# Horseshoes in Perturbations of Hamiltonian Systems with Two Degrees of Freedom*

Philip J. Holmes[1] and Jerrold E. Marsden[2]

1 Department of Theoretical and Applied Mechanics, Cornell University, Ithaca, NY 14853, USA
2 Department of Mathematics, University of California, Berkeley, CA 94720, USA

**Abstract.** This paper concerns Hamiltonian and non-Hamiltonian pertur-
bations of integrable two degree of freedom Hamiltonian systems which
contain homoclinic and periodic orbits. Our main example concerns per-
turbations of the uncoupled system consisting of the simple pendulum and the
harmonic oscillator. We show that small coupling perturbations with, possibly,
the addition of positive and negative damping breaks the integrability by
introducing horseshoes into the dynamics.

## 1. Introduction

This paper concerns Hamiltonian and non-Hamiltonian perturbations of inte-
grable two degree of freedom Hamiltonian systems which contain homoclinic and
periodic orbits. Our main example concerns perturbations of the uncoupled
system consisting of the simple pendulum and the harmonic oscillator. We show
that small coupling perturbations with, possibly, the addition of positive and
negative damping breaks the integrability by introducing horseshoes into the
dynamics.

We begin with an unperturbed $n+1$ degree of freedom Hamiltonian in
canonical coordinates $q = (q^1, ..., q^n)$, $p = (p_1, ..., p_n)$, $x$, $y$ of the form

$$H^0(q, p, x, y) = F(q, p) + G(x, y). \tag{1.1}$$

Starting in Sect. 3, we will assume $n = 1$, but for some of the development $n$ can be
arbitrary. Allowing $x$ and $y$ to be multidimensional will be the subject of another
publication.

We shall assume that $G$ admits action-angle variables; i.e. there is a canonical
change of coordinates to $(\theta, I)$ such that $\theta$ is $2\pi$ periodic, $I \geq 0$ and $G$ becomes a
function of $I$ alone; we write $G(I)$ for this function and assume that

$$G(0) = 0, \quad \Omega(I) \equiv G'(I) > 0 \quad \text{for} \quad I > 0. \tag{1.2}$$

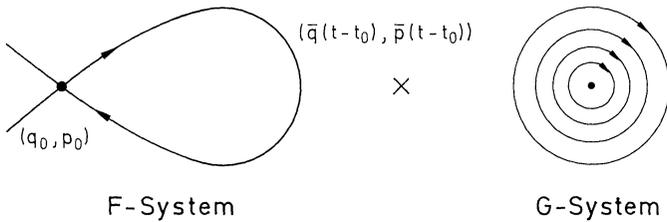F-System                                  G-System

**Fig. 1.** The unperturbed system

Note that (1.2) implies the existence of $G^{-1}$.

The equations of motion are

$$\dot{q}^i = \frac{\partial F}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial F}{\partial q^i}, \qquad i = 1, \ldots, n, \tag{1.3}$$

$$\dot{\theta} = \Omega(I), \qquad \dot{I} = 0. \tag{1.4}$$

We shall assume that the system (1.3) contains a homoclinic orbit $(\bar{q}(t-t_0), \bar{p}(t-t_0))$ joining a saddle point $(q_0, p_0)$ to itself. Of course (1.4) contains the $2\pi$-periodic orbits $\theta(t) = \theta_0 + t\Omega(I_0)$, $I(t) = I_0$. Thus, for the system (1.3)–(1.4), we have orbits which are the products of the homoclinic orbits and the periodic orbits (See Fig. 1.) [The case in which $F$ has a heteroclinic orbit may be treated by similar methods.]

Our principal example in this paper is the pendulum-oscillator Hamiltonian

$$H^0(\phi, v, x, y) = \frac{v^2}{2} - \cos\phi + \frac{1}{2}(y^2 + \omega^2 x^2), \tag{1.5}$$

which takes the form (1.1). Action-angle variables for the oscillator are $x = \sqrt{\frac{2I}{\omega}}\sin\theta$, $y = \omega\sqrt{\frac{2I}{\omega}}\cos\theta$, so that

$$H^0(\phi, v, \theta, I) = F(\phi, v) + G(I),$$

where $\qquad F(\phi, v) = \frac{1}{2}v^2 - \cos\phi,$ $\left.\begin{array}{c} \\ \\ \\ \end{array}\right\}$ $\qquad\qquad\qquad (1.6)$

and $\qquad\quad G(I) = \omega I.$

The Hamiltonian system associated with $F$ possesses the two homoclinic orbits

$$\begin{aligned} \phi(t) &= \pm 2\arctan[\sinh(t-t_0)], \\ v(t) &= \pm 2\,\mathrm{sech}(t-t_0). \end{aligned} \right\} \tag{1.7}$$

We deal with Hamiltonian perturbations of (1.1) in Sects. 2, 3, and 4. We assume that our perturbed Hamiltonian depends on a small parameter $\varepsilon$ in the form

$$H^\varepsilon(q, p, \theta, I) = F(q, p) + G(I) + \varepsilon H^1(q, p, \theta, I) + O(\varepsilon^2), \tag{1.8}$$

where $H^1$ is smooth and $2\pi$ periodic in $\theta$. We shall show that a Poincaré map associated with $H^\varepsilon$ contains Smale horseshoes on each energy surface for $\varepsilon$ small and $H^1$ satisfying certain conditions.

The equations of motion corresponding to $H^\varepsilon$ are

$$\dot{q}^i = \frac{\partial F}{\partial p_i} + \varepsilon \frac{\partial H^1}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial F}{\partial q^i} - \varepsilon \frac{\partial H^1}{\partial q^i}, \qquad i = 1, \ldots, n, \left.\rule{0pt}{40pt}\right\}$$
$$\dot{\theta} = \Omega(I) + \varepsilon \frac{\partial H^1}{\partial I}, \qquad \dot{I} = -\varepsilon \frac{\partial H^1}{\partial \theta}. \tag{1.9}$$

Our method for finding horseshoes involves the Melnikov function technique that has been used in Melnikov (1963), Arnold (1964), Holmes (1979, 1980), Holmes and Marsden (1981a) and Greenspan and Holmes (1981), to show the existence of transverse intersections of stable and unstable manifolds and hence the existence of horseshoes. The Melnikov technique is used after the system has been reduced to a non-autonomous single degree of freedom system (as in Whittaker (1959) Chap. 12, and Birkhoff (1966), Chap. VI, Sect. 3). In particular, in Sect. 4 we prove that the pendulum-oscillator (1.6) develops a horseshoe on each energy surface near the value $H = 1$, when it is perturbed using the coupling term

$$H^1(\phi, v, x, y) = \tfrac{1}{2}(x - \phi)^2. \tag{1.10}$$

Churchill (1980) suggested the possibility of this approach but did not examine any specific examples. Section 5 concerns the more delicate case in which (1.4) is given an additional non-Hamiltonian perturbation. We prove that at least one of the horseshoes persists under this perturbation provided there is a suitable energy transfer mechanism. In Sect. 6 we apply this theory to the pendulum oscillator example once more.

In another paper [Holmes and Marsden (1981b)] we use these methods to address the question of nearly integrable multidegree of freedom systems and Arnold diffusion [cf. Arnold (1964)]. Holmes and Marsden (1981c) treats Hamiltonian systems with symmetry in which (part of) the phase space is the coadjoint orbit of a Lie group. This provides a natural framework in which to consider non-integrable perturbations of rigid bodies.

In many examples of physical interest, such as weakly nonlinear problems, the unperturbed system $H = F(p, q) + G(I)$ does not possess a homoclinic orbit, but some *averaged* system, after truncation, does have homoclinic orbits [cf. McGehee and Meyer (1974)]. In such cases the Melnikov function, computed with the use of second order terms normally neglected in averaging, is typically exponentially small and conclusions on the intersections of manifolds do not immediately follow without a careful study of the errors. The elastic pendulum in the limit of a very stiff rod, with linearized frequency $\omega/\varepsilon$ and Hamiltonian

$$H = \frac{v^2}{2} - \cos\phi + \omega I - \varepsilon \sqrt{\frac{2I}{\omega}} \sin\left(\frac{\theta}{\varepsilon}\right) \cos\phi, \tag{1.11}$$

also falls into this class. The study of such systems is planned for a future publication. The problems of the motion of four point vortices treated by Ziglin (1980), the three wave interaction model, and the motion of charged particles in the earth's magnetic field [see Braun (1981)] possess related difficulties.

We expected that the methods developed here will be applicable to a number of Hamiltonian systems exhibiting complex dynamics. Two examples that seem to

involve homoclinic phenomena are the Hénon-Heiles system [see Churchill, Pecelli, and Rod (1979)] and the mixmaster model in cosmology [Barrow (1981)]. The results in Sect. 5 should also enable one to deal with nearby systems with forcing and dissipative terms.

For other papers in which horseshoes are found in two dimensional mappings by very different techniques, see Devaney and Nitecki (1979) and Tresser, Coullet, and Arneodo (1979). The Melnikov technique in a non-Hamiltonian context is also studied in Chow, Hale and Mallet-Paret (1980).

## 2. The Reduction Method

We now recall how to reduce the $n+1$ degree of freedom system (1.9) to an $n$ degree of freedom non-autonomous system. This is a special case of the general reduction procedure by which a Hamiltonian system with symmetry is reduced to another Hamiltonian system with fewer degrees of freedom. The standard reference is Whittaker (1959, Chap. 12); see also Birkhoff (1966) and Churchill (1980). The case of concern in this paper is the symmetry of time translations, with energy being the corresponding conserved quantity. The procedure is also a special case of that of Marsden and Weinstein (1974) in the context of time-dependent mechanics, as in Abraham and Marsden (1978, Sect. 5.1).

Energy is conserved for (1.9), so we consider the equation

$$H^{\varepsilon}(q, p, \theta, I) = h. \tag{2.1}$$

Now $\dfrac{\partial H^{\varepsilon}}{\partial I} = \Omega(I) + \varepsilon \dfrac{\partial H^1}{\partial I}$. On any compact subset of $(q, p, \theta, I)$ space not containing $I = 0$, we can choose $\varepsilon$ small enough so that $\dfrac{\partial H^{\varepsilon}}{\partial I} > 0$, since $\Omega(I) > 0$ for $I \neq 0$, by assumption. Thus, in such a region, we can solve (2.1) for $I$ to obtain

$$I = L^{\varepsilon}(q, p, \theta, h), \tag{2.2}$$

Now define $L^0$ and $L^1$ by writing

$$L^{\varepsilon}(q, p, \theta, h) = L^0(q, p, h) + \varepsilon L^1(q, p, \theta, h) + O(\varepsilon^2). \tag{2.3}$$

**2.1. Proposition.** *We have*

$$L^0(q, p, h) = G^{-1}(h - F(q, p)), \tag{2.4}$$

*and*

$$L^1(q, p, \theta, h) = -\frac{H^1(q, p, \theta, L^0(q, p, h))}{\Omega(L^0(q, p, h))}. \tag{2.5}$$

*Proof.* Substituting (2.2) into (2.1) gives

$$F(q, p) + G(L^0 + \varepsilon L^1 + O(\varepsilon^2)) + \varepsilon H^1(q, p, \theta, L^0 + \varepsilon L^1 + O(\varepsilon^2)) = h,$$

i.e.

$$[F(q, p) + G(L^0) - h] + \Omega(L^0)(\varepsilon L^1) + \varepsilon H^1(q, p, \theta, L^0) = O(\varepsilon^2).$$

The $\varepsilon^0$ and $\varepsilon^1$ terms of this expression give (2.4) and (2.5).   $\square$

Having passed to the level set $H^{\varepsilon}=h$ and thereby eliminated $I$, we now eliminate the variable conjugate to $H$, namely $t$. (In reduction, an even number of variables is always eliminated.) Since $\Omega(I)>0$ and $H^{\varepsilon}$ is not explicitly $t$-dependent, and $\theta$ is (for small $\varepsilon$) an increasing function of $t$, we can eliminate $t$ by inverting $\theta=\theta(t)$ and expressing $q$ and $p$ as functions of $\theta$. We write $q^{i'}=dq^i/d\theta$ and $p'_i=dp_i/d\theta$, $i=1,\dots,n-1$ so that

$$
q^{i'}=\frac{\dot{q}^i}{\dot{\theta}}=\frac{\partial H^{\varepsilon}}{\partial p_i}\bigg/\frac{\partial H^{\varepsilon}}{\partial I}\,,
$$

and

$$
p'_i=\frac{\dot{p}^i}{\dot{\theta}}=-\frac{\partial H^{\varepsilon}}{\partial q^i}\bigg/\frac{\partial H^{\varepsilon}}{\partial I}\,. \qquad (2.6)
$$

However, implicit differentiation of (2.1) gives

$$
\frac{\partial H^{\varepsilon}}{\partial q^i}+\frac{\partial H^{\varepsilon}}{\partial I}\frac{\partial L^{\varepsilon}}{\partial q^i}=0\,,
$$

and

$$
\frac{\partial H^{\varepsilon}}{\partial p_i}+\frac{\partial H^{\varepsilon}}{\partial I}\frac{\partial L^{\varepsilon}}{\partial p_i}=0\,. \qquad (2.7)
$$

Substituting (2.7) into (2.6) we get

$$
q^{i'}=-\frac{\partial L^{\varepsilon}}{\partial p_i}\,,
$$

and

$$
p'_i=\frac{\partial L^{\varepsilon}}{\partial q^i}\,. \qquad (2.8)
$$

Using (2.3), this becomes

$$
q^{i'}=-\frac{\partial L^0}{\partial p_i}-\varepsilon\frac{\partial L^1}{\partial p_i}+O(\varepsilon^2)\,,
$$

$$
p'_i=\frac{\partial L^0}{\partial q^i}+\varepsilon\frac{\partial L^1}{\partial q^i}+O(\varepsilon^2)\,. \qquad (2.9)
$$

Since $L^0$ depends only on $(q,p)$, but $L^1$ depends on $q$, $p$, and $\theta$, the system (2.9) has the form of a $2\pi$-periodically perturbed $n$ degree of freedom Hamiltonian system. For $n=1$, (2.9) becomes a forced oscillator equation. This is exactly the situation which occurs in our pendulum-oscillator problem.

## 3. Melnikov's Method: The Existence of Horseshoes

For $n=1$, the system (2.9) is in the form analyzed by Melnikov (1963), Holmes (1979, 1980) and Greenspan and Holmes (1981). [For $(\theta, I)$ vectorial or $n\geq 2$, analogous techniques were developed by Arnold (1964) and Holmes and Marsden (1981b) and will be the concern of a subsequent paper.]

For $\varepsilon=0$, the system (2.9) reduces to (1.3) and thus also contains a homoclinic orbit. The Melnikov method involves integration of the Poisson bracket $\{L^0, L^1\}$

around the homoclinic orbit of the unperturbed system. Let us first use (2.4) and (2.5) to express $\{L^0, L^1\}$ in terms of $\{F, H^1\}$.

**3.1. Proposition.** *Holding $\theta$ and $h$ fixed, we have*

$$\{L^0, L^1\} = \frac{1}{[\Omega(L^0)]^2} \{F, H^1\}. \tag{3.1}$$

*Proof.* Using (2.4), we have

$$\frac{\partial L^0}{\partial q^i} = (G^{-1})' \left(-\frac{\partial F}{\partial q^i}\right) \quad \text{and} \quad \frac{\partial L^0}{\partial p_i} = (G^{-1})' \left(-\frac{\partial F}{\partial p_i}\right),$$

while (2.5) gives

$$\frac{\partial L^1}{\partial q^i} = -\frac{1}{\Omega(L^0)}\left(\frac{\partial H^1}{\partial q^i} + \frac{\partial H^1}{\partial I}\frac{\partial L^0}{\partial q^i}\right) + \frac{1}{[\Omega(L^0)]^2}H^1\Omega'\frac{\partial L^0}{\partial q^i},$$

and

$$\frac{\partial L^1}{\partial p_i} = -\frac{1}{\Omega(L^0)}\left(\frac{\partial H^1}{\partial p_i} + \frac{\partial H^1}{\partial I}\frac{\partial L^0}{\partial p_i}\right) + \frac{1}{[\Omega(L^0)]^2}H^1\Omega'\frac{\partial L^0}{\partial p_i}.$$

Thus,

$$\begin{aligned}
\frac{\partial L^0}{\partial q^i}\frac{\partial L^1}{\partial p_i} - \frac{\partial L^0}{\partial p_i}\frac{\partial L^1}{\partial q^i} &= \frac{(G^{-1})}{\Omega(L_0)}\left[\frac{\partial F}{\partial q^i}\left(\frac{\partial H^1}{\partial p_i} + \frac{\partial H^1}{\partial I}\frac{\partial L^0}{\partial p_i} - \frac{1}{\Omega(L^0)}H^1\Omega'\frac{\partial L^0}{\partial p_i}\right) \right.\\
&\qquad \left. -\frac{\partial F}{\partial p_i}\left(\frac{\partial H^1}{\partial q^i} + \frac{\partial H^1}{\partial I}\frac{\partial L^0}{\partial q^i} - \frac{1}{\Omega(L^0)}H^1\Omega'\frac{\partial L^0}{\partial q^i}\right)\right]\\
&= \frac{(G^{-1})'}{\Omega(L^0)}\left[\frac{\partial F}{\partial q^i}\frac{\partial H^1}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial H^1}{\partial q^i}\right].
\end{aligned}$$

But $(G^{-1})' = 1/\Omega(L^0)$ and so we obtain (3.1).  □

The cancellations that occur to yield (3.1) reflect the general fact that the Poisson brackets before and after reduction correspond. Similarly, if $K$ is a function of $(q, p)$ we obtain the formula

$$\{K, L^1\} = -\frac{1}{\Omega}\{K, H^1\} + \frac{1}{\Omega}\frac{\partial}{\partial I}\left(\frac{H^1}{\Omega}\right)\{K, F\}. \tag{3.2}$$

Thus, if $K$ is a first integral for $F$, then (3.2) becomes

$$\{K, L^1\} = -\frac{1}{\Omega}\{K, H^1\}. \tag{3.3}$$

In particular, in the multidegree of freedom case in which all but the first of the $n$ variables $(q, p)$ are in action angle form so that $\{p_k, F\} = 0$, $k = 2, \ldots, n$ then with $K = p_k$, (3.3) becomes

$$\{p_k, L^1\} = -\frac{1}{\Omega}\{p_k, H^1\} = \frac{1}{\Omega}\frac{\partial H^1}{\partial q^k}. \tag{3.4}$$

Relations of this type are useful in the study of perturbations of integrable systems when $n \geq 2$.

In connection with the identities (3.1)–(3.4) the following observation is useful. Along an orbit for the unperturbed system, $F$ is constant, so if $h > F$, $L^0$ will not vanish and $\dfrac{1}{\Omega(L^0)}$ will be a finite constant. Thus, on such an orbit, $\{L^0, L^1\}$ will differ only by a multiplicative constant from $\{F, H^1\}$.

We are now ready to state our main result for Hamiltonian perturbations in case $n = 1$.

**3.2. Theorem.** *Consider a two degree of freedom Hamiltonian system of the form*

$$H^\varepsilon(q, p, \theta, I) = F(q, p) + G(I) + \varepsilon H^1(q, p, \theta, I), \qquad (3.5)$$

*and assume that $F$ contains a homoclinic orbit $(\bar{q}(t - t_0), \bar{p}(t - t_0))$ connecting a hyperbolic saddle to itself (or to another hyperbolic saddle point). Suppose $\Omega(I) = G'(I) > 0$ for $I > 0$. Let $h_1 = F(\bar{q}, \bar{p})$ be the energy of the homoclinic orbit and let $h > h_1$ and $\ell^0 = G^{-1}(h - h_1)$ be constants. Let $\{F, H^1\}(t - t_0)$ denote the Poisson bracket of $F(q, p)$ and $H^1(q, p, \Omega(\ell^0)t, \ell^0)$ evaluated at $\bar{q}(t - t_0)$ and $\bar{p}(t - t_0)$. Define*

$$M(t_0) = \int_{-\infty}^{\infty} \{F, H^1\}(t - t_0)dt, \qquad (3.6)$$

*and assume that $M(t_0)$ has simple zeros. Then for $\varepsilon > 0$ sufficiently small the Hamiltonian system corresponding to (3.5) has a Smale horseshoe in its dynamics on the energy surface $H^\varepsilon = h$.*

This result follows from our previous development (the reduction and Proposition 3.1) and the Melnikov theory given in the references at the beginning of this section. Equation (3.6) can also be obtained from the evolution equation for $F$ along the unperturbed orbit

$$\dot{F} = \{F, H^1\}, \qquad (3.7)$$

cf. Arnold (1964).

For our analysis in Sect. 5 we shall need some facts about the construction of the horseshoe, so we collect them here. First we pick an energy surface $H^\varepsilon = h > h_1$ and consider the Poincaré map $P^{\theta_0}_\varepsilon : \Sigma^{\theta_0} \to \Sigma^{\theta_0}$ (which we just denote $P_\varepsilon$ below) associated with the periodically perturbed system (2.9). Here

$$\Sigma^{\theta_0} = \{(q, p, \theta) | \theta = \theta_0 \in [0, 2\pi]\}$$

is a global cross section for the flow of (2.9). By hypothesis, for $\varepsilon = 0$, $P_\varepsilon$ has an invariant manifold filled with a continuous family of (nontransverse) homoclinic orbits. If $M(t_0)$ has simple zeros then this manifold breaks into a countable set of homoclinic orbits: the generic case found in advanced classical mechanics texts [cf. Arnold and Avez (1967), Abraham and Marsden (1978)]; see Fig. 2a. For more details on homoclinic orbits of maps see Moser (1973) or Newhouse (1980). Here we merely note that the Smale-Birkhoff homoclinic theorem asserts the existence, near any transverse homoclinic point, of a zero dimensional invariant Cantor set $\Lambda$ on which some power of the map, $P^N_\varepsilon$, is homeomorphic to a shift on
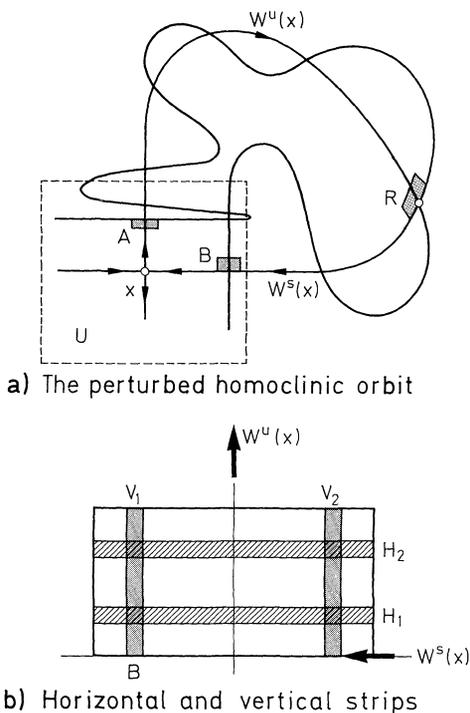
a) The perturbed homoclinic orbit



b) Horizontal and vertical strips

**Fig. 2a and b**

two symbols. Since $P_\varepsilon^N|\Lambda$ possesses a dense orbit, it follows [Moser (1973)] that (3.5) possesses no analytic second integral. Also see Smale (1967).

To construct the horseshoe one takes a small rectangle, $R$, partially bounded by pieces of the stable and unstable manifolds and containing a transverse homoclinic point. Integers $\ell_1, \ell_2$ can be chosen such that the forward and backward images $P^{\ell_1}(R), P_\varepsilon^{-\ell_2}(R)$ lie in a neighbourhood $U$ of the saddle point, $x$. The linearized map $DP_\varepsilon(x)$ can then be used to approximate the motions in $U$ and it can be shown that there are an integer $N < \infty$ and two disjoint "horizontal" strips $H_i \subset P^{\ell_1}(R) = B$ whose images under $P_\varepsilon^N$ are disjoint "vertical" strips $V_i \subset B$ (Fig. 2). The map $P_\varepsilon^N : H_i \to V_i$ is the horseshoe.

To obtain estimates necessary to prove hyperbolicity of $\Lambda$, one needs to find certain sector bundles which are mapped into themselves by $DP_\varepsilon^N$. In our case this implies that the choice of $N$ is related to $\varepsilon$, the perturbation strength, since the angle between the tangent vectors of the manifolds at a (transverse) homoclinic point is $0(\varepsilon)$ ($M(t_0)$ measures the $0(\varepsilon)$ component of the distance between the perturbed manifolds). In Appendix B we show that $N \sim \ln(1/\varepsilon)$. Thus, for each $\varepsilon > 0$ sufficiently small and each $h > h_1$, there is an invariant set $\Lambda_h$ near every transverse homoclinic point in each energy surface $H^\varepsilon = h$ (cf. Fig. 4, below). However as $\varepsilon$ gets smaller, $N$ must be increased. This dependence of $N$ on $\varepsilon$ plays an important role in our discussions of dissipative perturbations on Sect. 5.

## 4. Example: The Coupled Pendulum-Oscillator

We now apply Theorem 3.2 to the Hamiltonian (1.5) with $H^1$ given by (1.10), and the homoclinic orbit for $F$ given by (1.7). In terms of the variables $(\phi, v, \theta, I)$, we have

$$F(\phi, v) = \frac{v^2}{2} - \cos\phi, \tag{4.1}$$

and

$$H^1(\phi, v, \theta, I) = \frac{1}{2}\left(\sqrt{\frac{2I}{\omega}}\sin\theta - \phi\right)^2. \tag{4.2}$$

Thus

$$\{F, H^1\} = -v\phi + \sqrt{\frac{2I}{\omega}}\, v \sin\theta. \tag{4.3}$$

The energy of the homoclinic orbit (1.7) is $h_1 = 1$, so we let $h > 1$ and let $\ell^0 = \frac{1}{\omega}(h-1)$. Thus (3.6) gives

$$M(t_0) = -\int_{-\infty}^{\infty} \{4\,\mathrm{sech}(t-t_0)\arctan[\sinh(t-t_0)]$$

$$\pm \frac{2\sqrt{2(h-1)}}{\omega}\,\mathrm{sech}(t-t_0)\sin(t\omega)\}dt.$$

The first term is odd and so vanishes, leaving

$$M(t_0) = \pm \int_{-\infty}^{\infty} \frac{2\sqrt{2(h-1)}}{\omega}\,\mathrm{sech}(t-t_0)\sin(t\omega)dt. \tag{4.4}$$

This is evaluated by the method of residues as in Holmes (1979), yielding

$$M(t_0) = \pm 2\pi\sqrt{2(h-1)}\,\mathrm{sech}\left(\frac{\pi\omega}{2}\right)\sin\omega t_0. \tag{4.5}$$

Since $M(t_0)$ has simple zeros and is independent of $\varepsilon$ we conclude that, for $\varepsilon > 0$ sufficiently small, the conditions of Theorem 3.2 are satisfied and we have horseshoes in the Poincaré map associated with the pendulum-oscillator on each energy surface $h > 1$, where

$$H^\varepsilon = \frac{v^2}{2} - \cos\phi + \omega I + \frac{\varepsilon}{2}\left(\sqrt{\frac{2I}{\omega}}\sin\theta - \phi\right)^2 = h.$$

Thus, we have proved:

**4.1. Theorem.** *The Hamiltonian system with energy function*

$$H(\phi, v, x, y) = \frac{v^2}{2} - \cos\phi + \frac{1}{2}(y^2 + \omega^2 x^2) + \frac{\varepsilon}{2}(x - \phi)^2$$

*has horseshoes in its dynamics on each energy surface $H > 1$, for $\varepsilon$ sufficiently small, and hence possesses no analytic second integral.*

## 5. Non-Hamiltonian Perturbations

We now wish to consider perturbations under which the total energy $H$ is not conserved. In many physical problems one system produces energy which is subsequently adsorbed by a second system, so that the coupled systems can achieve a "dynamic equilibrium" in which (in a suitably time-averaged sense) energy is preserved. This often manifests itself in the presence of negative damping in one system and positive damping in the other. We will take an integrable Hamiltonian system which possesses continuous families of non-transverse homoclinic orbits, add a Hamiltonian perturbation $H^\varepsilon$, as before, which breaks these manifolds to give transverse homoclinic orbits, and then add dissipative effects which cause a net drift in the energy $H$ for perturbed orbits lying near the homoclinic manifold. Under suitable hypotheses, such dissipative perturbations can leave invariant isolated pieces of the continuous family $\Lambda_h$ of horseshoes discussed in Sect. 3. For simplicity we shall restrict our discussion to two degree of freedom systems ($n = 1$).

The Hamiltonian system (1.9) is modified to include dissipative terms as follows:

$$
\left.
\begin{aligned}
\dot{q} &= \frac{\partial F}{\partial p} + \varepsilon \frac{\partial H^1}{\partial p} + \varepsilon \gamma_1 f_1 \\[4pt]
\dot{p} &= -\frac{\partial F}{\partial q} - \varepsilon \frac{\partial H^1}{\partial q} + \varepsilon \gamma_2 f_2 \\[4pt]
\dot{\theta} &= \Omega(I) + \varepsilon \frac{\partial H^1}{\partial I} + \varepsilon \delta_1 g_1 \\[4pt]
\dot{I} &= -\varepsilon \frac{\partial H^1}{\partial \theta} + \varepsilon \delta_2 g_2,
\end{aligned}
\right\}
\tag{5.1}
$$

where $F$, $\Omega = G'$, and $H^1$ are as in the previous sections and $f_i$, $g_i$ are functions of $(q, p, I, \theta)$, $2\pi$ periodic in $\theta$. Specific hypotheses on $f_i$, $g_i$ and on the dissipation parameters $\gamma_i$, $\delta_i$ will be stated subsequently.

The energy function $H = H^\varepsilon = F(q, p) + G(I) + \varepsilon H^1(q, p, I, \theta)$ is no longer conserved, and our earlier Eq. (2.1) which was used to eliminate $I$ is now replaced by

$$
\begin{aligned}
\dot{H} &= \varepsilon \left( \gamma_1 \frac{\partial H^\varepsilon}{\partial q} f_1 + \gamma_2 \frac{\partial H^\varepsilon}{\partial p} f_2 + \delta_1 \frac{\partial H^\varepsilon}{\partial \theta} g_1 + \delta_2 \frac{\partial H^\varepsilon}{\partial I} g_2 \right) \\[4pt]
&= \varepsilon \left( \gamma_1 \frac{\partial F}{\partial q} f_1 + \gamma_2 \frac{\partial F}{\partial p} f_2 + \Omega(I) \delta_2 g_2 \right) + O(\varepsilon^2) \\[4pt]
&\equiv \varepsilon h(p, q, I) + O(\varepsilon^2).
\end{aligned}
\tag{5.2}
$$

Note that when $\gamma_i = \delta_i = 0$, (5.2) gives $\dot{H} = 0$ and (5.1) becomes (1.9) with conservation of energy.

The five Eq. (5.1)–(5.2) are redundant and we can eliminate the variable $I$ by regarding $I$ as a function of $p$, $q$, $\theta$, and $H$; i.e. by solving $H^\varepsilon = H$ implicitly for $I$ but remembering that $H$ is a variable with its own evolution Eq. (5.2). From (2.2)–(2.5) we have

$$
I = L^0(q, p, H) + \varepsilon L^1(q, p, \theta, H) + O(\varepsilon^2),
\tag{5.3}
$$

where

$$L^0 = F^{-1}(H - F(q, p)),\tag{5.4}$$

and

$$L^1 = \frac{-H^1(q, p, \theta, L^0(q, p, H))}{\Omega(L^0(q, p, H))}.\tag{5.5}$$

As before, we have

$$q' = \dot{q}/\dot{\theta}, \qquad p' = \dot{p}/\dot{\theta},\tag{5.6}$$

and from the implicit equation $H^\varepsilon(q, p, \theta, L^\varepsilon) = H$ we obtain

$$\frac{\partial H^\varepsilon}{\partial p} + \frac{\partial H^\varepsilon}{\partial I}\frac{\partial L^\varepsilon}{\partial p} = 0, \text{ etc.,}$$

so that

$$\dot{q} = \frac{\partial H^\varepsilon}{\partial p} + \varepsilon\gamma_1 f_1$$

becomes

$$
\begin{aligned}
q' &= \left(\frac{\partial H^\varepsilon}{\partial p} + \varepsilon\gamma_1 f_1\right)\Big/\left(\frac{\partial H^\varepsilon}{\partial I} + \varepsilon\delta_1 g_1\right)\\
&= \frac{1}{\Omega}\left(\frac{\partial F}{\partial p} + \varepsilon\left(\frac{\partial H^1}{\partial p} + \gamma_1 f_1\right)\right)\Big/\left(\Omega(I) + \varepsilon\left(\frac{\partial H^1}{\partial I} + \delta_1 g_1\right)\right)\\
&= -\frac{\partial L^0}{\partial p} - \varepsilon\left(\frac{\partial L^1}{\partial p} - \frac{\gamma_1 f_1}{\Omega} - \frac{\partial L^0}{\partial p}\frac{\delta_1 g_1}{\Omega}\right) + O(\varepsilon^2).
\end{aligned}
$$

A similar computation for $p'$ and use of (5.2) yields the three dimensional system

$$
\left.
\begin{aligned}
q' &= -\frac{\partial L^0}{\partial p} - \varepsilon\left(\frac{\partial L^1}{\partial p} - \frac{\gamma_1 f_1}{\Omega} - \frac{\partial L^0}{\partial p}\frac{\delta_1 g_1}{\Omega}\right) + O(\varepsilon^2),\\
p' &= \frac{\partial L^0}{\partial q} + \varepsilon\left(\frac{\partial L^1}{\partial q} - \frac{\gamma_2 f_2}{\Omega} - \frac{\partial L^0}{\partial q}\frac{\delta_1 g_1}{\Omega}\right) + O(\varepsilon^2),\\
H' &= \varepsilon\Omega\left(\delta_2 g_2 - \frac{\partial L^0}{\partial q}\gamma_1 f_1 - \frac{\partial L^0}{\partial p}\gamma_2 f_2\right) + O(\varepsilon^2).
\end{aligned}
\right\}\tag{5.7a, b, c}
$$

Equations (5.7a, b, c) constitute the system we now study, with the dependent variables $q$, $p$, $H$ and the independent time-like variable $\theta$. For $\gamma_i = \delta_i = 0$, (5.7) reduce to (2.9), as expected.

To deal conveniently with the slow variable $H$ compared with the fast variables $(q, p)$, we use a slight modification of the usual averaging theorem in which the $O(\varepsilon)$ term in the right hand side of (5.7c) is replaced by its $\theta$-average.

$$\varepsilon\bar{h} = \varepsilon\Omega\left(\delta_2\bar{g}_2 - \frac{\partial L^0}{\partial q}\gamma_1\bar{f}_1 - \frac{\partial L^0}{\partial p}\gamma_2\bar{f}_2\right),\tag{5.8}$$

Fig. 3
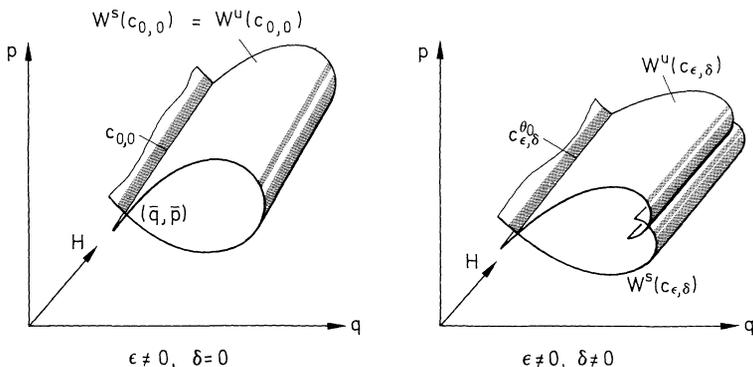
where an overbar denotes the time average $\dfrac{1}{2\pi}\displaystyle\int_0^{2\pi}(\cdot)d\theta$. The modified averaging procedure is discussed in Appendix A. Its use is not essential, but it makes calculations somewhat easier. Retaining the same notation for the averaged variables, our reduced system is now (5.7a, b) plus

$$H' = \varepsilon\bar{h}(q,p,H) + O(\varepsilon^2) . \tag{5.9}$$

Let $\delta = (\gamma_1, \gamma_2, \delta_1, \delta_2)$ denote the dissipation coefficients in our system. Also, let $P_{\varepsilon,\delta} = P_{\varepsilon,\delta}^{\theta_0}$ denote the Poincaré map associated with the system (5.7a, b) and (5.9). Thus, $P_{\varepsilon,\delta}^{\theta_0}$ maps (an open subset of) $\mathbb{R}^3$ to $\mathbb{R}^3$ and is given by advancing the independent variable $\theta$ by $2\pi$, with starting value $\theta_0$.

Let us assume that the homoclinic orbit $(\bar{q}(\theta), \bar{p}(\theta))$ for $F$ has energy $= h_1$ and joins a hyperbolic saddle point $(\bar{q}, \bar{p})$ to itself. (The case of a heteroclinic orbit is similar.) Thus, for each value of $H > h_1$, $(\bar{q}, \bar{p}, H)$ is a fixed point for $P_{0,0}^{\theta_0}$. Let $c_{0,0}$ denote this curve of fixed points. Since we are assuming $(\bar{q}, \bar{p})$ is a hyperbolic fixed point for $F$, $c_{0,0}$ is a hyperbolic invariant manifold for $P_{0,0}^{\theta_0}$, with $H$ restricted to an interval, say $h_1 < H_0 \leqq H \leqq H_1$. Since hyperbolic manifolds are preserved under perturbation, we have:

**5.1. Lemma.** *For $\delta$ bounded and $\varepsilon$ sufficiently small, there are invariant curves $c_{\varepsilon,\delta}^{\theta_0}$ close to $c_{0,0}$ for $P_{\varepsilon,\delta}^{\theta_0}$. Moreover, the stable and unstable manifolds of $c_{\varepsilon,\delta}^{\theta_0}$, denoted $W^s(c_{\varepsilon,\delta}^{\theta_0})$ and $W^u(c_{\varepsilon,\delta}^{\theta_0})$ are "$C^r$ close" to those of $c_{0,0}$ and each is two dimensional. (See Fig. 3).*

Notice that if $\delta = 0$ then $c_{\varepsilon,0}$ is still a curve of fixed points by conservation of energy and we recover the situation of Sect. 3. For $\delta \neq 0$, points on $c_{\varepsilon,\delta}^{\theta_0}$ can "drift" under iteration of $P_{\varepsilon,\delta}^{\theta_0}$ since energy is not conserved. However, they will stay on $c_{\varepsilon,\delta}^{\theta_0}$ (until $H$ leaves the interval $[H_0, H_1]$).

For $\delta = 0$, suppose Theorem 3.1 is used to show that $W^s(c_{\varepsilon,0}^{\theta_0})$ intersects $W^u(c_{\varepsilon,0}^{\theta_0})$ transversely. This persists for $\delta$ sufficiently small, by the stability of transversal intersections under perturbation. Because of potential drift in the $H$ variable, this alone does not permit us to conclude the existence of horseshoes for $\delta \neq 0$. Rather, we must control $H$. The crucial hypothesis that enables this to be done will be given next.

Let $N$ be an integer fixed so that $(P_{\varepsilon,0}^{\theta_0})^N$ has a horseshoe, as described following Theorem 3.2. Thus, $N$ is large enough so that $(P_{\varepsilon,0}^{\theta_0})^N$ maps two horizontal strips in $B$ back around to two vertical strips in $B$ again, where $B$ is a rectangle lying in a neighborhood $U$ of the saddlepoint (Fig. 2).

Let

$$\Delta H = \varepsilon \int_{-\pi N}^{\pi N} \frac{\bar{h}}{\Omega} d\theta = \varepsilon \int_{-\pi N/\Omega}^{\pi N/\Omega} \bar{h} dt, \tag{5.10}$$

where $\bar{h}$ is given by (5.8) and $\bar{h}, \Omega$ are evaluated on the homoclinic orbit $(\bar{q}(\theta), \bar{p}(\theta))$ at an energy value $H$. (Recall that $\Omega$ is constant on this orbit). From (5.9) we see that $\Delta H$ represents the approximate change in energy in following a point starting near the homoclinic orbit for $N$ iterates; i.e. for a total $\theta$-time $2\pi N$. For $N$ large but finite $P_{\varepsilon,\delta}$ maps points in the horizontal strips in rectangle $B$ in Fig. 2 back to the vertical strips in rectangle $B$ after $N$ iterates. Thus $\Delta H$ represents the leading term in the energy change while going from $B$ back to $B$. Of course $\Delta H$ is a function of $H$ and depends on $\varepsilon$, $\delta$ and $N$. Strictly speaking, (5.10) should be evaluated on trajectories just *inside* the homoclinic orbit, but as we show in Appendix B, as $\varepsilon \to 0$, $N \to \infty$ and the horizontal and vertical strips, $H_i, V_i \subset B$ must be taken closer and closer to the homoclinic orbit. Thus, since we only need $\Delta H$ to leading order in what follows, evaluation on the homoclinic orbit is sufficient.

Now we state our basic energy-transfer condition:

**Condition (H).** *Assume there is a value $H_c > 1$ of $H$ at which $\Delta H$ given by (5.10) changes sign transversely, i.e.*

$$\Delta H < 0, \quad if \quad H < H_c,$$
$$\Delta H > 0, \quad if \quad H > H_c,$$

*and*

$$\frac{d}{dH}(\Delta H)\bigg|_{H=H_c} \neq 0.$$

Under this condition and $\varepsilon$ sufficiently small, we have

**5.2. Lemma.** *There is a smooth function $H(q,p)$ defined on the rectangle $B$ such that if $(q,p) \in B$ then $(P_{\varepsilon,\delta}^{\theta_0})^N(q,p,H(q,p))$ has the form $(\tilde{q}, \tilde{p}, H(\tilde{q}, \tilde{p}))$.*

*Proof.* Since the exact energy change differs from (5.10) by $0(\varepsilon^2)$, persistence of transversality guarantees that condition $(H)$ is also true for the exact energy change for $\varepsilon$ sufficiently small. For $\delta = 0$, the surface $H = \text{constant} = H_c$ is preserved by $(P_{\varepsilon,0}^{\theta_0})^N$. For $\delta \neq 0$ the surface $H_c$ is preserved and contracting to first order in $\varepsilon$. By persistence of hyperbolic invariant manifolds, there is a nearby surface exactly invariant under $(P_{\varepsilon,\delta}^{\theta_0})^N$; this surface is the graph of $H$. $\quad\square$

Thus, we have identified a surface, say $\Sigma_c$ near $H = H_c$ such that $(P_{\varepsilon,\delta}^{\theta_0})^N$ maps $\Sigma_c$ to $\Sigma_c$.

Now we wish to show that there is a horseshoe in this surface $\Sigma_c$. By the arguments in Holmes and Marsden (1981a), Appendix A), we must check that $W^s(c_{\varepsilon,\delta}^{\theta_0})$ and $W^u(c_{\varepsilon,\delta}^{\theta_0})$ continue to intersect transversally, for $\delta \neq 0$. To do this, we form the Melnikov function at energy value $H_c$ for the system (5.7) and (5.8).

If $M(t_0)$ is given by (3.6), then, using Proposition 3.1, the Melnikov function for (5.7a) and (5.7b) is given by

$$M_\delta(t_0) = \frac{1}{\Omega^2} M(t_0) + \int_{-\infty}^{\infty} \frac{1}{\Omega}\left(\frac{\partial L^0}{\partial p}\gamma_2 f_2 - \frac{\partial L^0}{\partial q}\gamma_1 f_1\right) dt.$$

(Note that the $\delta_1 g_1$ term cancels out.) Thus, using (5.4), we get

**5.3. Lemma.**

$$M_\delta(t_0) = \frac{1}{\Omega^2}\left[M(t_0) + \int_{-\infty}^{\infty}\left(\frac{\partial F}{\partial q}\gamma_1 f_1 - \frac{\partial F}{\partial p}\gamma_2 f_2\right)(t - t_0) dt\right]. \qquad (5.11)$$

We will assume that $\delta$ is chosen such that $M(t_0)$ continues to have simple zeros.

In Appendix B we discuss the relationship between $N$ and $\varepsilon$. As $\varepsilon$ gets smaller, the number of iterates required of the Poincaré map to guarantee a horseshoe gets larger. It is shown that:

**5.4. Lemma.** *There are constants $\alpha$, $\beta$ and an integer $L$ such that*

$$N = L + \alpha \ln\left(\frac{\beta}{\varepsilon M(H_c)}\right), \qquad (5.12)$$

*where $M(H_c)$ is the supremum of $M(t_0)$ over $t_0$.*

This result applies to the case of Hamiltonian perturbations ($\delta = 0$). When $\delta \neq 0$ there is an analogous result $N = N(\varepsilon, \delta)$ in which $\alpha = \alpha(\delta)$ and $M(H_c)$ is replaced by $M_\delta(H_c)$. However, in our application we set $\delta = \varepsilon^\mu \bar{\delta}$ to be $O(\varepsilon^\mu)$, so that $\varepsilon\delta = \varepsilon^{1+\mu}\bar{\delta} \ll \varepsilon$ for $\varepsilon \ll 1$ and the effects of $\delta$ in (5.12) can be ignored, cf. Appendix B.

From (5.11) we obtain a condition on the size of $\gamma_1$ and $\gamma_2$; if $M(t_0)$ has simple zeros and oscillates with an amplitude $M(H_c)$, then (5.11) gives us conditions of the form

$$c_1\gamma_1 < M(H_c),$$
$$c_2\gamma_2 < M(H_c), \qquad (5.13)$$

which must be satisfied for $M_\delta$ to still have simple zeros.

We summarize our findings as follows:

**5.5. Theorem.** *Suppose that $N$, $H_c$, and $\delta$ can be chosen so that conditions (H), (5.12), and (5.13) all hold. Then the iterate $(P_{\varepsilon,\delta}^{\theta_0})^N$ of the Poincaré map of the reduced system (5.7) has, for $\varepsilon$ sufficiently small, a horseshoe in its dynamics; the horseshoe lies near the homoclinic orbit in the $(q, p)$ variables and near the (non-invariant) energy surface $H = H_c$.*

# 6. Example: The Oscillator-Pendulum with Positive and Negative Damping

We now wish to show that for $\varepsilon$ sufficiently small the pendulum-oscillator system considered in Sect. 4,

$$\dot{\phi} = v, \quad \dot{v} = -\sin\phi + \varepsilon(x - \phi),$$
$$\dot{x} = y, \quad \dot{y} = -\omega^2 x + \varepsilon(\phi - x),$$

continues to have a horseshoe when dissipation is included. Specifically, we add negative damping $(-\delta)$ to the oscillator so that it drives the pendulum, which now has positive damping $(\gamma)$:

$$
\left.\begin{aligned}
\dot{\phi} &= v, \\
\dot{v} &= -\sin\phi + \varepsilon(x-\phi) - \varepsilon\gamma v, \\
\dot{x} &= y, \\
\dot{y} &= -\omega^2 x + \varepsilon(\phi-x) + \varepsilon\delta y.
\end{aligned}\right\}
\tag{6.1}
$$

In action-angle variables, (6.1) becomes

$$
\left.\begin{aligned}
\dot{\phi} &= v, \\
\dot{v} &= -\sin\phi + \varepsilon\left(\sqrt{\frac{2I}{\omega}}\sin\theta - \phi\right) - \varepsilon\gamma v, \\
\dot{\theta} &= \omega + \varepsilon\left(\sqrt{\frac{2I}{\omega}}\sin\theta - \phi\right)\frac{\sin\theta}{\sqrt{2I\omega}} - \varepsilon\delta\sin\theta\cos\theta, \\
\dot{I} &= -\varepsilon\left(\sqrt{\frac{2I}{\omega}}\sin\theta - \phi\right)\sqrt{2I\omega}\cos\theta + \varepsilon\delta 2I\cos^2\theta.
\end{aligned}\right\}
\tag{6.2}
$$

Note that $\delta < 0$ represents damping while $\delta > 0$ represents negative damping (energy production). The energy evolution equation (5.2) is

$$
\dot{H} = \varepsilon(\delta y^2 - \gamma v^2),
$$

or

$$
H' = \frac{\varepsilon}{\omega}(2\delta\omega I\cos^2\theta - \gamma v^2) + O(\varepsilon^2),
\tag{6.3}
$$

where the dependent variable is now $\theta$. Using

$$
I = L^0(\phi, v, H) = \frac{1}{\omega}(H - (v^2/2 - \cos\phi)),
\tag{6.4}
$$

from (2.4), (6.3) becomes

$$
H' = \frac{\varepsilon}{\omega}[\delta(H + \cos\phi - v^2/2)(1 + \cos 2\theta) - \gamma v^2] + O(\varepsilon^2),
\tag{6.5}
$$

Although we do not need them explicitly in the calculations to follow, we also give the reduced evolution equations for $\phi, v$:

$$
\phi' = \frac{1}{\omega}\left[v + \frac{\varepsilon}{\omega}\left(v\left(\frac{\phi\sin\theta}{\omega\sqrt{2A}} - \frac{\sin^2\theta}{\omega} + \delta\sin\theta\cos\theta\right)\right)\right] + O(\varepsilon^2),
$$

$$
v' = \frac{1}{\omega}\left[-\sin\phi + \frac{\varepsilon}{\omega}\left(\sin\phi\left(\frac{\sin^2\theta}{\omega} - \frac{\phi\sin\theta}{\omega\sqrt{2A}} - \gamma\sin\theta\cos\theta\right)\right.\right.
$$

$$
\left.\left. + (\sqrt{2A}\sin\theta - \omega\phi - \gamma v)\right)\right] + O(\varepsilon^2),
\tag{6.6a, b}
$$

where $A = H + \cos\phi - v^2/2$. Equations (6.6a, b) and (6.5) correspond to (5.7a, b, c).

We now average (6.5). The transformation $(\phi, v, H) \rightarrow (\phi, v, \bar{H})$ is given by

$$(\phi, v, H) = (\phi, v, \bar{H} + \varepsilon u(\phi, v, \bar{H}, \theta)),$$

where

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial \phi}\left(\frac{v}{\omega}\right) + \frac{\partial u}{\partial v}\left(-\frac{\sin\phi}{\omega}\right) = \frac{\delta}{\omega}(H + \cos\phi - v^2/2)\cos 2\theta, \tag{6.7}$$

from (A.8). This is satisfied if we take

$$u = \frac{\delta}{2\omega}(H + \cos\phi - v^2/2)\sin 2\theta, \tag{6.8}$$

and then (6.5) becomes, dropping the overbars,

$$H' = \frac{\varepsilon}{\omega}[\delta(H + \cos\phi - v^2/2) - \gamma v^2] + O(\varepsilon^2). \tag{6.9}$$

We now check condition $(H)$. Inserting the unperturbed solution (1.7) we have

$$H + \cos\phi(t) - v(t)^2/2 \equiv H - 1$$

and (6.9) becomes

$$H' = \frac{\varepsilon}{\omega}\left[\delta(H - 1) - 4\gamma \operatorname{sech}^2\left(\frac{\theta}{\omega}\right)\right] + O(\varepsilon^2),$$

or

$$\dot{H} = \varepsilon[\delta(H - 1) - 4\gamma \operatorname{sech}^2 t] + O(\varepsilon^2). \tag{6.10}$$

Hence

$$\Delta H = \varepsilon \int_{-\pi N/\omega}^{\pi N/\omega} [\delta(H - 1) - 4\gamma \operatorname{sech}^2 t]dt,$$

$$= \varepsilon\left[\frac{2\pi N\delta}{\omega}(H - 1) - 8\gamma \tanh\left(\frac{\pi N}{\omega}\right)\right], \tag{6.11}$$

and so condition (H) is satisfied if we have

$$H_c = 1 + \frac{8\omega \tanh(\pi N/\omega)}{2\pi N}\left(\frac{\gamma}{\delta}\right) \approx \frac{8\omega}{2\pi N}\frac{\gamma}{\delta}(\text{ for } N \text{ large}). \tag{6.12}$$

We next compute the Melnikov function $M_{(\delta, \gamma)}(t_0)$ from (5.11). From Sect. 4 and (3.1) we have

$$M(t_0) = \pm \frac{1}{\omega^2} 2\pi \sqrt{2(H - 1)} \operatorname{sech}\left(\frac{\pi\omega}{2}\right)\sin\omega t_0. \tag{6.13}$$

Using $F = v^2/2 - \cos\phi$, $\gamma_1 f_1 \equiv 0$, and $\gamma_2 f_2 = -\gamma v$ in (5.11), the second term of $M_{(\delta, \gamma)}(t_0)$ is

$$-\int_{-\infty}^{\infty} \frac{1}{\omega^2} v(-\gamma v)dt = \frac{\gamma}{\omega^2}\int_{-\infty}^{\infty} (2\operatorname{sech}(t - t_0))^2 dt = \frac{8\gamma}{\omega^2}.$$

Thus we obtain

$$M_{(\delta,\gamma)}(t_0) = \pm \frac{1}{\omega^2} \left[ 2\pi \sqrt{2(H-1)} \, \text{sech} \left( \frac{\pi\omega}{2} \right) \sin \omega t_0 + 8\gamma \right]$$

$$\overset{\text{def}}{=} M(H) \sin \omega t_0 + 8\gamma/\omega^2. \tag{6.14}$$

To complete our verification of the hypotheses of Theorem 5.5, we note that, for $M_{(\delta,\gamma)}(t_0)$ to have simple zeros, we require

$$M(H_c) > 8\gamma/\omega^2,$$

or

$$H_c - 1 > \frac{8\gamma^2}{\pi^2 \, \text{sech}^2(\pi\omega/2)} \overset{\text{def}}{=} C_1 \gamma^2; \tag{6.15}$$

while for satisfaction of condition $(H)$ we have

$$H_c - 1 \approx \frac{8\omega\gamma}{2\pi N\delta} \overset{\text{def}}{=} C_2 \frac{\gamma}{N\delta}. \tag{6.16}$$

We also have the relationship

$$N = N(\varepsilon) = L + \alpha \ln(\beta/\varepsilon \sqrt{H_c - 1}) \tag{6.17}$$

from (5.12), where the remaining constants in $M_{(\delta,\gamma)}(t_0)$ are accumulated into $\beta$. From (6.16) and (6.17) we have

$$(H_c - 1) \left[ L + \alpha \ln(\beta/\varepsilon \sqrt{H_c - 1}) \right] = C_2 \gamma/\delta. \tag{6.18}$$

To satisfy (6.15) and (6.18) simultaneously we pick $\gamma$ and $\delta$ to be of the same order in $\varepsilon$, say $\gamma = \varepsilon^\mu \bar{\gamma}$, $\delta = \varepsilon^\mu \bar{\delta}$; $\mu > 0$. Then (6.18) may be rewritten as

$$(H_c - 1) \left[ 1 + C_3 \ln(\beta/\varepsilon \sqrt{H_c - 1}) \right] = C_4. \tag{6.19}$$

A simple exercise in calculus shows that, for $\varepsilon$ small, (6.19) has a unique solution $H_c$ near $H_c = 1$ and for any fixed $\sigma$, $0 < \sigma < 2$,

$$H_c - 1 > \varepsilon^\sigma, \tag{6.20}$$

for small $\varepsilon$; "how small" depends on how close $\sigma$ is to 0.

*Proof.* Rewrite (6.19) as

$$\varepsilon \sqrt{x} = \beta e^{1/C_3} e^{-C_4/C_3} x \overset{\text{def}}{=} \phi(x),$$

where $x = H_c - 1$. By considering the graphs of $y = \varepsilon \sqrt{x}$ and $y = \phi(x)$ one sees that for small $\varepsilon$, (6.19) has a unique small solution $x(\varepsilon) \to 0$ as $\varepsilon \to 0$. Moreover, as $\phi$ vanishes to all orders at $x = 0$, $\phi(x) \leq x^p$ for $p \geq 1/2$ and $x$ small. It follows that $x(\varepsilon)$ is larger than the solution of $\varepsilon \sqrt{x} = x^p$ for small $\varepsilon$; i.e. $x(\varepsilon) \geq \varepsilon^{2/(2p-1)}$ for small $\varepsilon$. Let $\sigma = 2/(2p-1)$. $\square$

Picking $\sigma < 2\mu$, (6.15) is now satisfied, since

$$H_c - 1 > c_\sigma \varepsilon^\sigma > C_1 \varepsilon^{2\mu} \bar{\gamma}^2 \tag{6.21}$$

for $\varepsilon$ sufficiently small. For example, we can take $\mu = 1/2$, since then the damping perturbations $\varepsilon(\varepsilon^\mu\bar\delta)$, $\varepsilon(\varepsilon^\mu\bar\gamma)$ appear at $O(\varepsilon^{3/2})$ and the $O(\varepsilon^2)$ terms ignored in our computations do not affect the results. Thus, we have proved:

**6.1. Theorem.** *The system* (6.1) *has a horseshoe in its dynamics provided we choose* $\varepsilon\gamma = \varepsilon^{3/2}\bar\gamma$ *and* $\varepsilon\delta = \varepsilon^{3/2}\bar\delta$ *and* $\varepsilon$ *sufficiently small.*

Of course it is possible to vary the orders of $\delta$ and $\gamma$ with some latitude and still maintain the hypotheses. Specifically, 6.1 remains valid if we choose $\gamma = \varepsilon^\nu\bar\gamma$ and $\delta = \varepsilon^\mu\bar\delta$, $0 \le \mu \le \nu < 1$, $\nu \ne 0$.

## Conclusions

In this paper we have developed applicable techniques for establishing the existence of chaotic dynamics in the sense of the presence of a horseshoe for both Hamiltonian and non-Hamiltonian perturbations of systems with two degrees of freedom containing homoclinic orbits and periodic orbits.

While horseshoes are not strange attractors, they are often visible and behave like them in numerical experiments (perhaps due to small background noise); cf. Franks (1981). Arnold diffusion is a higher dimensional manifestation of the same phenomenon and is certainly seen in many examples [see, for instance Lieberman (1980)].

For conservative perturbations the method is a straightforward combination of a classical reduction scheme with a method of Melnikov. For non-conservative perturbations a delicate energy balance argument is needed to ensure that at least one horseshoe survives near the energy balance point. Near other points there is a "ghost horseshoe" which decays because of energy drift. If the dissipation terms all contribute to energy loss then, while no invariant set remains near the homoclinic orbit (since $H$ decreases on all orbits), the manifolds $W^s(c^{\theta_0}_{\varepsilon,\delta})W^u(c^{\theta_0}_{\varepsilon,\delta})$ continue to intersect and the resulting ghost horseshoes would give rise to complicated dynamics on *finite* time intervals, as orbits move through the energy band.

The results are shown to apply to typical perturbations of the pendulum-oscillator system, thereby showing that this classical example has complex dynamics and, in particular, is non-integrable.

### Appendix A. A Modified Averaging Theorem

For the basic averaging theorem see Hale (1969). Here we consider a system of the form

$$\begin{aligned}\dot x &= f(x) + \varepsilon g(x, y, t),\\ \dot y &= \varepsilon h(x, y, t), \quad 0 \le \varepsilon \ll 1,\end{aligned} \tag{A.1}$$

where $x = x(t)$ and $y = y(t)$ are the fast and slow variables, the functions $f$, $g$, $h$ are sufficiently smooth and the latter two are $T$-periodic in $t$.

**A1. Proposition.** *There exists a near identity time dependent change of coordinates* $(x, y) \rightarrow (x, z)$ *under which* (A.1) *becomes*

$$\dot{x} = f(x) + \varepsilon g(x, z, t) + O(\varepsilon^2)$$
$$\dot{z} = \varepsilon \bar{h}(x, z) + O(\varepsilon^2), \qquad (A.2)$$

*where* $\bar{h} = \dfrac{1}{T} \int\limits_0^T h(x, y, t) dt$ *is the t-average of h.*

*Proof.* As in the usual averaging theorem, we set

$$y = z + \varepsilon u(x, z, t). \qquad (A.3)$$

Differentiating (A.3) with respect to time, we obtain

$$\dot{y} = \dot{z} + \varepsilon \dot{u} + \varepsilon D_x u \dot{x} + \varepsilon D_z u \dot{z}, \quad \text{where} \quad (\dot{\phantom{)}}) = \partial/\partial t. \qquad (A.4)$$

Using (A.1), (A.2), and (A.4):

$$(\mathrm{Id} - \varepsilon D_z u)\dot{z} = \dot{y} - \varepsilon \dot{u} - \varepsilon D_x u \dot{x} = \varepsilon [h(x, z + \varepsilon u, t)$$
$$- \varepsilon \dot{u}(x, z + \varepsilon u, t) - D_x u f(x)]. \qquad (A.5)$$

We can write

$$h(x, z, t) = \bar{h}(x, z) + \tilde{h}(x, z, t), \qquad (A.6)$$

where $\tilde{h}$ is $T$-periodic in $t$ and has zero mean. From (A.5)–(A.6) we have

$$(\mathrm{Id} - \varepsilon D_z u)\dot{z} = \varepsilon [\bar{h}(x, z) + \tilde{h}(x, z, t) - \dot{u}(x, z, t) - D_x u(x, z, t) f(x)] + O(\varepsilon^2). \quad (A.7)$$

Thus, if we set

$$\frac{\partial u}{\partial t} + D_x u f(x) = \tilde{h}(x, z, t), \qquad (A.8)$$

we have, from (A.7)

$$\dot{z} = \varepsilon \bar{h}(x, z) + O(\varepsilon^2), \qquad (A.9)$$

and, using (A.3) in (A.1)

$$\dot{x} = f(x) + \varepsilon g(x, z, t) + O(\varepsilon^2). \qquad (A.10)$$

It remains only to check that the linear partial differential Eq. (A.8) admits a solution $u = u(x, z, t)$. However, (A.8) has the solution

$$u(x(t), z, t) = u(x, z, 0) + \int\limits_0^t \tilde{h}(x(s), z, s) ds, \qquad (A.11)$$

where $x(t)$ satisfies $x(0) = x$ and $\dot{x} = f(x)$. $\quad\square$

Finally, the hyperbolicity noted in Sect. 5, together with the usual Gronwall estimates, implies that solutions of (A.2) near the homoclinic manifolds remain within $O(\varepsilon)$ of those of (A.1) for times of $O(1/\varepsilon)$. Since we wish only to integrate for times of $O(N) = O(\ln(1/\varepsilon))$ [Eqs. (5.10) and (5.12)] the averaged equation may be used in computations. Note that the transformation $u(x, z, t)$ is not in general

$T$-periodic, since solutions of (A.8) depend upon the (nonperiodic) flow $x(t)$ of $\dot{x} = f(x)$. One must therefore be careful in inferring the existence of $T$-periodic solutions of (A.1) corresponding to fixed points of (A.2), as in the usual averaging theorem.

## Appendix B. The Iteration Number: Proof of Lemma 5.4

In this appendix we derive a relationship between $N$, the number of iterates of the Poincaré map $P^{\theta_0}_{\varepsilon,0}$ necessary to guarantee that $F = (P^{\theta_0}_{\varepsilon,0})^N$ has a horseshoe, and $\varepsilon$, the perturbation parameter. Henceforth we drop the sub- and superscripts on $P^{\theta_0}_{\varepsilon,0}$.

Let $x$ be the (perturbed) saddle point of $P$ and $y$ a transverse homoclinic point lying outside a ball $B_\mu(x)$ of radius $\mu$ about $x$. The Melnikov theory tells us that the maximum distance between the manifolds near $y$ is

$$d_{\max} = \varepsilon K_1 M(H_c) + O(\varepsilon^2),\qquad\qquad (B.1)$$

where $M(H_c) = \sup_{t_0 \in [0, T)} M(t_0)$ and $K_1$ is a constant.

We next need a basic result from dynamical systems theory, the "lambda-lemma" [Palis (1969), Newhouse (1980)], which enables us to make our choices of horizontal and vertical strips in the horseshoe map more precise:

**B.1. Lemma.** *Let $x$ be a hyperbolic saddle point of a $C^r$ diffeomorphism $P$ and $D \subset W^u(x)$ an open disc in its unstable manifold. Let $\Delta$ be a disc of $\dim(W^u(x))$ meeting $W^s(x)$ transversely at a point $y$. Then $\bigcup_{n \geq 0} P^n(\Delta)$ contains discs arbitrarily $C^r$ close to $D$.*

This result implies that, if $y \in W^u(x) \cap W^s(x)$ is a transverse homoclinic point, then $W^u(x)$ and $W^s(x)$ accumulate on themselves, giving us the structure of Fig. B.1. We assume that the map is orientation preserving, as are the Poincaré maps occurring in the application of this paper.

We pick a rectangle $R$ bounded by pieces of $W^s(x)$, $W^u(x)$ as shown. Since $\mu$ is fixed independent of $\varepsilon$, there are fixed integers $L_1$, $L_2$ such that $P^{L_1}(R)$, $P^{-L_2}(R) \subset B_\mu(x)$ and the "height" and "width" of $P^{L_1}(R)$ and $P^{-L_2}(R)$ are $\varepsilon K_3 M(H_c)$, $\varepsilon K_4 M(H_c)$ respectively. Once in $B_\mu$ the dynamics is dominated by the linearized map, which, working in suitable coordinates, we can take to be

$$DP = \begin{pmatrix} \lambda & 0 \\ 0 & \gamma \end{pmatrix}, \quad |\lambda| < 1 < |\gamma|.$$

To obtain the horseshoe structure $P^{-(L_2 + N_2)}(R) \cap P^{L_1 + N_1}(R)$ as shown, we require further iteration numbers $N_i = N_i(\varepsilon)$ such that

$$\gamma^{N_1} \varepsilon K_3 M(H_c) \geq K_5 \mu \gamma^{-N_2}, \quad \lambda^{-N_2} \varepsilon K_4 M(H_c) \geq K_6 \mu \lambda^{N_1}.$$

A sufficient condition is thus

$$N_1 + N_2 = \alpha \ln(\beta / \varepsilon M(H_c)),\qquad\qquad (B.2)$$

where

$$\alpha = \max\{1/\ln(\gamma), 1/\ln(\lambda^{-1})\}, \quad \beta = \max\{\mu K_5 / K_3, \mu K_6 / K_4\}.$$

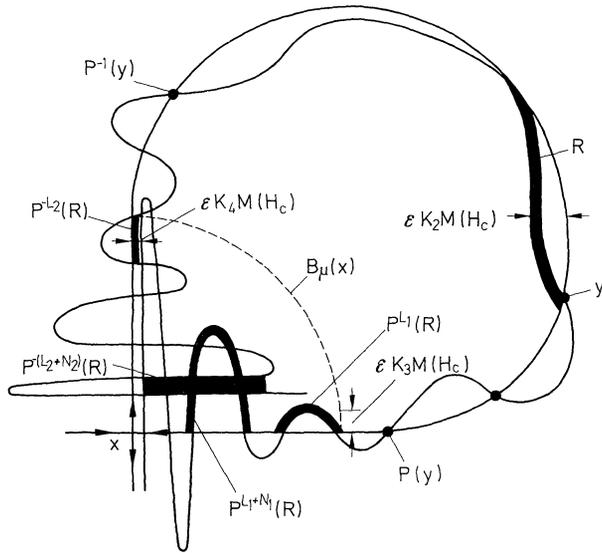**Fig. B1.** The iteration number

Thus the total number of iterates is $N = L_1 + L_2 + N_1 + N_2$, or

$$N = N(\varepsilon) = L + \alpha \ln(\beta/\varepsilon M(H_c)), \tag{B.3}$$

where $L$ is a fixed integer and $\alpha, \beta$ are constants.

When $\delta \neq 0$ $M(H_c)$ should be replaced by $M_\delta(H_c)$ and $\lambda, \gamma$ by $\lambda + K_7\delta, \gamma + K_8\delta$, leading to

$$N_i(\varepsilon, \delta) = \alpha(\delta) \ln(\beta/\varepsilon M_\delta(H_c)), \tag{B.4}$$

where

$$\alpha(\delta) = \max\{1/\ln(\gamma + K_8\delta), 1/\ln((\lambda + K_7\delta)^{-1})\} = \alpha + K_9\delta + O(\delta^2). \tag{B.5}$$

and $K_9$ is a positive constant. Thus $N(\varepsilon, \delta) > N(\varepsilon)$ in general. However, in our application we take $\delta$ of order $\varepsilon^\mu$, $\mu > 0$ (for example $\varepsilon^{1/2}$) and thus $\varepsilon\delta = e^{1 + \mu}\overline{\delta}$, say, and the dependence of $N$ on $\delta$ is weaker than its dependence on $\varepsilon$, and hence can effectively be ignored in the limit $\varepsilon \to 0$.

## References

Abraham, R., Marsden, J.: Foundations of mechanics. 2nd edn. Reading, Mass.: Addision Wesley 1978
Arnold, V.I.: Sov. Math. Dokl. **156**, 9–12 (1964)
    **156**, 9–12 (1964)
Arnold, V.I., Avez, A.: Théorie ergodique des systèmes dynamiques. Paris: Gauthier-Villars (1967)
    (English edition: Benjamin-Cummings, Reading Mass., 1968)
Barrow, J.: Chaotic behavior in general relativity (preprint) (1981)

Birkhoff, G.D.: Dynamical systems. Colloq. Publ. IX, 2nd ed. Providence, R.I.: Am. Math. Soc., 1966
Braun, M.: SIAM Rev. **23**, 61–93 (1981)
Chow, S.-N., Hale, J., Mallet-Paret, J.: J. Diff. Eq'ns. **37**, 351–373 (1980)
Churchill, R.C.: On proving the nonintegrability of a Hamiltonian system (preprint, Hunter College,
    C.U.N.Y.) (1980)
Churchill, R.C., Pecelli, G., Rod, D.L.: A survey of the Hénon-Heiles Hamiltonian with applications to
    related examples. In: Lecture Notes in Physics, Vol. 93. Berlin, Heidelberg, New York: Springer
    1979
Devaney, R., Nitecki, Z.: Commun. Math. Phys. **67**, 137–148 (1979)
Franks, J.: On the Hénon attractor (in preparation)
Greenspan, B., Holmes, P.J.: Homoclinic orbits, subharmonics and global bifurcations in forced
    oscillations. In: Non-linear dynamics and turbulence. Barenblatt, G., Iooss, G., Joseph, D.D. (eds.).
    Pitman 1981 (to appear)
Hale, J.K.: Ordinary differential equations. New York: Wiley 1969
Holmes, P.: Phil. Trans. Roy. Soc. A**292**, 419–448 (1979)
Holmes, P.: SIAM J. Appl. Math. **38**, 65–80; **40**, 167–168 (1980)
Holmes, P.J., Marsden, J.E.: Arch. Rat. Mech. Anal. **76**, 135–167 (1981a)
Holmes, P.J., Marsden, J.E.: Melnikov's method and Arnold diffusion for perturbations of integrable
    Hamiltonian systems. (1981b) (to appear in J. Math. Phys.)
Holmes, P.J., Marsden, J.E.: Horseshoes and Arnold diffusion for Hamiltonian systems on Lie groups.
    (Preprint) (1981c)
Lieberman, M.A.: Ann. N.Y. Acad. Sci. **357**, 119–142 (1980)
Marsden, J., Weinstein, A.: Rep. Math. Phys. **5**, 121–130 (1974)
McGehee, R., Meyer, K.: Homoclinic points of area preserving diffeomorphisms. Am. J. Math. **96**, 409–
    421 (1974)
Melnikov, V.K.: Trans. Mosc. Math. Soc. **12**, 1–57 (1963)
Moser, J.: Stable and random motions in dynamical systems, with special emphasis on celestial
    mechanics. Ann. Math. Studies, No. 77. Princeton, N.J.: Princeton University Press 1973
Newhouse, S.: Topology **12**, 9–18 (1974)
Newhouse, S.: Lectures on dynamical systems. In: Dynamical systems, Moser, J. (ed.), pp. 1–114.
    Boston: Birkhäuser 1980
Palis, J.: Topology **8**, 385–405 (1969)
Smale, S.: Bull. Am. Math. Soc. (1967)
Tresser, C., Coullet, P., Arneodo, A.: Topological horseshoe and numerically observed chaotic
    behavior in the Hénon mapping (preprint, C.N.R.S., Université de Nice, France) (1979)
Whittaker, E.T.: A treatise on the analytical dynamics of particles and rigid bodies, 4th edn.
    Cambridge: Cambridge University Press 1959
Ziglin, S.L.: Sov. Math. Dokl. **21**, 296–299 (1980)