twist map) F is a map from a connected subset U of the cotangent bundle of a manifold $\mathcal{M}$ (which can be non-compact) into U, which comes equipped with a twist generating function $\mathcal{L} : \mathcal{M} \times \mathcal{M} \to \mathbf{R}$ that satisfies

$$F^*(y\,\mathrm{d}x) - y\,\mathrm{d}x = Y\,\mathrm{d}X - y\,\mathrm{d}x = \mathrm{d}\mathcal{L}(x,X), \qquad (X,Y) = F(x,y),$$

where (x,y) are any cotangent coordinates on $T^*\mathcal{M}$, that is, x are coordinates on $\mathcal{M}$, extended to coordinates (x,y) in the obvious way. The canonical form ϕ_0 on $T^*\mathcal{M}$ reads as $\phi_0 = y\,\mathrm{d}x$ in cotangent coordinates. This can also be written in a coordinate free manner. Given $\mathcal{L}$, one can retrieve the map (at least implicitly) from $y = -\partial_1\mathcal{L}(x,X)$, and $Y = \partial_2\mathcal{L}(x,X)$. This can be done globally (i.e., $U = T^*\mathcal{M}$) only when $\mathcal{M}$ is diffeomorphic to a fiber of $T^*\mathcal{M}$, for example when $\mathcal{M}$ is the covering space of $\mathbf{T}^n$ or a manifold of constant negative curvature.

The form $F^*\phi_0 - \phi_0$ is exact, so F is exact. Let $S : U \to \mathbf{R}$ be the generating function of F, in the geometric sense of the previous definitions. Then, $S(x,y) = \mathcal{L}(x,X)$. The fact that S can be written in terms of old and new coordinates: (x,X), is the twist condition. In a coordinate free formulation it reads as

$$S(p) = \mathcal{L}(\pi(p), \pi(F(p))), \qquad \forall p \in U, \tag{2.9}$$

where $\pi : T^*\mathcal{M} \to \mathcal{M}$ is the canonical projection.

Now, we carry out the generalization of (2.9) for families of twist maps, depending (in a smooth way) on a small parameter ε. That is, we search for the relationship between the first order variations in ε of the twist and geometric generating functions.

Lemma 2.7 *Let $\{F_\varepsilon\}$ be a smooth family of twist maps. Let $\mathcal{L}_\varepsilon$ (resp. S_ε) be the twist (resp. geometric) generating function of F_ε. Set $\mathcal{L}_\varepsilon = \mathcal{L}_0 + \varepsilon\mathcal{L}_1 + \mathrm{O}(\varepsilon^2)$ and $S_\varepsilon = S_0 + \varepsilon S_1 + \mathrm{O}(\varepsilon^2)$. Then,*

$$\widehat{S}_1(p) = \mathcal{L}_1(\pi(p), \pi(F_0(p))), \qquad \forall p \in U, \tag{2.10}$$

where $\widehat{S}_1$ is the function given in (2.6).

Proof. Fix $p \in U$ and let (x,y) be cotangent coordinates in a neighbourhood of p. If we denote $(X_\varepsilon, Y_\varepsilon) = F_\varepsilon(x,y) = (X_0,Y_0) + \varepsilon(X_1,Y_1) + \mathrm{O}(\varepsilon^2)$, the $\mathrm{O}(\varepsilon)$ terms of the equality $S_\varepsilon(x,y) = \mathcal{L}_\varepsilon(x,X_\varepsilon)$ give

$$S_1(x,y) = \mathcal{L}_1(x,X_0) + \partial_2\mathcal{L}_0(x,X_0)X_1 = \mathcal{L}_1(x,X_0) + Y_0X_1.$$

Thus, from the definition of $\widehat{S}_1$ and using $\phi_0 = y\,\mathrm{d}x$ we get $\widehat{S}_1(x,y) = \mathcal{L}_1(x,X_0)$. □

Assume now that F_0 has a hyperbolic fixed point p_∞ with a separatrix $\Lambda \subset U$ and that $F_\varepsilon : U \to U$ are exact diffeomorphisms. The choice $\widehat{S}_1(p_\infty) = 0$ reads as $\mathcal{L}_1(x_\infty, x_\infty) = 0$ in the twist frame, where $x_\infty = \pi(p_\infty)$. From equation (2.10), it follows directly that the Melnikov potential (2.7) can be written as

$$L(p) = \sum_{k\in\mathbf{Z}} \mathcal{L}_1(x_k, x_{k+1}), \qquad x_k = \pi(p_k), \quad p_k = F_0^k(p). \tag{2.11}$$

This formula is simpler than (2.7), since only the first order term of the twist generating function $\mathcal{L}_\varepsilon$ appears in it.

2.5 Hamiltonian maps

One of the main ideas in dynamical systems is to study maps in order to understand flows. For example, the description of Hamiltonian systems can be carried out considering the time-T maps of their flows, which are exact maps. Thus, it is interesting to present the previous results from the Hamiltonian point of view. Besides, this allows us to compare the discrete and continuous frameworks.

Recall that a *non-autonomous Hamiltonian system* over an exact symplectic manifold $(\mathcal{P}, \omega = -\,\mathrm{d}\phi)$ is given by a real-valued function (called the *Hamiltonian*) $H : \mathcal{P} \times \mathbf{R} \to \mathbf{R}$. Then, the equations of motion have the form $\dot{p} = X_H(p,t)$, $p \in \mathcal{P}$, $t \in \mathbf{R}$, where for every fixed t, $X_H(\cdot,t)$ is the Hamiltonian field generated by $H(\cdot,t)$: $\mathrm{d}H(p,t) = \omega(p)(X_H(p,t),\cdot)$, $\forall p \in \mathcal{P}$. In symplectic coordinates (x,y) on $\mathcal{P}$, we have $\phi = y\,\mathrm{d}x$, $\omega = \mathrm{d}x \wedge \mathrm{d}y$ and the Hamiltonian equations take the canonical form

$$\dot{x} = \frac{\partial H}{\partial y}(x,y,t), \qquad \dot{y} = -\frac{\partial H}{\partial x}(x,y,t).$$

It is clear that X_H does not change if a function depending only on time is added to the Hamiltonian H. We will restrict ourselves to Hamiltonians such that generate a Hamiltonian flow, i.e., all the trajectories of X_H are defined for all time.

A *Hamiltonian map* F is the time-T map of some Hamiltonian H and for some $T > 0$, i.e., $F = \Psi^T : \mathcal{P} \to \mathcal{P}$, where $\Psi^t(p)$ stands for the solution of the Hamiltonian equations of H, with initial condition p at $t = 0$. Obviously, Hamiltonians maps are diffeomorphisms isotopic to the identity. Besides, they are exact over exact manifolds; if $i(X)\omega$ stands for the inner product of a form ω by a field X, and $\Phi : \mathcal{P} \times \mathbf{R} \to \mathcal{P} \times \mathbf{R}$ is given by $\Phi(p,t) = (\Psi^t(p), t)$, we get

$$\begin{aligned} F^*\phi - \phi &= (\Psi^T)^*\phi - (\Psi^0)^*\phi = \int_0^T \frac{\mathrm{d}}{\mathrm{d}t}[(\Psi^t)^*\phi]\,\mathrm{d}t \\ &= \int_0^T \Phi^*\{i(X_H)\,\mathrm{d}\phi + \mathrm{d}(i(X_H)\phi)\}\,\mathrm{d}t = \mathrm{d}\left[\int_0^T \Phi^*(i(X_H)\phi - H)\,\mathrm{d}t\right]. \end{aligned}$$

Thus, the generating function S of F is given by

$$S(p) = \int_{(p,0)}^{(F(p),T)} \lambda, \qquad \lambda = \phi - H\,\mathrm{d}t, \tag{2.12}$$

where the one-form λ is the so-called *Poincaré-Cartan invariant integral*, defined on the (extended) phase space $\mathcal{P} \times \mathbf{R}$, and the path of integration is the trajectory $\Phi(p,t)$, $0 \leq t \leq T$, of the (extended) flow. Now, we carry out the generalization of equation (2.12) for families of Hamiltonian maps, depending (in a smooth way) on a small parameter ε. That is, we look for the relationship between the first order variations in ε of the Hamiltonians and the generating functions of their Hamiltonian maps.

Lemma 2.8 *Let H_ε be a smooth family of non-autonomous Hamiltonians, and $\Psi^t_\varepsilon(p)$ the solution of its Hamiltonian equations with $\Psi^0_\varepsilon(p) = p$. Let F_ε and S_ε be the Hamiltonian map Ψ^T_ε and its generating function, respectively. Set $H_\varepsilon = H_0 + \varepsilon H_1 + \mathrm{O}(\varepsilon^2)$ and $S_\varepsilon = S_0 + \varepsilon S_1 + \mathrm{O}(\varepsilon^2)$. Then*

$$\widehat{S}_1(p) = -\int_0^T H_1(\Psi^t_0(p), t)\,\mathrm{d}t, \qquad \forall p \in \mathcal{P}, \tag{2.13}$$

where $\widehat{S}_1$ is the function given in (2.6).

Proof. Let $\gamma(p,\varepsilon)$ be the path in the (extended) phase space $(\Psi^t_\varepsilon(p), t)$, $0 \leq t \leq T$. Set $A_\varepsilon(p,t) = \phi(\Psi^t_\varepsilon(p))[\dot{\Psi}^t_\varepsilon(p)] - H_0(\Psi^t_\varepsilon(p), t)$, where the dot means the derivative with respect to the time t. We will use through the proof the following notations for the first variation of the considered objects:

$$F_1(p) = \left.\frac{\partial F_\varepsilon}{\partial \varepsilon}(p)\right|_{\varepsilon=0}, \quad \Psi^t_1(p) = \left.\frac{\partial \Psi^t_\varepsilon}{\partial \varepsilon}(p)\right|_{\varepsilon=0}, \quad A_1(p,t) = \left.\frac{\partial A_\varepsilon}{\partial \varepsilon}(p,t)\right|_{\varepsilon=0}.$$

Besides, we will prove below that

$$A_1(p,t) = \dot{B}_1(p,t), \qquad B_1(p,t) = \phi(\Psi^t_0(p))[\Psi^t_1(p)]. \tag{2.14}$$

From $S_\varepsilon(p) = \int_{\gamma(p,\varepsilon)}[\phi - H_\varepsilon\,\mathrm{d}t]$, $A_1 = \dot{B}_1$, $\Psi^T_1 = F_1$ and $\Psi^0_1 \equiv 0$, we get

$$\begin{aligned}
S_\varepsilon(p) &= \int_{\gamma(p,\varepsilon)}[\phi - H_0\,\mathrm{d}t] - \varepsilon\int_{\gamma(p,\varepsilon)} H_1\,\mathrm{d}t + \mathrm{O}(\varepsilon^2) \\
&= \int_0^T A_\varepsilon(p,t)\,\mathrm{d}t - \varepsilon\int_0^T H_1(\Psi^t_\varepsilon(p), t)\,\mathrm{d}t + \mathrm{O}(\varepsilon^2) \\
&= S_0(p) + \varepsilon\int_0^T \dot{B}_1(p,t)\,\mathrm{d}t - \varepsilon\int_0^T H_1(\Psi^t_0(p), t)\,\mathrm{d}t + \mathrm{O}(\varepsilon^2) \\
&= S_0(p) + \varepsilon\phi(F_0(p))[F_1(p)] - \varepsilon\int_0^T H_1(\Psi^t_0(p), t)\,\mathrm{d}t + \mathrm{O}(\varepsilon^2),
\end{aligned}$$

and the terms $\mathrm{O}(\varepsilon)$ in this equation give (2.13).

To end the proof, it only remains to check that (2.14) holds. For simplicity, we prove it using symplectic coordinates. Given $p \in \mathcal{P}$ and $t \in \mathbf{R}$, let (x,y) be symplectic coordinates in a neighbourhood of $\Psi^t_0(p)$. We denote the coordinates of $\Psi^t_\varepsilon(p)$ by $(x_\varepsilon, y_\varepsilon) = (x_0, y_0) + \varepsilon(x_1, y_1) + \mathrm{O}(\varepsilon^2)$. Thus,

$$\begin{aligned}
A_\varepsilon(p,t) &= y_\varepsilon \dot{x}_\varepsilon - H_0(x_\varepsilon, y_\varepsilon, t) \\
&= A_0(p,t) + \varepsilon[y_0\dot{x}_1 + y_1\dot{x}_0 - \partial_x H_0(x_0,y_0,t)x_1 - \partial_y H_0(x_0,y_0,t)y_1] + \mathrm{O}(\varepsilon^2) \\
&= A_0(p,t) + \varepsilon\,\mathrm{d}[y_0 x_1]/\,\mathrm{d}t + \mathrm{O}(\varepsilon^2),
\end{aligned}$$

where we have used the canonical form of Hamiltonian equations in symplectic coordinates. Finally, since in this coordinates $B_1 = y_0 x_1$, equation (2.14) follows. □

Henceforth, we restrict ourselves to time-periodic Hamiltonians H_ε, being T their period.

Assume now that F_0 has a hyperbolic fixed point p_∞ with a separatrix Λ. In the Hamiltonian frame, the choice $\widehat{S}_1(p_\infty) = 0$ becomes $\int_0^T H_1(\Psi_0^t(p_\infty), t)\, \mathrm{d}t = 0$. Indeed, it is possible (and more usual) to determine the Hamiltonian in such a way that it verifies the stronger condition $H_1(\Psi_0^t(p_\infty), t) \equiv 0$. From equation (2.13), it follows easily that the Melnikov potential (2.7) can be written as

$$L(p) = -\int_{\mathbf{R}} H_1(\Psi_0^t(p), t)\, \mathrm{d}t, \tag{2.15}$$

since $\Psi_0^t(F_0^k(p)) = \Psi_0^{t+kT}(p)$, for all integer k and real t, and H_1 is T-periodic in t. (This is the reason to consider only periodic Hamiltonians.)

We want to emphasize that the Hamiltonian version of the Melnikov potential can be deduced directly in the continuous frame, without appealing to discrete tools. However, taking into account the theory already developed in this paper, it has been easier to work directly on Hamiltonian maps.

Remark 2.8 Usually, the unperturbed Hamiltonian H_0 is time independent. In fact, in most of the applications it is Liouville integrable.

Remark 2.9 Using the Lagrangian formalism instead of the Hamiltonian one, a similar formula to (2.15) can be obtained for *Lagrangian maps* (i.e., time-T maps of some Euler-Lagrangian flow), but with $-H_1$ replaced by the first order in ε of the Lagrangian.

2.6 Lower Bounds

Along this subsection, we will assume without explicit mention that: (a) the invariant manifolds are doubled, that is, $\mathcal{W}_0^{\mathrm{u}} = \mathcal{W}_0^{\mathrm{s}}$, and (b) the bifurcation set is minimal, i.e., $\Sigma = \{p_\infty\}$. (Remember that the hyperbolic fixed point p_∞ is always contained in the bifurcation set Σ, see (i) of lemma 2.5.) These hypothesis are equivalent to require that the separatrix is $\Lambda = \mathcal{W}_0^{\mathrm{u,s}} \backslash \{p_\infty\}$. We will say that the invariant manifolds are *completely doubled* in this case. Besides, we also assume $n > 1$, to avoid trivial degenerate cases. (In particular, the separatrix is connected.)

To avoid a tedious exposition, several standard computations about Betti numbers are omitted. The expert reader in differential and algebraic topology will be able to fill in the gaps without difficulty, and we prefer to give the appropriate references for the novice one, instead of writing here a treatise. Thus, for a general discussion of Morse theory we refer to [Hir76], and for thorough discussions of homology the reader is urged to consult [Swi75,GH81].

The quotient manifold $\Lambda^* := \Lambda/F_0$, consisting of unperturbed homoclinic orbits of Λ, will be called the *reduced separatrix* (of the unperturbed map). It is shown below that

Λ^* is a compact manifold without boundary. Since the Melnikov potential L is invariant under F_0, we can consider it defined over the reduced separatrix. (The new function is called L, too.) We search for lower bounds of the number of homoclinic orbits and the main idea is to apply the Morse's inequalities to the map $L : \Lambda^* \to \mathbf{R}$.

The presence of symmetries and/or reversions usually leads to better results concerning the existence of homoclinic orbits. Let us introduce the (anti)symmetries that allow us to improve the lower bounds. We will say that the family $\{F_\varepsilon\}$ is *antisymmetric* if $\{F_\varepsilon\}$ is I-symmetric, for some involution I preserving the symplectic potential such that $DI(p_\infty) = -\mathrm{Id}$. As it is well-known, involutions are locally conjugate to their linear parts at fixed points. Thus, there exist coordinates $z = (z_1, \ldots, z_{2n})$ in some neighbourhood of p_∞ such that $I(z) = -z$, that is, the maps F_ε are *odd* in some coordinates defined close to p_∞. The definition above of antisymmetric maps is intended to translate the main features of odd maps on $(\mathbf{R}^{2n}, \mathrm{d}x \wedge \mathrm{d}y)$ to maps on general exact manifolds.

Under these hypotheses, lemma 2.6 claims that the Melnikov potential is I-invariant. Thus, we can consider L defined over the quotient manifold $\Lambda^*_I := \Lambda/\{F_0, I\}$, which has a richer topological structure than Λ^*, in the sense that Morse theory gives better lower bounds of the number of homoclinic orbits.

We recall that a real-valued smooth function over a compact manifold without boundary is called a *Morse function* when all its critical points are non-degenerate. It is very well-known that the set of Morse functions is open and dense in the set of real-valued smooth functions [page 147] Thus, to be a Morse function is a condition of generic position. Now we can state a result about the number of primary homoclinic orbits that persist under a general perturbation. In section 3, we will verify the optimality of this result for specific examples.

Theorem 2.2 *Assume that $L : \Lambda^* \to \mathbf{R}$ is a Morse function. Then the number of primary homoclinic orbits is at least 4. If the family $\{F_\varepsilon\}$ is antisymmetric, there exist at least $2n$ antisymmetric pairs of primary homoclinic orbits, and so at least $4n$ primary homoclinic orbits.*

Proof. From the celebrated Morse's inequalities, a Morse function over a n-dimensional compact manifold without boundary X has at least $SB(X; R) := \sum_{q=0}^{n} \beta_q(X; R)$ critical points, where $\beta_q(X; R)$ are the *R-Betti numbers* of X and R is any field. Let us recall that $\beta_q(X; R)$ is the dimension of the q-th singular homology R-vector space of X, noted $H_q(X, R)$.

In the antisymmetric case, $I^2(p) = p \neq I(p)$, for all $p \in \Lambda$. Thus (Λ^*, Π) is a covering space of Λ^*_I of *two* sheets, where $\Pi : \Lambda^* \to \Lambda^*_I$ is the canonical projection onto the quotient of Λ^* by the antisymmetry I. In particular, $L : \Lambda^*_I \to \mathbf{R}$ is a Morse function if and only if the same happens to $L : \Lambda^* \to \mathbf{R}$, and each critical point $\mathcal{Q}$ of $L : \Lambda^*_I \to \mathbf{R}$ corresponds to an antisymmetric pair of critical points $\Pi^{-1}(\mathcal{Q}) = \{\mathcal{O}, I(\mathcal{O})\}$ of $L : \Lambda^* \to \mathbf{R}$, for some unperturbed homoclinic orbit $\mathcal{O} \in \Lambda^*$.

Now the theorem follows from the formulae $SB(\Lambda^*; \mathbf{Z}_2) = 4$ and $SB(\Lambda^*_I; \mathbf{Z}_2) = 2n$. The rest of the proof is devoted to check that these formulae hold.

Since Betti numbers are topological invariants, we look for topological spaces homeomorphic to Λ^* and Λ_I^* whose homologies can be easily computed. To accomplish it, let us consider the restriction $f^{\mathrm{u,s}}$ of F_0 to $\mathcal{W}_0^{\mathrm{u,s}}$, and denote $B^{\mathrm{u,s}} = Df^{\mathrm{u,s}}(p_\infty)$. Since F_0 is symplectic, $\det(B^{\mathrm{u}}) \cdot \det(B^{\mathrm{s}}) = 1$, so $\det(B^{\mathrm{u}})$ and $\det(B^{\mathrm{s}})$ have the same sign. When these signs are positive (resp. negative) the map F_0 preserves (resp. reverses) the orientation of Λ, and we denote by $\sigma = +$ (resp. $\sigma = -$) the so-called *index of orientation*. In the following lemma it is shown that the topological classification of f^{u} only depends on σ. This will allow us to classify Λ^* and Λ_I^* just in terms of σ.

Lemma 2.9 *Let $A_\pm : \mathbf{R}^n \to \mathbf{R}^n$ be the linear isomorphisms given by:*

$$A_\pm(x) = 2x_\pm, \qquad x = (x_1, \ldots, x_n), \quad x_\pm = (\pm x_1, x_2, \ldots, x_n).$$

Then, there exists a global topological conjugation between f^{u} and A_σ, that is, a homeomorphism $g : \mathbf{R}^n \to \mathcal{W}_0^{\mathrm{u}}$ such that $f^{\mathrm{u}} \circ g = g \circ A_\sigma$. In the antisymmetric case, the conjugation g can be chosen in such a way that $g(-x) = I(g(x))$.

Proof. We note that p_∞ is a hyperbolic fixed point of f^{u}, and all the eigenvalues of B^{u} have modulus greater than one. From [Th. 5.5, II §5] is locally conjugated at p_∞ to A_+ (resp. A_-) in the orientation-preserving (resp. orientation-reversing) case. This local conjugation can be extended to a global one, using that f^{u} and A_σ are global repulsors. The existence of an antisymmetric conjugation (certainly, a very intuitive fact) follows the same lines. We omit the details. □

Thanks to Lemma 2.9, we now easily introduce *time-energy coordinates* (t, a) on Λ. First, we give some notations. We denote by $\mathbf{S}^n$, $\mathbf{T}^n$, and $\mathbf{P}^n$, the n-dimensional sphere, the n-dimensional torus, and the n-dimensional projective space, respectively. Besides, we introduce the n-dimensional manifold

$$\mathbf{X}^n := \mathbf{R} \times \mathbf{S}^{n-1},$$

and the homeomorphism $\eta : \mathbf{X}^n \to \mathbf{R}^n \setminus \{0\}$, $\eta(t, a) = 2^t a$, whose inverse is given by $\eta^{-1}(x) = (\hat{t}(x), \hat{a}(x)) = (\log_2 |x|, x/|x|)$. Then, $\hat{t}(A_\pm x) = \hat{t}(2x_\pm) = \hat{t}(x) + 1$ and $\hat{a}(A_\pm x) = \hat{a}(2x_\pm) = (\hat{a}(x))_\pm$, so $A_\pm \circ \eta = \eta \circ \rho_\pm$, where the map $\rho_\pm : \mathbf{X}^n \to \mathbf{X}^n$ is

$$\rho_\pm(t, a) = (t + 1, a_\pm), \qquad a = (a_1, \ldots, a_n), \quad a_\pm = (\pm a_1, a_2, \ldots, a_n).$$

Thus, $F_0 : \Lambda \to \Lambda$ and $\rho_\sigma : \mathbf{X}^n \to \mathbf{X}^n$ are topologically conjugated by $g \circ \eta$, where g is the conjugation given in lemma 2.9. This proves that $\Lambda^* = \Lambda/F_0$ and $\mathbf{X}_\sigma^n := \mathbf{X}^n/\rho_\sigma$ are homeomorphic. Hence, $SB(\Lambda^*; \mathbf{Z}_2) = SB(\mathbf{X}_\sigma^n; \mathbf{Z}_2)$.

Concerning the antisymmetric case, we note that $\eta \circ \jmath = -\eta$, where

$$\jmath : \mathbf{X}^n \to \mathbf{X}^n, \qquad \jmath(t, a) = (t, -a).$$

Thus, the pairs of maps $F_0, I : \Lambda \to \Lambda$ and $\rho_\sigma, \jmath : \mathbf{X}^n \to \mathbf{X}^n$ are *simultaneously* topologically conjugated by $g \circ \eta$. This proves that $\Lambda_I^* = \Lambda/\{F_0, I\}$ and $\mathbf{Y}_\sigma^n := \mathbf{X}^n/\{\rho_\sigma, \jmath\}$ are homeomorphic. Hence, $SB(\Lambda_I^*; \mathbf{Z}_2) = SB(\mathbf{Y}_\sigma^n; \mathbf{Z}_2)$.

Consequently, it only remains to prove that $SB(\mathbf{X}^n_\pm;\mathbf{Z}_2)=4$ and $SB(\mathbf{Y}^n_\pm;\mathbf{Z}_2)=2n$.

First, we consider the case $\sigma=+$. In this case, $\mathbf{X}^n_+=\mathbf{S}^1\times\mathbf{S}^{n-1}$ and $\mathbf{Y}^n_+=\mathbf{S}^1\times\mathbf{P}^{n-1}$, since $\mathbf{S}^1=\mathbf{R}/\{t=t+1\}$ and $\mathbf{P}^{n-1}=\mathbf{S}^{n-1}/\{a=-a\}$. Therefore, from the well-known $\mathbf{Z}_2$-homologies

$$H_q(\mathbf{S}^m;\mathbf{Z}_2)\cong\left\{\begin{array}{ll}\mathbf{Z}_2 & \text{if } q=0,m\\ 0 & \text{otherwise}\end{array}\right. \qquad H_q(\mathbf{P}^m;\mathbf{Z}_2)\cong\left\{\begin{array}{ll}\mathbf{Z}_2 & \text{if } 0\leq q\leq m\\ 0 & \text{otherwise}\end{array}\right.,$$

and the Künneth's Formula $H_q(X\times Y;\mathbf{Z}_2)\cong\bigoplus_{p=0}^q H_p(X;\mathbf{Z}_2)\otimes H_{q-p}(Y;\mathbf{Z}_2)$, we get

$$H_q(\mathbf{X}^2_\pm;\mathbf{Z}_2)\cong\left\{\begin{array}{ll}\mathbf{Z}_2 & \text{if } q=0,2\\ \mathbf{Z}_2\oplus\mathbf{Z}_2 & \text{if } q=1\\ 0 & \text{otherwise}\end{array}\right., \quad H_q(\mathbf{X}^n_\pm;\mathbf{Z}_2)\cong\left\{\begin{array}{ll}\mathbf{Z}_2 & \text{if } q=0,1,n-1,n\\ 0 & \text{otherwise}\end{array}\right.$$

for all $n>2$, and

$$H_q(\mathbf{Y}^n_\pm;\mathbf{Z}_2)\cong\left\{\begin{array}{ll}\mathbf{Z}_2 & \text{if } q=0,n\\ \mathbf{Z}_2\oplus\mathbf{Z}_2 & \text{if } q=1,\ldots,n-1\\ 0 & \text{otherwise}\end{array}\right.,$$

for all $n>1$. Adding dimensions, we get $SB(\mathbf{X}^n_+;\mathbf{Z}_2)=4$ and $SB(\mathbf{Y}^n_+;\mathbf{Z}_2)=2n$.

Finally, a standard Mayer-Vietoris sequence argument shows that the $\mathbf{Z}_2$-homologies of $\mathbf{X}^n_\sigma$ and $\mathbf{Y}^n_\sigma$ do not depend on σ, so $SB(\mathbf{X}^n_-;\mathbf{Z}_2)=4$ and $SB(\mathbf{Y}^n_-;\mathbf{Z}_2)=2n$. □

Remark 2.10 Since the case $\sigma=-$ is more intricate, one could believe that it is better to replace the maps with their squares to get $\sigma=+$. However, it should be noted that the lower bounds obtained in this way are worse since a single homoclinic orbit consist of two different ones for the square map: one gets 2 and $2n$, instead of 4 and $4n$, as the number of homoclinic orbits. Thus, the case $\sigma=-$ deserves its own separate study. We also remark that this case cannot appear in the continuous frame, since the maps generated by a flow are isotopic to the identity.

3 Standard-like maps

As a first example we deal with standard-like maps over the symplectic manifold $(\mathcal{P},\omega)=(\mathbf{R}^{2n},\,\mathrm{d}x\wedge\mathrm{d}y)$, $n>1$, which are ones of the most celebrated examples of twist maps. Among them, we consider perturbations of maps with *central symmetry*, since then the dynamics over the unperturbed separatrix is essentially one-dimensional and gives rise to explicit computations, as already announced in [DR97c]. In the sequel, given $x,y\in\mathbf{R}^n$, $x^\top y$ and $|x|$ stand for the scalar product $\sum_{i=0}^n x_i y_i$ and the Euclidean norm $\sqrt{x^\top x}$.

3.1 Central standard-like maps

Let $V : \mathbf{R}^n \to \mathbf{R}$ be a function. The map $F : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ with equations $F(x, y) = (y, -x + \nabla V(y))$ is called the *standard-like map* with *potential* V. It is immediate to check that $\mathcal{L}(x, X) = -x^\top X + V(X)$ is a twist generating function of F, so F is a twist map. When V is even, F is odd.

It is worth mentioning that standard-like maps are also expressed in the literature as $F(x', y') = (x' + y' + \nabla U(x'), y' + \nabla U(x'))$, for some function U. The symplectic linear change of variables $(x', y') = (y, y - x)$ is the bridge between these two equivalent formulations, and the relation between the potentials is given by $V(y) = |y|^2 + U(y)$. Thus, it makes no difference which formulation is used, since we deal with symplectic invariants.

A *central standard-like map* is a standard-like map with a central potential, i.e., $V(y) = V_{\mathrm{c}}(|y|^2)$ for some function $V_{\mathrm{c}} : [0, \infty) \to \mathbf{R}$. Central standard-like maps are odd and have the "angular momenta" $A_{ij}(x, y) = x_i y_j - x_j y_i$ as first integrals. We denote by $\mathcal{A}_0^{n+1} = \{(x, y) : A_{ij}(x, y) = 0\}$ the $(n+1)$-dimensional manifold in $\mathbf{R}^{2n}$ of zero angular momenta. Clearly, $\mathcal{A}_0^{n+1} = \{(qa, pa) : a \in \mathbf{S}^{n-1},\ (q, p) \in \mathbf{R}^2\}$.

Let F be a central standard-like map with potential V, and $f : \mathbf{R}^2 \to \mathbf{R}^2$ the standard-like area preserving map defined by $f(q, p) = (p, -q + 2V_{\mathrm{c}}'(p^2)p)$. We will call f the *reduced map* (in $\mathcal{A}_0^{n+1}$) of F. This definition becomes clear when it is noted that

$$f(q, p) = (Q, P) \Longleftrightarrow F(qa, pa) = (Qa, Pa), \qquad \forall (q, p) \in \mathbf{R}^2,\ a \in \mathbf{S}^{n-1}. \tag{3.1}$$

Our interest in central standard-like maps is motivated by the following lemma, which follows easily from (3.1).

Lemma 3.1 *Let F be a central standard-like map and f its reduced map. Assume that* $\operatorname{Spec}[Df(0)] = \{\mathrm{e}^{\pm h}\}$, *for some $h > 0$, and hence that the origin is a hyperbolic fixed point of f. Then:*

(i) The origin is a hyperbolic fixed point of F. Moreover, $\operatorname{Spec}[DF(0)] = \{\mathrm{e}^{\pm h}\}$.

(ii) Suppose now that f has a separatrix Γ. Then, the invariant manifolds of F are completely doubled, giving rise to the separatrix

$$\Lambda = \{(qa, pa) : (q, p) \in \Gamma,\ a \in \mathbf{S}^{n-1}\}.$$

(iii) Let $\sigma = (q, p) : \mathbf{R} \to \Gamma$ be a natural parametrization *of the separatrix Γ, i.e., σ is a diffeomorphism that satisfies $f(\sigma(t)) = \sigma(t + h)$, for all $t \in \mathbf{R}$. Then, the diffeomorphism $\lambda : \mathbf{R} \times \mathbf{S}^{n-1} \to \Lambda$ defined by $\lambda(t, a) := (q(t)a, p(t)a)$ satisfies*

$$F(\lambda(t, a)) = \lambda(t + h, a), \quad \forall t \in \mathbf{R},\ a \in \mathbf{S}^{n-1}. \tag{3.2}$$

We note that f is odd, so when it has a separatrix, it has in fact a double (symmetric) loop.

The separatrix Λ is analytically diffeomorphic to $\mathbf{R} \times \mathbf{S}^{n-1}$, by means of λ. Thus, from now on, the functions defined over Λ will be expressed as functions of the time-energy coordinates $(t, a) \in \mathbf{R} \times \mathbf{S}^{n-1}$.

Now, we introduce the McLachlan map [McL94] as the central standard-like map with potential $V_0(y) = \mu \ln(1 + |y|^2)$ ($\mu \in \mathbf{R}$). It has the expression

$$F_0(x, y) = \left(y, -x + \frac{2\mu y}{1 + |y|^2} \right). \tag{3.3}$$

It is easy to check that for $\mu > 1$ the reduced map of (3.3)—usually called McMillan map—has a separatrix to the origin. (See figure 2 for a representation of the invariant curves.) In addition, the following natural parametrization of its separatrix can be found in [GPB89,DR96]: $\sigma(t) = (q(t), p(t))$, where $q(t) = p(t - h)$ and $p(t) = \sinh(h) \operatorname{sech}(t)$. Thus, using lemma 3.1, the McLachlan map has its invariant manifolds completely doubled, and the function λ given by

$$\lambda(t, a) = (p(t - h)a, p(t)a), \quad p(t) = \sinh(h) \operatorname{sech}(t), \quad \cosh(h) = \mu (> 1), \tag{3.4}$$

verifies equation (3.2).

Remark 3.1 The McLachlan map has n first integrals H_j ($j = 1, \ldots, n$), independent over its separatrix: $H_1(x, y) = |x|^2 + |y|^2 + |x|^2 |y|^2 - 2\mu x^\top y$, and the angular momenta $H_j = A_{1j}$ ($j = 2, \ldots, n$). This is not important for our purposes, but it would be essential for the study of non-symplectic perturbations with the Melnikov function (2.8).

3.2 Standard-like perturbations

Let us consider a general perturbation of (3.3) that preserves the standard character, i.e.,

$$F_\varepsilon(x, y) = \left(y, -x + \frac{2\mu y}{1 + |y|^2} + \varepsilon \nabla V(y) \right), \quad \mu > 1, \, \varepsilon \in \mathbf{R}, \tag{3.5}$$

where $V : \mathbf{R}^n \to \mathbf{R}$. We determine V by imposing $V(0) = 0$. Then, the twist generating function of F_ε that vanishes at the origin is $\mathcal{L}_\varepsilon = \mathcal{L}_0 + \varepsilon \mathcal{L}_1$, where $\mathcal{L}_0(x, X) = -x^\top X + \mu \ln(1 + |X|^2)$ and $\mathcal{L}_1(x, X) = V(X)$.

Using formulae (2.7), (3.2) and (3.4), the Melnikov potential of the problem is

$$L : \mathbf{R} \times \mathbf{S}^{n-1} \to \mathbf{R}, \quad L(t, a) = \sum_{k \in \mathbf{Z}} V(p(t + hk)a), \quad p(t) = \frac{\sinh(h)}{\cosh(t)}. \tag{3.6}$$

Obviously, L is h-periodic in t, so we can consider t defined modulo h and L as a function over $\mathbf{S}^1 \times \mathbf{S}^{n-1}$. Henceforth it will be assumed that $h > 0$, $\cosh(h) = \mu$.

Now, we focus our attention on entire perturbations, i.e., maps (3.5) with V an entire function. The result about the splitting in this case is the following one.

Theorem 3.1 *If V is entire but not identically zero, then the manifolds $\mathcal{W}_\varepsilon^{\mathrm{u,s}}$ of the map (3.5) split, for $0<|\varepsilon|\ll 1$.*

Proof. By theorem 2.1, it is sufficient to check that the Melnikov potential (3.6) is not constant.

First, we note that the only singularities of $p(t)$ are simple poles at any point $t_p \in \pi\,\mathrm{i}/2+\pi\,\mathrm{i}\mathbf{Z}$, and therefore it is analytic at t_p+hk for $k\in\mathbf{Z}\setminus\{0\}$. Now, let $V_a, f_a:\mathbf{R}\to\mathbf{R}$ be the functions defined by $V_a(t):=V(ta)$ and $f_a(t)=V_a(p(t))=V(p(t)a)$ $(a\in\mathbf{S}^{n-1})$. Since V is a non-zero entire function, there exists $\widehat{a}\in\mathbf{S}^{n-1}$ such that $V_{\widehat{a}}$ is a non-zero entire function. Thus, $f_{\widehat{a}}$ has non-removable singularities at any point $t_p\in\pi\,\mathrm{i}/2+\pi\,\mathrm{i}\mathbf{Z}$, and however it is analytic at t_p+hk for $k\in\mathbf{Z}\setminus\{0\}$. Consequently, $L_{\widehat{a}}(t):=L(t,\widehat{a})=\sum_{k\in\mathbf{Z}} f_{\widehat{a}}(t+kh)$ has a non-removable singularity at any point in $\pi\,\mathrm{i}/2+h\mathbf{Z}+\pi\,\mathrm{i}\mathbf{Z}$. This proves that L is not constant. □

Remark 3.2 The assumption of entire function on V has only been used to ensure that there exist $t_p\in\pi\,\mathrm{i}/2+\pi\,\mathrm{i}\mathbf{Z}$ and $\widehat{a}\in\mathbf{S}^{n-1}$, such that $f_{\widehat{a}}(t)$ has an isolated singularity at t_p, and however is analytic on t_p+hk for $k\in\mathbf{Z}\setminus\{0\}$. Thus, this assumption on V can be relaxed, although the entire case is the simplest case to study.

We observe that for even V, the maps F_ε are odd and hence the family $\{F_\varepsilon\}$ is antisymmetric. Therefore, theorem 2.2 gives the following corollary.

Corollary 3.1 *Assume that the function L given in (3.6) is a Morse function. Then, the map (3.5) has at least 4 primary homoclinic orbits, for $0<|\varepsilon|\ll 1$. If, in addition, the potential V is an even function, there exist at least $2n$ antisymmetric pairs of primary homoclinic orbits, and so at least $4n$ primary homoclinic orbits.*

3.3 Polynomial perturbations: explicit computations

We show here that explicit computations of Melnikov potentials can be performed, for any polynomial perturbations of the McLachlan map, i.e., maps (3.5) with $V(y)=\sum_{\ell=1}^N V_\ell(y)$, for some finite N, where V_ℓ denotes a homogeneous polynomial of order ℓ.

In this case, the Melnikov potential (3.6) turns out to be a linear combination of products of certain elliptic functions Σ_ℓ in the variable $t\in\mathbf{C}$ (of periods h, $2\pi\,\mathrm{i}$) and the homogeneous polynomials V_ℓ restricted to $\mathbf{S}^{n-1}$:

$$L(t,a)=\sum_{\ell=1}^N \sinh^\ell(h)V_\ell(a)\Sigma_\ell(t),\qquad \Sigma_\ell(t)=\sum_{k\in\mathbf{Z}}\left[\operatorname{sech}(t+hk)\right]^\ell. \tag{3.7}$$

Using the *Summation Formula* of the Appendix, all the elliptic functions Σ_ℓ (and consequently, the Melnikov potentials) can be explicitly computed. However, using the Summation Formula to find Σ_ℓ for big values of ℓ is rather tedious. It is better to use an idea contained in [GPB89]. The point is to note that the odd (respectively, even) powers of the hyperbolic function sech can be expressed as a linear combination, with rational coefficients, of the even derivatives of sech (respectively, sech^2). This allows us to write Σ_ℓ as a linear combination, with rational coefficients, of the even derivatives of Σ_1 (if ℓ is odd) or Σ_2 (if ℓ is even). For example, $\mathrm{sech}^3 = (\mathrm{sech} - \mathrm{sech}'')/2$ and $\mathrm{sech}^4 = [4\,\mathrm{sech}^2 - (\mathrm{sech}^2)'']/6$, so $\Sigma_3 = (\Sigma_1 - \Sigma_1'')/2$ and $\Sigma_4 = (4\Sigma_2 - \Sigma_2'')/6$. Consequently, it is enough to compute Σ_ℓ for $\ell = 1, 2$. This is done in lemma A.1 (see the Appendix) and the result is:

$$\begin{aligned} \Sigma_1(t) &= \left(\frac{2K_{2\pi}}{h}\right)\left[\sqrt{m_{2\pi}}\,\mathrm{cn}\left(\frac{4K_{2\pi}t}{h}\middle|\, m_{2\pi}\right) + \mathrm{dn}\left(\frac{4K_{2\pi}t}{h}\middle|\, m_{2\pi}\right)\right], \\ \Sigma_2(t) &= \left(\frac{2K_\pi}{h}\right)^2\left[\frac{E'_\pi}{K'_\pi} - 1 + \mathrm{dn}^2\left(\frac{2K_\pi t}{h}\middle|\, m_\pi\right)\right], \end{aligned}$$

where, if $K(m)$ and $E(m)$ are the elliptic integrals of the first and second kind, the parameter $m = m_T$ ($T = \pi, 2\pi$) of the Jacobian elliptic functions is determined by the equation $K(1-m_T)/K(m_T) = T/h$; and $K_T = K(m_T)$, $K'_T = K(1-m_T)$, $E'_T = E(1-m_T)$. It is equivalent to choose $q = q_T = \mathrm{e}^{-\pi T/h}$ as the nome of the elliptic functions. For the notations about elliptic functions we refer again to the Appendix.

Assume now that $V = V_2$, i.e., V is a quadratic form or, in other words, the perturbation ∇V is linear. We can write $V(y) = y^\top B y$, for some symmetric $n \times n$ matrix B. Then, there exists an orthogonal matrix $Q = (q_1 \cdots q_n)$ such that $\mathrm{diag}(b_1, \ldots, b_n) = Q^\top B Q$, where b_i are the eigenvalues of B and q_i are their respective (normalized) eigenvectors.

Proposition 3.1 *Suppose* $\det(B) \neq 0$ *and that* B *does not have multiple eigenvalues. Then:*

1. *The invariant manifolds* $\mathcal{W}_\varepsilon^{\mathrm{u,s}}$ *are transverse along exactly* $4n$ *primary homoclinic orbits* $\mathcal{O}_{\sigma,\pm i}(\varepsilon)$ ($\sigma \in \{0,1\}$, $i \in \{1,\ldots,n\}$), *for* $0 < |\varepsilon| \ll 1$. *These perturbed homoclinic orbits are created from the unperturbed ones*

$$\mathcal{O}_{\sigma,\pm i}(0) = (\lambda\,(\sigma h/2 + kh, \pm q_i))_{k\in\mathbf{Z}}\,, \quad \sigma \in \{0,1\},\ i \in \{1,\ldots,n\}.$$

2. *The homoclinic area between the primary homoclinic orbits* $\mathcal{O}_{\sigma,\pm i}(\varepsilon)$ *and* $\mathcal{O}_{\tau,\pm j}(\varepsilon)$ *is given by the asymptotic expression*

$$\Delta W\left[\mathcal{O}_{\sigma,\pm i}(\varepsilon), \mathcal{O}_{\tau,\pm j}(\varepsilon)\right] = \varepsilon\Delta_{\sigma,\tau,i,j} + \mathrm{O}(\varepsilon^2),$$

where

$$\Delta_{\sigma,\tau,i,j} = \Delta_{\sigma,\tau,i,j}(h) = \sinh^2(h)(2K_\pi/h)^2\left[b_i\delta_\sigma - b_j\delta_\tau\right],$$

with $\delta_0 = E'_\pi/K'_\pi$ *and* $\delta_1 = E'_\pi/K'_\pi - m_\pi$.

Proof. We note that $Q(\mathbf{S}^{n-1}) = \mathbf{S}^{n-1}$, so we can perform the change of variables $a \leftrightarrow Qa$ in $\mathbf{S}^{n-1}$ and then, $V(Qa) = \sum_{i=1}^n b_i(a_i)^2$, where $b_i \neq 0$, for all i, and $b_i \neq b_s$, for all $i \neq s$. It is easy to check that the only critical points of the restriction of V to $\mathbf{S}^{n-1}$ are

$\{\pm q_i : 1 \leq i \leq n\}$, all of them being non-degenerate. Moreover, from the properties of the Jacobian elliptic function $\mathrm{dn}(u|m)$, the real critical points of Σ_2 are $\{kh/2 : k \in \mathbf{Z}\}$, that are also non-degenerate. Consequently, L is a Morse function over $(\mathbf{R}/h\mathbf{Z}) \times \mathbf{S}^{n-1}$ and its critical points are $(\sigma h/2, \pm q_i)$, for $\sigma \in \{0,1\}$, $i \in \{1,\ldots,n\}$. Now the first part of the proposition follows from theorem 2.1.

For the second part, it is enough to observe that

$$\Delta W\left[\mathcal{O}_{\sigma,\pm i}(\varepsilon), \mathcal{O}_{\tau,\pm j}(\varepsilon)\right] = \varepsilon[L(\sigma h/2, \pm q_i) - L(\tau h/2, \pm q_j)] + \mathrm{O}(\varepsilon^2),$$

and $L(\sigma h/2, \pm q_i) = \sinh^2(h) V(\pm q_i) \Sigma_2(\sigma h/2) = \sinh^2(h) b_i (2K_\pi/h)^2 \delta_\sigma$, where we have used that $\mathrm{dn}(0|m) = 1$ and $\mathrm{dn}(K|m) = \sqrt{1-m}$. □

Finally, we study the linear potentials (constant perturbations ∇V), that is, $V = V_1$. Thus, $V(y) = b^\top y$, for some vector $b \in \mathbf{R}^n \setminus \{0\}$, and the critical points of V in $\mathbf{S}^{n-1}$ are $\pm q$, where $q = b/|b|$. Of course, they are non-degenerate. Then, using the same arguments as in the proof of the preceding proposition, we get the following result.

Proposition 3.2 *With the previous notations and assumptions:*

1. *The invariant manifolds $\mathcal{W}_\varepsilon^{\mathrm{u,s}}$ are transverse along exactly* 4 *primary homoclinic orbits, $\widehat{\mathcal{O}}_{\pm\sigma}(\varepsilon)$ ($\sigma \in \{0,1\}$), for $0 < |\varepsilon| \ll 1$. These perturbed homoclinic orbits are created from the unperturbed ones*

$$\widehat{\mathcal{O}}_{\pm\sigma}(0) = \left(\lambda(\sigma h/2 + kh, \pm q)\right)_{k\in\mathbf{Z}}, \quad \sigma \in \{0,1\}.$$

2. *The homoclinic area between the primary homoclinic orbits $\widehat{\mathcal{O}}_{\pm\sigma}(\varepsilon)$ and $\widehat{\mathcal{O}}_{\pm\tau}(\varepsilon)$ is given by the asymptotic expression*

$$\Delta W\left[\widehat{\mathcal{O}}_{\pm\sigma}(\varepsilon), \widehat{\mathcal{O}}_{\pm\tau}(\varepsilon)\right] = \varepsilon\widehat{\Delta}_{\pm\sigma,\pm\tau} + \mathrm{O}(\varepsilon^2),$$

where

$$\widehat{\Delta}_{\pm\sigma,\pm\tau} = \widehat{\Delta}_{\pm\sigma,\pm\tau}(h) = \sinh(h)\,|b|\,(2K_{2\pi}/h)\left[\widehat{\delta}_{\pm\sigma} - \widehat{\delta}_{\pm\tau}\right],$$

with $\widehat{\delta}_{\pm\sigma} = \pm\widehat{\delta}_\sigma$, ($\sigma \in \{0,1\}$), and $\widehat{\delta}_0 = 1 + \sqrt{m_{2\pi}}$, $\widehat{\delta}_1 = 1 - \sqrt{m_{2\pi}}$.

The conditions $\det(B) \neq 0$, B without multiple eigenvalues (for the quadratic potentials) and $b \neq 0$ (for the linear ones) are the conditions of generic position for L to be a Morse function. The condition B without multiple eigenvalues is equivalent to the complete breakdown of the central symmetry.

The examples of this subsection show that the lower bounds on the number of homoclinic orbits provided by theorem 2.2 are optimal.

3.4 Polynomial perturbations: weakly hyperbolic cases

It is a very well-known fact that the splitting size for analytic area preserving maps in the plane is exponentially small in the hyperbolicity parameter h, for families of maps with degenerate to the identity when $h=0$ [FS90]. Here, $\mathrm{e}^{\pm h}$ stands for the eigenvalues of the differential of the perturbed map on the perturbed *weakly* hyperbolic fixed point. Then, there arises the natural question about whether a similar result holds for analytic and symplectic maps in higher dimensions. We show here some results that lead us to believe that the answer is affirmative.

For the sake of brevity, we restrict ourselves to the case $V(y) = y^\top By$, but the same study can be carried out for any concrete polynomial perturbation. Using that $q_\pi = \mathrm{e}^{-\pi^2/h}$ and the formula $\sqrt{2Km^{1/2}/\pi} = 2\sum_{k\geq 0} q^{(k+1/2)^2}$ [page 479]

$$\Delta_{0,1,i,i}(h) = 16\pi^2 b_i h^{-2}\sinh^2(h)\mathrm{e}^{-\pi^2/h}\left\{\sum_{k\geq 0}\exp[-\pi^2 k(k+1)/h]\right\}^4 .$$

Thus, the homoclinic area between $\mathcal{O}_{0,\pm i}(\varepsilon)$ and $\mathcal{O}_{1,\pm i}(\varepsilon)$ $(i\in\{1,\dots,n\})$, is *a priori* exponentially small in h. A priori means that the first order term in ε is exponentially small in h. Of course, this does not imply that the higher order terms are also exponentially small in h. All the others homoclinic areas are not a priori exponentially small, or are trivially zero because of the odd character of F_ε.

It is important to remark that this is only a partial result: we have assumed that h is small enough, but *fixed*, and $\varepsilon\to 0$. If ε and h tend simultaneously to zero, then one is confronted with the difficult problem of justifying that some errors that seem to be $\mathrm{O}(\varepsilon^2)$ can be neglected in front of the main term that is $\mathrm{O}(\mathrm{e}^{-\pi^2/h})$. Thus, the question is whether some asymptotic formulae like

$$\Delta W\left[\mathcal{O}_{0,\pm i}(\varepsilon),\mathcal{O}_{1,\pm i}(\varepsilon)\right] \sim \varepsilon\Delta_{0,1,i,i}(h) \sim 16\pi^2 b_i\varepsilon\mathrm{e}^{-\pi^2/h},$$

hold, when ε and h tend to zero in any independent way. At the present moment, we do not have an analytical proof of these asymptotic formulae, but, concerning the planar case $(n=1)$, in [DR97a] we have succeeded in proving that the Melnikov method gives the correct asymptotic exponentially small behaviour under a generic assumption on the perturbative potential $V(y)$, for $\varepsilon=\mathrm{O}(h^p)$ and $p>6$. Besides, there is numerical evidence that the hypothesis $\varepsilon=\mathrm{O}(h^p)$, $p>6$, can be improved up to $\varepsilon=\mathrm{o}(1)$ [DR97b]. (It is important to remark here that such numerical experiments require an expensive multiple-precision arithmetic in order to detect the exponentially small size of the splitting.)

Nevertheless, from the computations above, it turns out that the exponentially small splitting can only take place along the direction of the t coordinate over Λ, since a directional derivative of L is exponentially small only in the t direction. (Recall that the differential of L measures the distance between the perturbed invariant manifolds.) This leads us to propose an affirmative answer about the exponentially small character of the splitting of the separatrices, at least in one direction. To give a dynamical interpretation of this distinguished direction, we note that if $h \to 0$ the action of the unperturbed map over Λ tends to a *flow* whose orbits are the coordinate curves $\{a = \text{constant}\}$ of the parametrization $\lambda(t, a)$. It is important to observe that this direction does not depend on the perturbation.

Moreover, the computations above show that the distinguished pairs of homoclinic orbits which give a priori exponentially small splittings are just the *interlaced* pairs, i.e., the pairs created from unperturbed orbits situated on the same coordinate curve $\{a = \text{constant}\}$ (in a interlaced way) of the separatrix Λ.

Finally, we want to stress that a priori exponentially small asymptotic expressions can be computed for the splitting angles in the t-direction over Λ. However, it seems better to work with the homoclinic area since it is an homoclinic invariant, whereas the splitting angles are not.

4 A magnetized spherical pendulum

Finally, as a second example, we focus our attention on Hamiltonian maps that arise from perturbations of a central field. The exact manifold is the same as in the previous example.

4.1 Unperturbed problem

First, we give some well-known definitions and results. Let $T : \mathbf{R}^n \to \mathbf{R}$ be the so-called *kinetic energy* $T(y) = \frac{1}{2}|y|^2$ and let $V : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$ be the *potential energy*. The Hamiltonians $H : \mathbf{R}^{2n} \times \mathbf{R} \to \mathbf{R}$ of the form $H(x, y, t) = T(y) + V(x, t)$, are called *natural*. The Hamiltonian equations can be written as $\ddot{x} = -\partial V(x, t)/\partial x$. Notice that if $V(x, t)$ is even in the spatial variable x, the Hamiltonian map is odd.

When $V(x, t) = V_{\mathrm{c}}(|x|^2)$, for some function $V_{\mathrm{c}} : [0, \infty) \to \mathbf{R}$, the Hamiltonian field is an (autonomous) *central field*, and hence the angular momenta are preserved. Let $\mathcal{A}_0^{n+1} = \{(ra, \dot{r}a) : a \in \mathbf{S}^{n-1}, (r, \dot{r}) \in \mathbf{R}^2\}$ be the manifold of zero angular momenta. Using the central symmetry, we can reduce on $\mathcal{A}_0^{n+1}$ the Hamiltonian system to one degree of freedom: $\ddot{r} = -2V_{\mathrm{c}}'(r^2)r$; that is, if $r(t)$ is a solution of the reduced system, then $\lambda(t, a) = (r(t)a, \dot{r}(t)a)$ is a solution of the original system, for all $a \in \mathbf{S}^{n-1}$.

In [Gru85], one of the first papers on the generalization of the Melnikov method for high-dimensional (continuous) systems, an *(undamped) magnetized spherical pendulum* was considered. It is given by the (autonomous) central field with $V_{\mathrm{c}}(r^2) = (r^4 - r^2)/2$. Obviously, the cases $n > 2$ have no real physical meaning and the cited reference does not deal with them, but the generalization is trivial and it is interesting in order to compare with the section before. The following lemma follows from a straightforward computation on the reduced system $\ddot{r} = r - 2r^3$, i.e., a Duffing equation.

Lemma 4.1 *Let $\Psi_0^t(p)$ be the solution of the Hamiltonian equations of this magnetized spherical pendulum, with initial condition p at $t = 0$. Given $h > 0$, let F_0 be the Hamiltonian map $\Psi_0^h : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$. Then:*

(i) The origin is a hyperbolic fixed point of F_0. Moreover, $\mathrm{Spec}[DF_0(0)] = \{\mathrm{e}^{\pm h}\}$.

(ii) The invariant manifolds of F_0 are completely doubled, giving rise to the separatrix

$$\Lambda = \{(ra, \dot{r}a) : \dot{r}^2 = r^2 - r^4, r \neq 0, a \in \mathbf{S}^{n-1}\}.$$

(iii) The diffeomorphism $\lambda : \mathbf{R} \times \mathbf{S}^{n-1} \to \Lambda$ defined by

$$\lambda(t, a) = (r(t)a, \dot{r}(t)a), \qquad r(t) = \operatorname{sech} t, \tag{4.1}$$

verifies

$$\Psi_0^s(\lambda(t, a)) = \lambda(t + s, a), \qquad \forall t, s \in \mathbf{R}, a \in \mathbf{S}^{n-1}. \tag{4.2}$$

4.2 Perturbed problem

Let us consider a perturbation that preserves the natural character, i.e., the perturbed Hamiltonians are

$$H_\varepsilon(x, y, t) = T(y) + (|x|^4 - |x|^2)/2 + \varepsilon V(x, t/h), \qquad h > 0, \varepsilon \in \mathbf{R},$$

where $V = V(x, \varphi)$ is 1-periodic in φ. We determine V by imposing $V(0, \varphi) \equiv 0$. Small values of h correspond to a rapidly forced pendulum of angular frequency (radians per second) $\omega = 2\pi/h$. We denote by F_ε the Hamiltonian map Ψ_ε^h, where $\Psi_\varepsilon^t(p)$ is the solution of the Hamiltonian equations of H_ε, with initial condition p. (The dependence on the parameter h is omitted to simplify the notation.)

Using equations (2.15), (4.2) and (4.1), the Melnikov potential $L : \mathbf{R} \times \mathbf{S}^{n-1} \to \mathbf{R}$ of the problem turns out to be

$$L(t, a) = -\int_{\mathbf{R}} V(r(t+s)a, s/h)\, \mathrm{d}s = -\int_{\mathbf{R}} V(r(s)a, (s-t)/h)\, \mathrm{d}s, \quad r(s) = \operatorname{sech} s. \tag{4.3}$$

Now, we consider polynomial perturbations, that is, we assume that the Taylor-Fourier expansion of the potential V has a finite number of terms. We write

$$V(x, \varphi) = \sum_{(k,\ell)\in\mathcal{K}} [C_{k,\ell}(x)\cos(2\pi k\varphi) + S_{k,\ell}(x)\sin(2\pi k\varphi)], \tag{4.4}$$

where $\mathcal{K}$ is a finite subset of $\{(k, \ell) \in \mathbf{Z}^2 : k \geq 0, \ell \geq 1\}$ and $C_{k,\ell}$, $S_{k,\ell}$ are homogeneous polynomials of degree ℓ. In this case, the Melnikov potential can be explicitly computed. The result is summarized in the following lemma, whose proof is straightforward.

Lemma 4.2 *Let $P_\ell(\omega)$ ($\ell \geq 0$) be the polynomials generated by the recurrences*

$$P_0(\omega) = 1, \quad P_1(\omega) = \omega, \quad P_{\ell+1}(\omega) = \frac{\omega^2 + \ell^2}{\ell(\ell+1)} P_{\ell-1}(\omega). \tag{4.5}$$

Then, the Melnikov potential (4.3) with V given in (4.4) is

$$L(t, a) = \pi \sum_{(k,\ell)\in\mathcal{K}} \{\operatorname{sech}(\pi k\omega/2) P_{\ell-1}(k\omega)[C_{k,\ell}(a)\cos(k\omega t) - S_{k,\ell}(a)\sin(k\omega t)]\}, \tag{4.6}$$

where $\omega = 2\pi/h$ is the frequency of the perturbation.

A typical difference between the continuous and discrete frames is revealed here: the Melnikov potential (4.6) is an entire periodic function in the complex variable t, whereas the Melnikov potential (3.6) is a doubly periodic one with singularities. Another difference is that a theorem like 3.1 does not hold for the pendulum, since there exist perturbative potentials $V(x, \varphi)$ such that the Melnikov potential (4.6) vanishes identically.

We also notice that $\operatorname{sech}(\pi k\omega/2) = \operatorname{sech}(k\pi^2/h) \sim e^{-\pi^2/h}$, when $h \to 0$. Thus, a discussion on *a priori* exponentially small splittings for this rapidly forced magnetized pendulum, along the lines of the previous section, can be given for any polynomial perturbation. As in the previous section, the exponentially small asymptotic expressions predicted by the Melnikov method are far of being proved for $n > 1$. However, it is well-known that for some perturbations of the rapidly forced planar pendulum [DS92], the Melnikov method gives the right answer.

Finally, we consider the perturbative potential

$$V(x_1, x_2, \varphi) = \frac{2}{\omega^2 + 1} x_2(x_1^2 + x_2^2) \cos(2\pi\varphi),$$

which was already studied in [Gru85]. In that paper, the general (non-Hamiltonian) case is considered, and consequently the symplectic structure is not taken into account, even in the examples where it was possible, like the one above. Using the formula (4.6), we get the Melnikov potential $L(t, a) = \pi \operatorname{sech} \frac{\pi\omega}{2} \sin\vartheta \cos\omega t$, where $a = (\cos\vartheta, \sin\vartheta) \in \mathbf{S}^1$. Its gradient is just the vector-valued Melnikov function used in [Gru85] to measure the splitting. Obviously, it is easier to compute a real-valued function than a vector-valued one. For higher dimensional cases, the saving of work is still bigger.

Appendix: Elliptic functions

A function that plays an important role in the computation of the infinite sums that appear in Melnikov potentials, is a complex function χ satisfying the following properties, where $T, h > 0$ are given parameters:

(C1) χ is meromorphic on $\mathbf{C}$.

(C2) χ is $T\mathrm{i}$-periodic and its derivative is h-periodic.

(C3) The set of poles of χ is $h\mathbf{Z} + T\mathrm{i}\mathbf{Z}$, and all of them are simple and of residue 1.

Remark A.1 Conditions (C1)–(C3) determine a function except for an additive constant: if χ_1 satisfies also (C1)-(C3), $(\chi - \chi_1)'$ is an entire doubly periodic function, and it must be a constant; thus, $\chi(z) - \chi_1(z) = az + b$, but $a = 0$ due to the $T\mathrm{i}$-periodicity.

The function χ can be expressed in terms of Jacobian elliptic functions, Theta functions, or Weierstrassian functions. The Jacobian elliptic functions are well adapted to pencil-and-paper computations, whereas the Theta functions are the best from the numerical point of view, and the Weierstrassian functions are the natural choice for theoretical work on account of their symmetry in the periods. Here, we deal with pencil-and-paper computations, so our choice are the Jacobian elliptic functions.

For a general background on elliptic functions of any kind, we refer to [AS72,WW27]. We follow the notation of the first reference.

Given the *parameter* $m \in [0,1]$, we recall that

$$K = K(m) := \int_0^{\pi/2} (1 - m \sin \vartheta)^{-1/2} \, \mathrm{d}\vartheta, \quad E = E(m) := \int_0^{\pi/2} (1 - m \sin \vartheta)^{1/2} \, \mathrm{d}\vartheta,$$

are the *complete elliptical integrals of the first and second kind* and that

$$E(u) = E(u|m) := \int_0^u \mathrm{dn}^2(v|m) \, \mathrm{d}v,$$

is the *incomplete elliptic integral of the second kind*, where dn is one of the well-known *Jacobian elliptic functions*. Moreover, introducing $K' = K'(m) := K(1-m)$, $E' = E'(m) := E(1-m)$, we also recall that the *nome* q, $|q| < 1$, is defined by $q = q(m) := \mathrm{e}^{-\pi K'/K}$. If any of the numbers m, q, K, K', E, E' or K'/K is given, all the rest are determined. From a numerical point of view, it is better to fix first the nome q, and after compute the rest of parameters and elliptic functions, since the q-series are rapidly convergent.

It is not difficult to check (see [DR96]) that

$$\chi_T(z) = (2K_T/h)^2 (E'_T/K'_T - 1) z + (2K_T/h) E(2K_T z/h + K'_T \mathrm{i} | m_T)$$

verifies (C1)-(C3), where the nome is determined by $q = q_T = \mathrm{e}^{-\pi T/h}$, and m_T, K_T, K'_T, E_T, E'_T are the associated parameters. (The dependence on h is not explicitly written.) Thus,

$$K'_T/K_T = \pi^{-1} \log(1/q_T) = T/h. \tag{A.1}$$

Given an isolated singularity $z_0 \in \mathbf{C}$ of a function f, let us denote $a_{-j}(f, z_0)$ the coefficient of $(z-z_0)^{-j}$ in the Laurent expansion of f around z_0. Obviously, $a_{-j}(f, z_0) = 0$ if z_0 is a pole of f and j is greater than its order.

Proposition A.1 (Summation Formula) *Let f be a function verifying:*

(P1) f is analytic in $\mathbf{R}$ and has only isolated singularities on $\mathbf{C}$.

(P2) f is $T\mathrm{i}$-periodic for some $T > 0$.

(P3) $|f(t)| \le A\mathrm{e}^{-c|\Re t|}$ when $|\Re t| \to \infty$, for some constants $A, c \ge 0$.

Then, $\Sigma(t) := \sum_{k \in \mathbf{Z}} f(t + hk)$ is analytic in $\mathbf{R}$, has only isolated singularities in $\mathbf{C}$, and is doubly periodic with periods h and $T\mathrm{i}$. Moreover, $\Sigma(t)$ can be expressed by the following sum

$$\Sigma(t) = - \sum_{z \in \mathrm{Sing}_T(f)} \mathrm{res}(\chi_T(\cdot - t) f(\cdot), z) = - \sum_{z \in \mathrm{Sing}_T(f)} \sum_{j \ge 0} \frac{a_{-(j+1)}(f,z)}{j!} \chi_T^{(j)}(z - t), \tag{A.2}$$

where $\mathrm{Sing}_T(f)$ is the set of singularities of f in $\mathcal{I}_T = \{z \in \mathbf{C} : 0 < \Im z < T\}$.

Proof. See [Prop. 3.1] □

If f is meromorphic in $\mathbf{C}$, the same happens to Σ, and then Σ is elliptic. From a computational point of view, this is the interesting case, since then (A.2) is a finite sum and can be explicitly computed, as the following lemma, used in section 3, shows.

Lemma A.1 *Let $\Sigma_\ell(t) = \sum_{k\in\mathbf{Z}} f^\ell(t+kh)$, where $f = \mathrm{sech}$. Then:*

$$\begin{aligned}
\Sigma_1(t) &= \left(\frac{2K_{2\pi}}{h}\right)\left[\sqrt{m_{2\pi}}\,\mathrm{cn}\left(\frac{4K_{2\pi}t}{h}\middle| m_{2\pi}\right) + \mathrm{dn}\left(\frac{4K_{2\pi}t}{h}\middle| m_{2\pi}\right)\right],\\
\Sigma_2(t) &= \left(\frac{2K_\pi}{h}\right)^2\left[\frac{E'_\pi}{K'_\pi} - 1 + \mathrm{dn}^2\left(\frac{2K_\pi t}{h}\middle| m_\pi\right)\right].
\end{aligned}$$

Proof. Clearly, $f = \mathrm{sech}$ satisfies properties (P1)-(P3) with $T = 2\pi$. Moreover, the singularities of f in $\mathcal{I}_{2\pi} = \{z \in \mathbf{C} : 0 < \Im z < 2\pi\}$ are simple poles: $\pi\mathrm{i}/2$ and $3\pi\mathrm{i}/2$, with $a_{-1}(f, \pi\mathrm{i}/2) = -a_{-1}(f, 3\pi\mathrm{i}/2) = -\mathrm{i}$. Thus, from (A.2) we get

$$\Sigma_1(t) = \mathrm{i}\left[\chi_{2\pi}(\pi\mathrm{i}/2 - t) - \chi_{2\pi}(3\pi\mathrm{i}/2 - t)\right].$$

From equation (A.1) with $T = 2\pi$, and using that $E(u + 2K'\mathrm{i}) - E(u)$ is a constant, and that $E(-u) = -E(u)$, we have

$$\Sigma_1(t) = \mathrm{i}(2K_{2\pi}/h)[\mathrm{i}(K'_{2\pi} - E'_{2\pi}) - E(v/2 + K'_{2\pi}\mathrm{i}|m_{2\pi}) + E(v/2|m_{2\pi})],$$

where $v = u - K'_{2\pi}\mathrm{i}$ and $u = 4K_{2\pi}t/h$.

In [pages 520 and 508]

$$\begin{aligned}
E(v + K'i) - E(v) &= \mathrm{i}(K' - E') + \mathrm{cn}(v)\,\mathrm{ds}(v),\\
\mathrm{cn}(v/2)\,\mathrm{ds}(v/2) &= \frac{\mathrm{dn}(v) + \mathrm{cn}(v)}{\mathrm{sn}(v)} = \mathrm{ds}(v) + \mathrm{cs}(v).
\end{aligned}$$

Therefore, we arrive at the following expression for Σ_1

$$\Sigma_1(t) = -\mathrm{i}(2K_{2\pi}/h)[\mathrm{ds}(u - K'_{2\pi}\mathrm{i}|m_{2\pi}) + \mathrm{cs}(u - K'_{2\pi}\mathrm{i}|m_{2\pi})],$$

and the formula for Σ_1 follows from $\mathrm{ds}(u - K'\mathrm{i}) = \mathrm{i}\sqrt{m}\,\mathrm{cn}(u)$ and $\mathrm{cs}(u - K'\mathrm{i}) = \mathrm{i}\,\mathrm{dn}(u)$.

The formula for Σ_2 is easier, since $f^2 = \mathrm{sech}^2$ also verifies the properties (P1)-(P3), but with $T = \pi$ instead of $T = 2\pi$. It has only one singularity in $\mathcal{I}_\pi$: $\pi\mathrm{i}/2$. Moreover, $\pi\mathrm{i}/2$ is a double pole with $a_{-1}(f^2, \pi\mathrm{i}/2) = 0$ and $a_{-2}(f^2, \pi\mathrm{i}/2) = -1$. Thus, by (A.2) we get $\Sigma_2(t) = \chi'_\pi(\pi\mathrm{i}/2 - t)$. But $E'(u) = \mathrm{dn}^2(u)$ is an even $2K'\mathrm{i}$-periodic function, so the formula for Σ_2 follows from (A.1) for $T = \pi$. □

Acknowledgements

This work has been partially supported by the EC grant ERBCHRXCT-940460 and the NATO grant CRG950273. Research by Amadeu Delshams is also supported by the Spanish grant DGICYT PB94-0215 and the Catalan grant CIRIT 1996SGR–00105. Research by Rafael Ramírez-Ros is also supported by the U.P.C. grant PR9409.

Both authors wish to express our appreciation to J. Amorós, M.A. Barja and P. Pascual for their help on the topics related to algebraic topology. It is also a pleasure to thank V. Gelfreich, A. Haro, R. de la Llave, J. Ortega, C. Simó, E. Tabacman and D. Treschev for very stimulating discussions and fruitful remarks.

References

[AA89] V.I. Arnold and A. Avez. *Ergodic Problems of Classical Mechanics.* Advanced Book Classics. Addison-Wesley, 1989.

[AM78] R.H. Abraham and J.E. Marsden. *Foundations of Mechanics.* Benjamin/Cummings, Reading, Mass., 1978.

[Ang93] S. Angenent. A variational interpretation of Melnikov's function and exponentially small separatrix splitting. In D. Salamon, editor, *Symplectic geometry*, volume 192 of *London Math. Soc. Lecture Note Ser.*, pages 5–35. Cambridge Univ. Press, 1993.

[Arn76] V.I. Arnold. *Les méthodes mathématiques de la mécanique classique.* Mir, Moscou, 1976.

[AS72] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions.* Dover, 1972.

[BF96] M. Baldomà and E. Fontich. Poincaré-Melnikov theory for n-dimensional diffeomorphisms. Preprint, 1996. To appear in Applicationes Mathematicae.

[BG96] P. Boyland and C. Golé. Lagrangian systems on hyperbolic manifolds. Preprint 1996/1, Institute for Mathematical Sciences, Stony Brook, January 1996.

[BGK95] T.C. Bountis, A. Goriely, and M. Kollmann. A Melnikov vector for n-dimensional mappings. *Phys. Lett. A*, 206:38–48, 1995.

[Bol94] S.V. Bolotin. Homoclinic orbits to invariant tori of Hamiltonian systems. Preprint, 1994. To appear in Adv. Sov. Math.

[BW95] K. Burns and H. Weiss. A geometric criterion for positive topological entropy. *Comm. Math. Phys.*, 172:95–118, 1995.

[DR96] A. Delshams and R. Ramírez-Ros. Poincaré-Melnikov-Arnold method for analytic planar maps. *Nonlinearity*, 9(1):1–26, 1996.

[DR97a] A. Delshams and R. Ramírez-Ros. Exponentially small splitting of separatrices for standard-like maps. Preprint, 1997.

[DR97b] A. Delshams and R. Ramírez-Ros. Melnikov methods and exponentially small splitting of separatrices for planar maps: An experimental study. Preprint in progress, 1997.

[DR97c] A. Delshams and R. Ramírez-Ros. Poincaré-Melnikov-Arnold method for symplectic maps. In Simó [Sim97].

[DRS97] A. Delshams, R. Ramírez-Ros, and T.M. Seara. Splitting of separatrices in Hamiltonian systems and symplectic maps. In Simó [Sim97].

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

[Eas91] R.W. Easton. Transport through chaos. *Nonlinearity*, 4:583–590, 1991.

[FS90] E. Fontich and C. Simó. The splitting of separatrices for analytic diffeomorphisms. *Ergodic Theory Dynamical Systems*, 10:295–318, 1990.

[GG73] M. Golubitsky and V. Guillemin. *Stable Mappings and Their Singularities*, volume 14 of *Graduate Texts in Math.* Springer, New York, 1973.

[GH81] M.J. Greenberg and J.R. Harper. *Algebraic Topology— A First Course*, volume 58 of *Math. Lecture Note Ser.* Addison-Wesley, Reading, Massachussets, 1981.

[Gol94a] C. Golé. Optical Hamiltonians and symplectic twist maps. *Phys. D*, 71:185–195, 1994.

[Gol94b] C. Golé. Periodic orbits for Hamiltonian in cotangent bundles. *Trans. Amer. Math. Soc.*, 343(1):327–347, 1994.

[GPB89] M.L. Glasser, V.G. Papageorgiou, and T.C. Bountis. Melnikov's function for two-dimensional mappings. *SIAM J. Appl. Math.*, 49(3):692–703, 1989.

[Gru85] J. Gruendler. The existence of homoclinic orbits and the method of Melnikov for systems in $\mathbf{R}^n$. *SIAM J. Math. Anal.*, 16(5):907–931, 1985.

[GS77] V. Guillemin and S. Sternberg. *Geometric Asymptotics.* Amer. Math. Soc., Providence, R.J., 1977.

[Hir76] M.W. Hirsch. *Differential Topology*, volume 33 of *Graduate Texts in Math.* Springer-Verlag, New York, Heidelberg, Berlin, 1976.

[Lom93] H.E. Lomelí. Applications of the Melnikov method to twist maps in higher dimensions using the variational approach. Preprint, Univ. of Minnesota, 1993. To appear in Ergodic Theory Dynamical Systems.

[Lom96a] H.E. Lomelí. Perturbations of elliptic billiards. *Phys. D*, 99:59–80, 1996.

[Lom96b] H.E. Lomelí. Saddle connections and heteroclinic orbits for standard maps. *Nonlinearity*, 9:649–668, 1996.

[LT93] P. Levallois and M.B. Tabanov. Séparation des séparatrices du billard elliptique pour une perturbation algébrique et symétrique de l'ellipse. *C. R. Acad. Sci. Paris Sér. I Math.*, 316:589–592, 1993.

[Mat86] J.N. Mather. A criterion for non-existence of invariant circles. *Publ. Math. I.H.E.S.*, 63:153–204, 1986.

[McL94] R.I. McLachlan. Integrable four–dimensional maps of standard type. *Phys. Lett. A*, 177:211–214, 1994.

[McM71] E.M. McMillan. A problem in the stability of periodic systems. In E. Brittin and H. Odabasi, editors, *Topics in modern physics, a tribute to E.V. Condon*, pages 219–244. Colorado Assoc. Univ. Press, Boulder, CO, 1971.

[Mei92] J.D. Meiss. Symplectic maps, variational principles, and transport. *Rev. Modern Phys.*, 64(3):795–848, 1992.

[MMP84] R.S. MacKay, J.D. Meiss, and I.C. Percival. Transport in Hamiltonian systems. *Phys. D*, 13:55–81, 1984.

[PM82] J. Palis, Jr. and W. de Melo. *Geometric Theory of Dynamical Systems.* Springer-Verlag, New York, 1982.

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

[Rom95] V. Rom-Kedar. Secondary homoclinic bifurcations theorems. *Chaos*, 5(2):385–401, 1995.

[RW88] V. Rom-Kedar and S. Wiggins. Transport in two-dimensional maps. *Arch. Rational Mech. Anal.*, 109:239–298, 1988.

[Sim97] C. Simó, editor. *Hamiltonian Systems with Three or More Degrees of Freedom.* NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. Held in S'Agaró, Spain, 19–30 June 1995. Kluwer Acad. Publ., Dordrecht, Holland, to appear in 1997.

[[Sun96] J.H. Sun. Melnikov vector function for high-dimensional maps. *Phys. Lett. A*, 216:47–52, 1996.

[Sur89] Yu B. Suris. Integrable mappings of the standard type. *Functional Anal. Appl.*, 23:74–76, 1989.

[Swi75] R.M. Switzer. *Algebraic Topology— Homotopy and Homology*, volume 212 of *Grundlehren Math. Wiss.* Springer-Verlag, New York, Heidelberg, Berlin, 1975.

[Tab94] M.B. Tabanov. Separatrices splitting for Birkhoff's billiard in a symmetric convex domain, close to an ellipse. *Chaos*, 4(4):595–606, 1994.

[Tab95] E. Tabacman. Variational computation of homoclinic orbits for twist maps. *Phys. D*, 85:548–562, 1995.

[Tre94] D.V. Treschev. Hyperbolic tori and asymptotic surfaces in Hamiltonian systems. *Russ. J. Math. Phys.*, 2(1):93–110, 1994.

[Wei73] A. Weinstein. Lagrangian submanifolds and Hamiltonian systems. *Ann. of Math. (2)*, 98:377–410, 1973.

[Wig91] S. Wiggins. *Introduction to applied nonlinear dynamical systems and chaos*, volume 3 of *Texts Appl. Math.* Springer, New York, 1991.

[WW27] E.T. Whittaker and G.N. Watson. *A course of modern analysis.* Cambridge Univ. Press, 1927.

Internet access:

All the authors' quoted preprints are available at `http://www-ma1.upc.es` in the preprints page, or at `ftp://www-ma1.upc.es`, in the `pub/preprints` directory.

E-mail addresses of the authors: `amadeu@ma1.upc.es`, `rafael@tere.upc.es`.