Chaos in a Model of the Forced and Damped Sine-Gordon Equation

Thesis by

Gregor Kovačič

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

June, 1990

(submitted September 26, 1989)

Acknowledgements

I would first like to thank my advisor, Professor Steve Wiggins, for guiding me through this work. I would also like to thank Professors Philip Holmes and David McLaughlin, and fellow students Dana Hobson and Tasso Kaper who gave me valuable advice about various aspects of the project. To Miriam Herrera and Tasso Kaper I express my gratitude for helping me with editing and proofreading the text of this thesis. I owe much of my present outlook on mathematics to the late Professor Paco Lagerstrom. In the two years I knew him before his sudden death he taught me many deep truths about the subject. And last but not least, I would like to thank my parents for their constant love and support.

I acknowledge the support from Caltech in the form of teaching and research assistantships.

Abstract

We analytically determine two of the mechanisms which cause chaotic dynamics to appear in a model of the forced and damped Sine-Gordon equation. In particular, we find orbits homoclinic to periodic orbits, and orbits homoclinic to fixed points which satisfy conditions sufficient to guarantee the existence of nearby chaotic invariant sets. One of these homoclinic orbits is a so-called Šilnikov-type loop. A proof the existence of a symmetric pair of such loops is our main result. This proof consists of a modified Melnikov perturbation analysis, augmented by some techniques from the field of geometric singular perturbation theory.

Table of Contents

List of Figures

CHAPTER 1

INTRODUCTION

1.1. Goal of This Thesis. The aim of this work is to analytically determine the mechanisms which cause chaotic dynamics to appear in a model of the forced and damped Sine-Gordon equation. In particular, we find orbits homoclinic to periodic orbits, and orbits homoclinic to fixed points which satisfy conditions sufficient to guarantee the existence of nearby chaotic invariant sets. One of these homoclinic orbits is a so called Šilnikov-type loop. A proof of its existence is our main result. We had to modify the standard Melnikov perturbation analysis for this problem, and to augment it with some techniques from the field of geometric singular perturbation theory to create a new, powerful technique which we hope to use on many similar problems.

The motivation for this work came through papers of Bishop, Ercolani, Forest, McLaughlin, Overman, and others, on a particular type of chaos occurring in the forced and damped Sine-Gordon equation and their attempts to model it by a two degree of freedom system of ordinary differential equations. They performed numerical experiments on both the Sine-Gordon equation and the model, and found similar behavior in the two systems. (See Bishop, Forest, McLaughlin, and Overman [1986], Bishop, Forest, McLaughlin [1988], Bishop, Flesch, Forest, McLaughlin, and Overman [1989], and references therein.) In particular, using nonlinear numerical

spectral analysis, they determined that most of the chaotic dynamics in the Sine-Gordon equation take place on a low-dimensional submanifold, spanned by only a few nonlinear modes. This suggested the creation of a model with only two linear modes which we describe in this work.

To better understand the role of the model, we now briefly describe the numerical experiment that indicated the existence of chaos in the Sine-Gordon equation. All the numerical studies have been performed on the equation

$$u_{tt} - u_{xx} + \sin u = \epsilon[-\alpha u_t + \Gamma \cos \omega t]\,, \qquad (1.1)_\epsilon$$

with periodic boundary conditions

$$u(x = -\frac{L}{2}, t) = u(x = \frac{L}{2}, t)$$

for all t and with even spatial symmetry

$$u(x,t) = u(-x,t)$$

for all t. The authors chose a fixed small value of $\epsilon\alpha$, and fixed ω and L and in addition assumed that ω is close to, but less than, one; *i.e.*, it is $\omega = 1 - \epsilon\tilde{\omega}$ with both ϵ and $\tilde{\omega}$ being positive. The initial condition they chose was a single hump Sine-Gordon "breather" solution, and they observed its evolution in space and time as they increased the forcing $\epsilon\Gamma$.

The complete bifurcation sequence is described, for instance, in Bishop, Flesch, Forest, McLaughlin, and Overman [1989]. Here we only present its main point,

which is that above a certain value of $\epsilon\Gamma$, chaotic jumping of the solution occurs between two "breathers," one peaked in the middle and the other one at the ends of the interval $[-\frac{L}{2}, \frac{L}{2}]$, with the solution passing near a spatially flat state at every jump. Comparing the situation with the unperturbed Sine-Gordon equation, they found that the latter possesses linearly unstable spatially flat states connected to themselves by homoclinic orbits: the two types of "breathers" with a spatial hump structure which exhibit chaotic behavior in the perturbed problem. (See Ercolani, Forest, McLaughlin [1989].) From this they inferred that a Melnikov-type analysis should be appropriate in order to see how these unperturbed structures are distorted under perturbation to give rise to chaos. As a preliminary step in their analysis of the perturbed Sine-Gordon equation, they chose to develop a simple model capturing as much of the unperturbed and perturbed structure of the Sine-Gordon equation as possible.

1.2. The Model. We now derive the model following Bishop, Flesch, Forest, McLaughlin, and Overman [1989], and then discuss its relationship with the Sine-Gordon equation. As a preliminary step we recall that ω in $(1.1)_\epsilon$ is assumed to be of the form $\omega = 1 - \epsilon\tilde{\omega}$. This permits us to derive a nonlinear Schroedinger envelope equation for the quantity $B(X,T)$ defined by

$$u_\epsilon(x,t) = 2\sqrt{\epsilon\tilde{\omega}}[B(X,T)e^{i\omega t} + B^*(X,T)e^{-i\omega t}] + \mathcal{O}(\epsilon)\,,$$

where $u_\epsilon(x,t)$ is a solution of $(1.1)_\epsilon$ with $X = \sqrt{2\epsilon\tilde{\omega}}x, T = \epsilon\tilde{\omega}t$. The resulting

envelope equation is

$$-iB_T + B_{XX} + (|B|^2 - 1)B = i\tilde{\alpha}B + \tilde{\Gamma}\,, \tag{1.2}$$

with $\tilde{\alpha} = \frac{\alpha}{2\tilde{\omega}}$ and $\tilde{\Gamma} = \frac{\Gamma}{8\epsilon^{\frac{1}{2}}\tilde{\omega}^{\frac{3}{2}}}$. It is defined on the X interval $[-\frac{L_X}{2}, \frac{L_X}{2}]$ where $L_X = \sqrt{2\epsilon\tilde{\omega}}L$.

Note that (1.2) is still integrable when $\tilde{\alpha}$ and $\tilde{\Gamma}$ are zero, and that we have eliminated the periodic dependence on t from the forcing. The whole approximation is somewhat similar to a crude form of averaging, since steady solutions of (1.2) correspond to periodic solutions of $(1.1)_\epsilon$, and periodic solutions of (1.2) correspond to quasiperiodic solutions of $(1.1)_\epsilon$ in case the period of the solution of (1.2) is rationally incommensurable with ω. Chaotic solutions are, of course, chaotic in both equations.

Following the numerical experiments of Bishop, *et.al.*, we are interested in small $\tilde{\alpha}$ and $\tilde{\Gamma}$. Hence, we rename the variables $\tilde{\alpha}$ and $\tilde{\Gamma}$ as $\epsilon\alpha$ and $\sqrt{2}\epsilon\Gamma$, respectively, where ϵ is assumed to be small. We also rename the variables X and T into x and t, respectively.

In correspondence with what has been observed in the Sine-Gordon equation we assume a two mode approximation

$$B(x,t) = \frac{1}{\sqrt{2}}c(t) + b(t)\cos kx + \cdots\,,$$

with $k = \frac{2\pi}{L_x}$. After inserting into (1.2) and neglecting the higher Fourier modes this approximation gives the following equations for the two complex amplitudes

$c = c_1 + ic_2$ and $b = b_1 + ib_2$:

$$-i\dot{c} + (\tfrac{1}{2}|c|^2 + \tfrac{1}{2}|b|^2 - 1)c + \tfrac{1}{2}(cb^* + bc^*)b = i\epsilon\alpha c + i\epsilon\Gamma \qquad (1.3a)_\epsilon$$

$$-i\dot{b} + (\tfrac{1}{2}|c|^2 + \tfrac{3}{4}|b|^2 - (1 + k^2))b + \tfrac{1}{2}(cb^* + bc^*)c = i\epsilon\alpha b\,. \qquad (1.3b)_\epsilon$$

The equations $(1.3)_\epsilon$ are the model we will analyze. In Chapter 2 we establish the following properties of the unperturbed equations, which are $(1.3)_\epsilon$ with $\epsilon = 0$ and are refered to by $(1.3)_0$:

(1) The equations $(1.3)_0$ constitute a completely integrable Hamiltonian system, *i.e.*, the equations possess two independent constants of motion whose mutual Poisson bracket is zero;

(2) The phase space of $(1.3)_0$ possesses a circle of fixed points $c = \sqrt{2}e^{i\gamma}$, $b = 0$, with $0 \leq \gamma \leq 2\pi$, which is connected to itself by two symmetric two-dimensional surfaces of heteroclinic orbits. These orbits connect pairs of fixed points on the circle.

The circle of fixed points and its heteroclinic orbits correspond to the flat state and the "breathers" in the Sine-Gordon equation, respectively. Of course, there are serious differences in the phase spaces of these equations. (See Bishop, Flesch, Forest, McLaughlin, and Overman [1989].) But the above analogies in the geometry of the homoclinic structures are sufficient to prompt research on the model. In fact, we are convinced that the methods we have developed for the perturbed model will, because of these common geometric features, need to be only slightly modified

in order to yield a similar result for the nonlinear Schroedinger equation (1.2). The modifications will consist mostly of extending finite-dimensional results about invariant manifolds and stable fibers (see Chapters 3 and 4) to the case of infinite dimensions and of taking care of the neutral modes.

1.3. Mechanisms for Chaos in the Model. The results we have found indicate two different mechanisms for chaos in the model. The first mechanism is orbits homoclinic to certain nonresonant periodic orbits in the case when $\alpha = 0$. Similar results have been independently found by Ercolani, McLaughlin, and Forest [1989] who used a different Melnikov-type method from the one we use to obtain the same conclusion.

The second mechanism, which is the main result of this work, is the Šilnikov orbit connecting a fixed point to itself. This fixed point is one of the two survivors from the above mentioned circle of fixed points, which under perturbation develops into a resonance band. We show that the Šilnikov orbit does not exist in the phase space of the equations $(1.3)_\epsilon$ as they stand. The orbit does, however, exist in the phase space of the modified system

$$-i\dot{c} + (\tfrac{1}{2}|c|^2 + \tfrac{1}{2}|b|^2 - 1)c + \tfrac{1}{2}(cb^* + bc^*)b = i\epsilon\alpha c + i\epsilon\Gamma \qquad (1.4a)_\epsilon$$

$$-i\dot{b} + (\tfrac{1}{2}|c|^2 + \tfrac{3}{4}|b|^2 - (1 + k^2))b + \tfrac{1}{2}(cb^* + bc^*)c = i\epsilon(1 + \lambda)\alpha b \qquad (1.4b)_\epsilon$$

with any positive λ. The source of this modification can be traced back to the Nonlinear Schroedinger and Sine-Gordon equations, which have to be altered so

that they include viscous damping. In their modified versions they are

$$u_{tt} - u_{xx} + \sin u = \epsilon[-\alpha u_t - \Lambda u_{txx} + \Gamma \cos \omega t] \qquad (1.5)_\epsilon$$

and

$$-iB_T + B_{XX} + (|B|^2 - 1)B = i\tilde{\alpha}B + i\tilde{\Lambda}B_{XX} + i\tilde{\Gamma}\,. \qquad (1.6)$$

The relationship between the parameters λ, $\tilde{\Lambda}$ and Λ is given by the equations $\tilde{\Lambda} = \epsilon\Lambda$ and $\epsilon\lambda = \frac{\tilde{\Lambda}k^2}{\alpha}$. The effect of the added terms is that damping is stronger for higher modes than for the lower ones. We see immediately that our nonexistence result was to be expected, since the original partial differential equations and the model were the borderline case between viscous damping and viscous forcing, the latter being unphysical.

As we have mentioned above, the Šilnikov loop in our modified model is homoclinic to a fixed point inside a resonance band, which is the leftover of the circle of fixed point after the perturbation has been applied. Methods for dealing with this delicate resonant structure are a combination of the Melnikov technique and geometric singular perturbation theory, which do not appear to have been used in this conjunction before. We hope to use these methods on other problems, founding our hopes on the fact that resonance phenomena are a common occurrence in systems with many degrees of freedom. A peculiarity of our method is that it is used to describe orbits homoclinic to a structure which is born out of the perturbation, yet we maximally use the unperturbed solutions as well.

At present, we speculate that the same method can be modified and used for certain partial differential equations such as the previously-mentioned nonlinear Schroedinger equation (1.2). In any case, we believe that the investigation of orbits homoclinic to resonance bands promises many new and interesting results.

CHAPTER 2

PHASE SPACE OF THE UNPERTURBED SYSTEM

2.1. Preliminaries. The goal of this section is to completely describe the geometry of the phase space of the unperturbed system. In the course of this description, we will identify the invariant sets possessing orbits homoclinic to themselves since these are the mechanisms most likely to cause chaos when the system is perturbed. We present explicit solutions for these homoclinic orbits in this chapter because they will be used as the starting point in the perturbation analysis we will develop in Chapters 3 and 4.

The unperturbed equations are obtained from the equations $(1.3)_\epsilon$ for the two complex amplitudes c and b by setting $\epsilon = 0$. They are

$$-i\dot{c} + \left(\tfrac{1}{2}|c|^2 + \tfrac{1}{2}|b|^2 - 1\right) c + \tfrac{1}{2}\left(cb^* + bc^*\right) b = 0 \qquad (2.1a)$$

$$-i\dot{b} + \left(\tfrac{1}{2}|c|^2 + \tfrac{3}{4}|b|^2 - \left(1 + k^2\right)\right) b + \tfrac{1}{2}\left(cb^* + c^*b\right) c = 0\,. \qquad (2.1b)$$

This system possesses two constants of motion, the Hamiltonian

$$H = \tfrac{1}{8}|c|^4 + \tfrac{1}{2}|b|^2|c|^2 + \tfrac{3}{16}|b|^4 - \tfrac{1}{2}\left(1 + k^2\right)|b|^2 - \tfrac{1}{2}|c|^2 + \tfrac{1}{8}\left(b^2c^{*2} + b^{*2}c^2\right) \qquad (2.2a)$$

and

$$I = \tfrac{1}{2}(|c|^2 + |b|^2)\,. \qquad (2.2b)$$

The system (2.1) can be written in the Hamiltonian form

$$\dot{c} = \left(\frac{\partial H}{\partial c_2} - i \frac{\partial H}{\partial c_1} \right)$$
$$\dot{b} = \left(\frac{\partial H}{\partial b_2} - i \frac{\partial H}{\partial b_1} \right) ,$$

where we define $c = c_1 + ic_2$ and $b = b_1 + ib_2$.

Before analyzing equations (2.1) in detail, we identify the three important symmetries they possess. By inspection, it is easy to see that (2.1) is invariant under the transformations

$$(c, b) \mapsto (-c, b) , \tag{2.3a}$$

$$(c, b) \mapsto (c, -b), \tag{2.3b}$$

and

$$(c, b) \mapsto (c, b)e^{i\chi} \tag{2.3c}$$

where χ is any real number. The first two symmetries imply that the planes $\Pi_c = \{(b, c) \,|\, b = 0\}$ and $\Pi_b = \{(b, c) \,|\, c = 0\}$ are both invariant under the flow generated by (2.1). The third symmetry is responsible for the constant of motion I which foliates the planes Π_c and Π_b into circular periodic orbits with constant frequencies. We will give a detailed derivation of these facts and describe two more invariant surfaces, also foliated by circular periodic orbits, below. First, however, we rewrite equations (2.1) in a simpler form.

2.2. The Symplectic Reduction. The analysis of equation (2.1) is greatly facilitated by finding a coordinate system in which the form (2.1) assumes will be

as simple as possible. The fact that we are dealing with a Hamiltonian system with an additional constant of motion plays a significant role in the search for the appropriate coordinate system since there exists a standard way to simplify such systems. The standard procedure consists of making a symplectic coordinate change such that the extra constant and its conjugate variable, which we call the conjugate angle, are among the new coordinates. This symplectic coordinate change effectively decouples these two new variables from the rest: The angle is not present in any of the equations, because, as a variable canonically conjugate to a conserved quantity, it does not enter the Hamiltonian; and, in addition, the constant of motion acts as a parameter in the system for the remaining variables. Theoretically, this symplectic coordinate change can always be made when the system possesses a second constant of motion. This fact is proved, for instance, in Olver [1986]. In practice, however, no recipe exists to find the conjugate angle and a suitable change of the other coordinates. Thus one must, in general, resort to trial and error in order to find the appropriate transformation.

For our system, we will follow Zufiria [1988] and take the argument of one of the two complex variables in (2.1) as the angle conjugate to the constant I. We have two choices for the form of the transformation, and we will have to use both in order to avoid singularities.

For most of this work we will use the coordinate system obtained by taking the argument of the complex variable c to be the angle conjugate to the constant

I. Hence, our first transformation is:

$$c = |c|e^{i\gamma} \tag{2.4a}$$

$$b = (x + iy)e^{i\gamma} . \tag{2.4b}$$

We then take I, γ, x and y as the new variables. The proof that this transformation is canonical is tedious, and we relegate it to Appendix B. The Hamiltonian in these variables is

$$H = \tfrac{1}{2}I^2 - I - \tfrac{7}{16}x^4 - \tfrac{3}{8}x^2y^2 + \tfrac{1}{16}y^4 + (I - \tfrac{1}{2}k^2)x^2 - \tfrac{1}{2}k^2y^2 . \tag{2.5}$$

Equations (2.1) transform to

$$\dot{x} = -k^2y - \tfrac{3}{4}x^2y + \tfrac{1}{4}y^3 \tag{2.6a}$$

$$\dot{y} = \left(k^2 - 2I\right)x + \tfrac{7}{4}x^3 + \tfrac{3}{4}xy^2 \tag{2.6b}$$

$$\dot{I} = 0 \tag{2.6c}$$

$$\dot{\gamma} = 1 - I - x^2 . \tag{2.6d}$$

From (2.2b) we see that (2.6) is only valid in the interior of the paraboloid $x^2 + y^2 < 2I$. The boundary of this paraboloid corresponds to the plane Π_b. The transformation is singular there.

We now analyze the structure of system (2.6) in detail. The angle γ is indeed completely decoupled from the rest of the equations. We can treat (2.6a) and (2.6b) as a one-parameter family of planar Hamiltonian systems for the variables x and

y since I does not change in time. We will call this system the reduced system. The common way of depicting the dynamics of the equations (2.6) is to sketch the phase portraits of the reduced system on various levels of I and suppress the angle (see Figure 2.1). In such a representation, fixed points of the reduced system correspond to periodic orbits if the frequency is non zero, or closed curves of fixed points if the frequency is zero in the full equations. Periodic orbits correspond to invariant tori, and separatrices connecting fixed points to themselves to families of orbits homoclinic to periodic orbits or heteroclinic connections between pairs of points on a curve of equilibria. We can also use this geometric approach to explicitly compute the orbits of (2.6). We first calculate x and y as functions of time from the reduced system, then insert them in the equation (2.6d) and integrate (2.6d) to compute the angle γ. This will be carried out in Section 2.4.

The above discussion focused on (2.6) when I is fixed at a particular value. But indeed, I is more than just a parameter for (2.6), it is one of its variables. We may expect (and we will see shortly that this is the case) that the fixed points of the reduced system vary smoothly with I and that, therefore, in the x-y-I coordinates, the equations (2.6a), (2.6b) and (2.6c) possess smooth curves of equilibria with their x and y coordinates being expressed as smooth functions of I.

When we add the angle γ, these curves become two-dimensional surfaces, foliated by periodic orbits and occasional closed curves of fixed points, in the full four-dimensional space. In the same spirit, we may expect to see smooth one-

parameter families of orbits homoclinic to the fixed points lying on (possibly parts of) the smooth curves in the x-y-I space that we have just discussed. In the x-y-I coordinates, these families of homoclinic orbits will form two-dimensional homoclinic surfaces. In the full four-dimensional system (2.6) we will thus have to deal with three-dimensional manifolds, homoclinic to two-dimensional invariant surfaces. These two-dimensional invariant surfaces, together with their homoclinic manifolds which, obviously, are also invariant, are the backbone of the structure of our phase space (See Figure 2.2). What makes these structures especially interesting is that they are robust in their ability to persist under perturbation. In Section 3.3, we will discuss the exact nature and the extent of their persistence.

In order to at least locally avoid the singularity on Π_b we now make the second possible transformation by first writing

$$b = |b|e^{i\beta} \tag{2.7a}$$

$$c = (u + iv)e^{i\beta}\,, \tag{2.7b}$$

and then taking I, β, u and v as the new variables. The Hamiltonian, rewritten in these variables, is

$$H = \tfrac{3}{4}I^2 + (1+k^2)I - \tfrac{7}{16}u^4 - \tfrac{3}{8}u^2v^2 + \tfrac{1}{16}v^4 + \tfrac{1}{2}(k^2 + \tfrac{3}{2}I)u^2 + \tfrac{1}{2}(k^2 - \tfrac{1}{2}I)v^2\,, \tag{2.8}$$

and the equations (2.1) transform to

$$\dot{u} = \left(k^2 - \tfrac{1}{2}I\right)v - \tfrac{3}{4}u^2v + \tfrac{1}{4}v^3 \tag{2.9a}$$

$$\dot{v} = -\left(k^2 + \tfrac{3}{2}I\right) u + \tfrac{7}{4}u^3 + \tfrac{3}{4}uv^2 \qquad (2.9b)$$

$$\dot{I} = 0 \qquad (2.9c)$$

$$\dot{\beta} = 1 + k^2 - \tfrac{3}{2}I - \tfrac{3}{4}u^2 + \tfrac{1}{4}v^2 \,. \qquad (2.9d)$$

From (2.2b) it should be clear that the domain of this system is the interior of the paraboloid $u^2 + v^2 < 2I$ and Π_b is described by $u = v = 0$. The transformation is of course again singular at the boundary which is now mapped onto Π_c. Note that the axis of the paraboloid in the u-v coordinates corresponds to the boundary of the paraboloid in the x-y coordinates and vice versa. This completes our discussion of the reduced system.

2.3. Fixed Points of the Reduced Systems and Their Meaning in the Original Variables. The simplest structure to be analyzed in any dynamical system are its fixed points. On the other hand, knowledge of their position and stability type provides a great deal of insight into the geometric structure of the phase space. This is particularly true for planar Hamiltonian systems such as (2.6a), (2.6b) and (2.9a), (2.9b), since in this case the knowledge of the fixed points, their stability, and the separatrices that emerge from the unstable fixed points, completely determine the phase portraits. Besides that, for us, the fixed points of the reduced system are even more important since, due to a clever choice of coordinates, they represent either periodic orbits or circles of fixed points for the full equations (2.6) and (2.9). We will therefore devote the first part of this section to the analysis of

the fixed points of the reduced systems, and the second part to their interpretation in the full phase space.

We begin by listing the fixed points of (2.6a), (2.6b) and their stability types (see Figure 2.1). Throughout this discussion we will use (2.6a) and (2.6b) whenever possible; we will only use (2.9a) and (2.9b) to describe Π_b since the transformation leading to (2.6) is singular there. We also write down the eigenvalues of the linearizations of (2.6a), (2.6b) around these fixed points which reveal their stability. We find:

(a) The origin with the eigenvalues $\pm k\sqrt{2I - k^2}$ is a center for $I < k^2/2$ and a saddle for $I > k^2/2$.

(b) The points $x = 0$, $y = \pm 2k$, with the eigenvalues $\pm 2ik\sqrt{I - 2k^2}$ are centers for $I > 2k^2$ and hit the boundary of the paraboloid at $I = 2k^2$.

(c) The points $x = \pm 2\sqrt{\frac{1}{7}(2I - k^2)}$, $y = 0$, with the eigenvalues

$$\pm 2i\sqrt{\tfrac{1}{7}\,(2I - k^2)\,(2k^2 + 3I)}$$

are centers for $I > k^2/2$ and merge with the origin in a pitchfork bifurcation at $I = k^2/2$.

The boundary of the paraboloid contains up to four of what seem to be fixed points. Since the change of variables leading to (2.6) is singular there, these points are only fictitious and do not have a clear meaning, and therefore they can not be analyzed in the x-y coordinates. The correct thing to do is to look at

(d) The origin in the u-v coordinates with the eigenvalues $\pm\sqrt{(\frac{1}{2}I - k^2)\,(k^2 + \frac{3}{2}I)}$.

It is a center for $I < 2k^2$ and a saddle for $I > 2k^2$.

Making use of the Hamiltonian (2.5), we arrive at the phase portraits presented in Figure 2.1. The phase portraits in the u-v coordinates look similar and can be obtained by interchanging the role of the boundary and the centers of the circles.

The sequence of bifurcations of the reduced system (2.6) is as follows. For $I < k^2/2$ the origin, $x = y = 0$, is a center. At $I = k^2/2$ the fixed points (c) bifurcate off from it in a pitchfork bifurcation. The point $x = y = 0$ then becomes a saddle, connected to itself by a pair of homoclinic orbits described implicitly by the equation

$$H(x, y, I) - H(0, 0, I) = 0. \tag{2.10}$$

The reason that (2.10) is valid on the separatrices is that, by continuity, the value of the Hamiltonian on a homoclinic orbit must be the same as its value at the fixed point to which the orbit is homoclinic. At $I = 2k^2$ the fixed points (b) enter the disc $x^2 + y^2 < 2I$, and heteroclinic orbits appear which connect pairs of the fictitious fixed points on the circle $x^2 + y^2 = 2I$. These orbits enclose the fixed points (b).

Because of the singularity of the x-y-I coordinates at $x^2 + y^2 = 2I$, we have to look at this last bifurcation in the u-v-I coordinates. In that coordinate system the bifurcation at $I = 2k^2$ corresponds to a pitchfork bifurcation of the point $u = v = 0$, producing the fixed points (b) and a pair of homoclinic orbits connecting $u = v = 0$

to itself. In mixed coordinates these homoclinic orbits are described by the equation

$$H(x,y,I) - H(u=0,v=0,I) = 0 \tag{2.11}$$

with $H(x,y,I)$ given by formula (2.5) and $H(u,v,I)$ by formula (2.8).

At $I = 4k^2$, a global bifurcation occurs in which the orbits homoclinic to $x = y = 0$ and those homoclinic to $u = v = 0$ pass through each other giving rise to four heteroclinic connections. For $I > 4k^2$ the phase portraits look the same as those just before the global bifurcation rotated through an angle of ninety degrees. The separatrices are described by the same implicit equations as before.

Let us now interpret these fixed points of the reduced system in the (c,b) phase space.

(a) We know already that the line $x = y = 0$ corresponds to the plane Π_c. Inserting these values in (2.6d) and integrating (2.6d) gives the time evolution of the angle γ to be $\gamma = (1-I)t + \gamma_0$. From (2.4) and (2.2b) we then find that Π_c is foliated by circular periodic orbits $c = \sqrt{2I}e^{i\gamma}$. A circle of fixed points occurs at $I = 1$.

(b) The points $x = 0$, $y = \pm 2k$, correspond to the circles $c = \pm\sqrt{2(I-2k^2)}e^{i\gamma}$, $b = 2ike^{i\gamma}$, where $\gamma = (1-I)t + \gamma_0$. For $I = 1$ we get two circles of fixed points, one for each sign of c.

(c) The points $x = \pm 2\sqrt{\frac{1}{7}(2I-k^2)}$, $y = 0$ correspond to the circles

$$c = \pm\sqrt{\tfrac{2}{7}(3I+2k^2)}e^{i\gamma}, \qquad b = 2\sqrt{\tfrac{1}{7}(2I-k^2)}e^{i\gamma},$$

where $\gamma = [1 + \frac{1}{7}(4k^2 + 15I)]t + \gamma_0$. The two circles of fixed points are given by $I = \frac{1}{15}(4k^2 + 7)$.

(d) Finally the line $u = v = 0$ corresponds to the plane Π_b foliated by the periodic orbits $c = 0$, $b = \sqrt{2I}e^{i\beta}$, where $\beta = (1 + k^2 - \frac{3}{2}I)t + \beta_0$. Again,we find a circle of fixed points, this time at $I = \frac{2}{3}(1 + k^2)$.

Note that not all the circles of fixed points need to exist for all values of k and that there are at most six of them in the entire space. For convenience, we now list their position, stability, and the intervals of k in which they exist.

(a) The circle in Π_c has coordinates $c = \sqrt{2}e^{i\gamma}$, $b = 0$. The eigenvalues of the reduced system at this circle are $\pm k\sqrt{2 - k^2}$. It exists for all k and it is stable for $k > \sqrt{2}$ and unstable for $k < \sqrt{2}$.

(b) Two circles $c = \pm\sqrt{2(1 - 2k^2)}e^{i\gamma}$, $b = 2ike^{i\gamma}$ with the eigenvalues equal to $\pm 2ik\sqrt{1 - 2k^2}$ exist for $k < 1/\sqrt{2}$ and are stable.

(c) Two circles $c = \pm\sqrt{\frac{2}{5}(2k^2 + 1)}e^{i\gamma}$, $b = 2\sqrt{\frac{1}{15}(2 - k^2)}e^{i\gamma}$ have the eigenvalues $\pm\frac{2}{5}i\sqrt{7(2 - k^2)(2k^2 + 1)}$. They exist for $k < \sqrt{2}$ and are stable.

(d) The circle in Π_b lies at $c = 0$, $b = 2\sqrt{\frac{1}{3}(1 + k^2)}e^{i\beta}$ with the eigenvalues $\pm\sqrt{\frac{1}{3}(1 - 4k^4)}$ exists for all k and is stable for $k > 1/\sqrt{2}$ and unstable for $k < 1/\sqrt{2}$.

We now take a more global view of the structures we described above. The periodic orbits we have discovered form invariant two-dimensional surfaces, Π_c, Π_b

and two more. In the x-y-I coordinates, when we suppress the angle γ, these surfaces appear as curves of fixed points as shown on Figure 2.1. We see that the annuli $k^2/2 < I < 4k^2$ and $I > 4k^2$ in Π_c and $2k^2 < I < 4k^2$ and $I > 4k^2$ in Π_b are connected to themselves by three-dimensional homoclinic manifolds. (At $I = 4k^2$ these manifolds pass through each other in a global bifurcation.) In the x-y-I coordinates, the homoclinic manifolds are represented as two-dimensional surfaces; the projection of the manifold homoclinic to the annulus $\frac{k^2}{2} < I < 4k^2$ is presented on Figure 2.2.

The homoclinic manifolds are foliated into surfaces of constant I. In the x-y-I projection, these are just the individual separatrices forming the homoclinic surface. In the full four-dimensional phase space, though, they assume the form of "pinched" two-tori. In a rather unusual coordinate system, one of them is shown on Figure 2.3. The middle line, along which such a torus is pinched, is either a periodic orbit or a circle of fixed points, depending on the frequency. The rest of the torus consists of two symmetric two-dimensional surfaces. If the middle line is a periodic orbit, the two surfaces are further foliated by orbits homoclinic to it. If the middle line is a circle of fixed points, the surfaces are foliated by either heteroclinic orbits connecting pairs of fixed points on the circle, or else orbits homoclinic to the points on the circle. Which one of these two situations occurs depends, as we will see, on the value of the parameter k, and will play a great role in the perturbed system.

It is also interesting to note that, since the Hamiltonian is independent of the

decoupled angle, the manifolds homoclinic to annuli in Π_c are still described by the equation (2.10) and those homoclinic to annuli in Π_b by the equation (2.11). This representation is important, since it enables us to calculate tangent spaces and normals to the homoclinic manifolds.

2.4. Explicit Solutions of the Homoclinic Orbits. Due to the fact that the homoclinic connections are the primary candidates for being the sources of chaos, we now calculate closed form solutions for them. We obtain manageable expressions by further transforming the systems (2.6) and (2.9) to symplectic polar coordinates.

To transform (2.6) write

$$x + iy = \sqrt{2B}e^{i\theta} \tag{2.12}$$

and use B and θ as new variables. The Hamiltonian is transformed to

$$H = \tfrac{1}{2}I^2 - I - \tfrac{3}{4}B^2 + (I - k^2)B + B(I - B)\cos 2\theta\,, \tag{2.13}$$

and the equations are now

$$\dot{B} = -2B(I - B)\sin 2\theta \tag{2.14a}$$

$$\dot{\theta} = k^2 - I(1 + \cos 2\theta) + B(\tfrac{3}{2} + 2\cos 2\theta) \tag{2.14b}$$

$$\dot{I} = 0 \tag{2.14c}$$

$$\dot{\gamma} = 1 - I - B(1 + \cos 2\theta)\,. \tag{2.14d}$$

The analogous transformation of (2.9) is achieved by first writing

$$u + iv = \sqrt{2C}e^{i\phi} \tag{2.15}$$

and then using C and ϕ as new variables. The resulting Hamiltonian is

$$H = \tfrac{3}{4}I^2 - (1 + k^2)I - \tfrac{3}{4}C^2 + (\tfrac{1}{2}I + k^2)C + C(I - C)\cos 2\phi\,, \tag{2.16}$$

and the equations are

$$\dot{C} = -2C(I - C)\sin 2\phi \tag{2.17a}$$

$$\dot{\phi} = -k^2 - I(\tfrac{1}{2} + \cos 2\phi) + C(\tfrac{3}{2} + 2\cos 2\phi) \tag{2.17b}$$

$$\dot{I} = 0 \tag{2.17c}$$

$$\dot{\beta} = 1 + k^2 - \tfrac{3}{2}I - C(\tfrac{1}{2} + \cos 2\phi)\,. \tag{2.17d}$$

The variables in the systems (2.14) and (2.17) have more in common than we would expect at first glance. The relations

$$B + C = I \tag{2.18}$$

and

$$\theta = -\phi = \beta - \gamma \tag{2.19}$$

which we compute from (2.4), (2.7), (2.12), and (2.15) show that the transformation between the systems is a simple one. We also note that we can calculate β instead of γ in (2.14d) or γ instead of β in (2.17d) and use (2.19) to recover the other angle.

Let us first calculate the solutions on the orbits homoclinic to Π_c. By continuity, the value of the Hamiltonian (2.13) on the homoclinic orbits will be $\frac{1}{2}I^2 - I$, the same as at the origin $B = 0$. Inserting this in (2.13) and cancelling a common factor B we find the following connection between B and the angle θ to hold on the heteroclinic orbits:

$$B = I - \frac{4k^2 - I}{3 + 4\cos 2\theta}\,. \tag{2.20}$$

We now want to find the differential equations for the two remaining angles which will give us the final answer. Inserting (2.20) into (2.14b) we get the equation

$$\dot{\theta} = I(1 + \cos 2\theta) - k^2\,, \tag{2.21}$$

and since it is more convenient to calculate the angle β instead of γ we use (2.17d), (2.18) and (2.20) to find

$$\dot{\beta} = 1 - I - \tfrac{1}{4}B\,. \tag{2.22}$$

We first integrate the equation (2.21), then express B from (2.20), and finally integrate (2.22), to get the desired solution. As mentioned before, we can then calculate γ from (2.19), and we can also transform the solution back to I, γ, x, and y coordinates using (2.12).

Both families of homoclinic orbits, the one for $I > 4k^2$ and that for $I < 4k^2$, can be obtained from the above equations (2.20), (2.21), and (2.22) since in our derivation we never needed to specify the value of k. Also, since our equations are autonomous, all the solutions on a given homoclinic orbit can be obtained from a single one by time translation $t \mapsto t - t_0$, where t_0 is a suitable constant. This gives us great freedom at finding the solutions; in particular, we can choose especially simple initial conditions to calculate one of the solutions and then obtain the other solutions on the same homoclinic orbit by time translation. The suitable initial conditions, which give rise to solutions whose B and θ components are symmetric with respect to time reversal, are easily spotted on Figure 2.1. For brevity, we only display these particular solutions, but note again that any other solution on the same homoclinic orbit can be obtained by simply writing $t - t_0$ instead of t in all our formulas.

Consider first the case $I > 4k^2$. We use the initial conditions $\theta(t=0) = \frac{\pi}{2}$, $\beta(t=0) = \beta_0$ to compute

$$B = \frac{4k^2\left(2I - k^2\right)}{\left(I - 4k^2\right)\cosh\left(2k\sqrt{2I - k^2}t\right) + I + 3k^2} \tag{2.23a}$$

$$\cot\theta = \frac{k}{\sqrt{2I - k^2}}\tanh\left(k\sqrt{2I - k^2}t\right) \tag{2.23b}$$

$$\beta = -\frac{1}{\sqrt{7}}\tanh^{-1}\left[\sqrt{\frac{7k^2}{2I - k^2}}\tanh\left(k\sqrt{2I - k^2}t\right)\right] + (1 - I)t + \beta_0\,. \tag{2.23c}$$

When $I < 4k^2$, we take the initial conditions $\theta(t=0) = 0$, $\beta(t=0) = \beta_0$ and find

$$B = \frac{4k^2\left(2I - k^2\right)}{\left(4k^2 - I\right)\cosh\left(2k\sqrt{2I - k^2}t\right) + I + 3k^2} \tag{2.24a}$$

$$\tan\theta = \frac{\sqrt{2I - k^2}}{k}\tanh\left(k\sqrt{2I - k^2}t\right) \tag{2.24b}$$

$$\beta = -\frac{1}{\sqrt{7}}\tanh^{-1}\left[\sqrt{\frac{2I - k^2}{7k^2}}\tanh\left(k\sqrt{2I - k^2}t\right)\right] + (1 - I)t + \beta_0\,. \tag{2.24c}$$

Now let us look at the orbits homoclinic to Π_b. The points with the same value of the Hamiltonian (2.16) as the origin are given by

$$C = I - \frac{I - 4k^2}{3 + 4\cos 2\phi}\,. \tag{2.25}$$

Inserting in (2.17b) we compute

$$\dot{\phi} = k^2 + I(\tfrac{1}{2} + \cos 2\phi)\,. \tag{2.26}$$

Finally, using (2.14d), (2.18) and (2.25) gives

$$\dot{\gamma} = 1 + k^2 - \tfrac{3}{2}I + \tfrac{1}{4}C\,. \tag{2.27}$$

For $I > 4k^2$ we pick the initial conditions $\phi(t=0) = 0$, $\gamma(t=0) = \gamma_0$. Together with the above equations this gives

$$C = \frac{2\left(I - 2k^2\right)\left(3I + 2k^2\right)}{2\left(I - 4k^2\right)\cosh\left(\sqrt{\left(I - 2k^2\right)\left(3I + 2k^2\right)}t\right) + 5I - 6k^2} \tag{2.28a}$$

$$\tan\phi = \sqrt{\frac{3I + 2k^2}{I - 2k^2}}\tanh\left(\tfrac{1}{2}\sqrt{\left(I - 2k^2\right)\left(3I + 2k^2\right)}t\right) \tag{2.28b}$$

$$\gamma = -\frac{1}{\sqrt{7}}\tanh^{-1}\left[\sqrt{\frac{3I + 2k^2}{7\left(I - 2k^2\right)}}\tanh\left(\tfrac{1}{2}\sqrt{\left(I - 2k^2\right)\left(3I + 2k^2\right)}t\right)\right] + \left(\left(1 + k^2\right) - \tfrac{3}{2}I\right)t + \gamma_0\,. \tag{2.28c}$$

Finally, for $I < 4k^2$ we get, using the initial conditions $\phi(t=0) = \frac{\pi}{2}$, $\gamma(t = 0) = \gamma_0$, the following expressions

$$C = \frac{2\left(I-2k^2\right)\left(3I+2k^2\right)}{2\left(4k^2-I\right)\cosh\left(\sqrt{\left(I-2k^2\right)\left(3I+2k^2\right)}t\right)+5I-6k^2} \tag{2.29a}$$

$$\cot\phi = \sqrt{\frac{I-2k^2}{3I+2k^2}}\tanh\left(\tfrac{1}{2}\sqrt{\left(I-2k^2\right)\left(3I+2k^2\right)}t\right) \tag{2.29b}$$

$$\gamma = -\frac{1}{\sqrt{7}}\tanh^{-1}\left[\sqrt{\frac{7\left(I-2k^2\right)}{3I+2k^2}}\tanh\left(\tfrac{1}{2}\sqrt{\left(I-2k^2\right)\left(3I+2k^2\right)}t\right)\right] + \left(\left(1+k^2\right)-\tfrac{3}{2}I\right)t+\gamma_0\,. \tag{2.29c}$$

The motion on the heteroclinic connections at $I = 4k^2$ is described by

$$C = \frac{I}{1+e^{\pm\frac{\sqrt{7}}{2}I(t-t_0)}}\,, \tag{2.30a}$$

$$\cos 2\phi = -\tfrac{3}{4}\,, \tag{2.30b}$$

$$\begin{aligned}\beta &= \beta_0+\left(1-\tfrac{5}{4}I\right)(t-t_0)\mp\tfrac{1}{2\sqrt{7}}\log\left[1+e^{\mp\frac{\sqrt{7}}{2}I(t-t_0)}\right]\\ &= \beta_0+(1-I)(t-t_0)\mp\tfrac{1}{2\sqrt{7}}\log\left[1+e^{\pm\frac{\sqrt{7}}{2}I(t-t_0)}\right]\,.\end{aligned} \tag{2.30c}$$

For $I = \frac{4}{5}$ and $I = 1$ we get heteroclinic connections of fixed points to periodic orbits. We again point out that it is easy to transform these expressions to B, θ, and γ using (2.18) and (2.19) and also to all other variables, which we will not do.

2.5. Nature of the Orbits Homoclinic to the Circles of Fixed Points.

In this section we present a detailed analysis of the orbits homoclinic to the circles of fixed points. More precisely, we calculate the distance in angle between the initial and the end point on such an orbit from the formulas given in section 2.4. We will only consider the orbits homoclinic to the circle (a) in Π_c. The orbits homoclinic to the circle (d) in Π_b can be treated in the same way and have the same properties. However, since we will not use them, we do not present their analysis.

We look at the orbits (2.23) for $k < \frac{1}{2}$ and (2.24) for $k > \frac{1}{2}$ separately, beginning with (2.23). We parametrize the circle (a) by the angle γ via the equation $c = \sqrt{2}e^{i\gamma}$. The quantity we want to find is the difference $\Delta\gamma = \gamma(\infty) - \gamma(-\infty)$ between the two fixed points on the circle, connected by an orbit of the type (2.23). We calculate $\Delta\gamma$ from (2.23) and (2.19) with the aid of Figure 2.1. The equation (2.23c) first gives

$$\beta(-\infty) = \frac{1}{\sqrt{7}} \tanh^{-1}\left(\sqrt{\frac{7k^2}{2I-k^2}}\right) + \beta_0 \,. \tag{2.31a}$$

and

$$\beta(\infty) = -\frac{1}{\sqrt{7}} \tanh^{-1}\left(\sqrt{\frac{7k^2}{2I-k^2}}\right) + \beta_0 \,. \tag{2.31b}$$

From the equation (2.23b) and Figure 2.1. we get

$$\theta(-\infty) = \pi - \cot^{-1}\left(\frac{k}{\sqrt{2I-k^2}}\right) \tag{2.31c}$$

and

$$\theta(\infty) = \cot^{-1}\left(\frac{k}{\sqrt{2I-k^2}}\right) \tag{2.31d}$$

We use the equation (2.19) to first calculate $\beta_0 = \gamma_0 + \frac{\pi}{2}$ at $t = 0$ ad then combine (2.19) and (2.31) to get

$$\gamma(-\infty) = \gamma_0 - \frac{\pi}{2} + \cot^{-1}\left(\frac{k}{\sqrt{2I-k^2}}\right) + \frac{1}{\sqrt{7}}\tanh^{-1}\left(\sqrt{\frac{7k^2}{2I-k^2}}\right) \tag{2.32a}$$

and

$$\gamma(\infty) = \gamma_0 + \frac{\pi}{2} - \cot^{-1}\left(\frac{k}{\sqrt{2I-k^2}}\right) - \frac{1}{\sqrt{7}}\tanh^{-1}\left(\sqrt{\frac{7k^2}{2I-k^2}}\right) . \tag{2.32b}$$

The difference $\Delta\gamma$ is then equal to

$$\Delta\gamma = \pi - 2\cot^{-1}\left(\frac{k}{\sqrt{2-k^2}}\right) - \frac{2}{\sqrt{7}}\tanh^{-1}\left(\frac{\sqrt{7}k}{\sqrt{2-k^2}}\right) . \tag{2.33}$$

When k varies between 0 and $\frac{1}{2}$, this difference decreases monotonically from 0 to $-\infty$. Hence there are infinitely many cases when $\Delta\gamma = 2n\pi$ and the orbits are homoclinic. From the analysis it is also evident that the values of k for which the orbits become homoclinic possess a limit point at $k = \frac{1}{2}$. The phase portrait shows, though, that at that point the orbits become heteroclinic, connecting fixed points to periodic circles.

On the orbits (2.24) we compute

$$\beta(-\infty) = \frac{1}{\sqrt{7}} \tanh^{-1}\left(\sqrt{\frac{2I-k^2}{7k^2}}\right) + \beta_0 \qquad (2.34a)$$

$$\beta(\infty) = -\frac{1}{\sqrt{7}} \tanh^{-1}\left(\sqrt{\frac{2I-k^2}{7k^2}}\right) + \beta_0 \qquad (2.34b)$$

$$\theta(-\infty) = -\tan^{-1}\left(\frac{\sqrt{2I-k^2}}{k}\right) \qquad (2.34c)$$

$$\theta(\infty) = \tan^{-1}\left(\frac{\sqrt{2I-k^2}}{k}\right) . \qquad (2.34d)$$

The initial condition $\beta_0 = \gamma_0$ then gives

$$\gamma(-\infty) = \gamma_0 + \tan^{-1}\left(\frac{\sqrt{2-k^2}}{k}\right) + \frac{1}{\sqrt{7}} \tanh^{-1}\left(\frac{\sqrt{2-k^2}}{\sqrt{7}k}\right) , \qquad (2.35a)$$

$$\gamma(\infty) = \gamma_0 - \tan^{-1}\left(\frac{\sqrt{2-k^2}}{k}\right) - \frac{1}{\sqrt{7}} \tanh^{-1}\left(\frac{\sqrt{2-k^2}}{\sqrt{7}k}\right) , \qquad (2.35b)$$

and

$$\Delta\gamma = -2\tan^{-1}\left(\frac{\sqrt{2-k^2}}{k}\right) - \frac{2}{\sqrt{7}} \tanh^{-1}\left(\frac{\sqrt{2-k^2}}{\sqrt{7}k}\right) . \qquad (2.36)$$

For k between $\frac{1}{2}$ and $\sqrt{2}$ this function increases monotonically from $-\infty$ to 0. Hence we get again infinitely many values of k with homoclinic orbits accumulating on $k = \frac{1}{2}$.

We have thus proven that, except for isolated values of k, the orbits connecting points on the circle of equilibria are heteroclinic. The difference in the angle γ between their forward and backward limit points increases monotonically for k between $k = 0$ and $k = \frac{1}{2}$ until it becomes infinite at $k = \frac{1}{2}$. In between, there are infinitely many values of k for which a homoclinic orbit exists. At the n-th such value of k, the homoclinic orbit will wrap around the "Pinched" homoclinic torus (similar to the one on Figure 2.3) n times before returning to the fixed point it emerged from. The same situation, only in reverse order, occurs for k between $k = \frac{1}{2}$ and $k = \sqrt{2}$. Besides being interesting in its own right, the precise knowledge of these heteroclinic connections will play a major role in Chapter 4, where we give criteria for the existence of a pair of Šilnikov-type saddle connections.

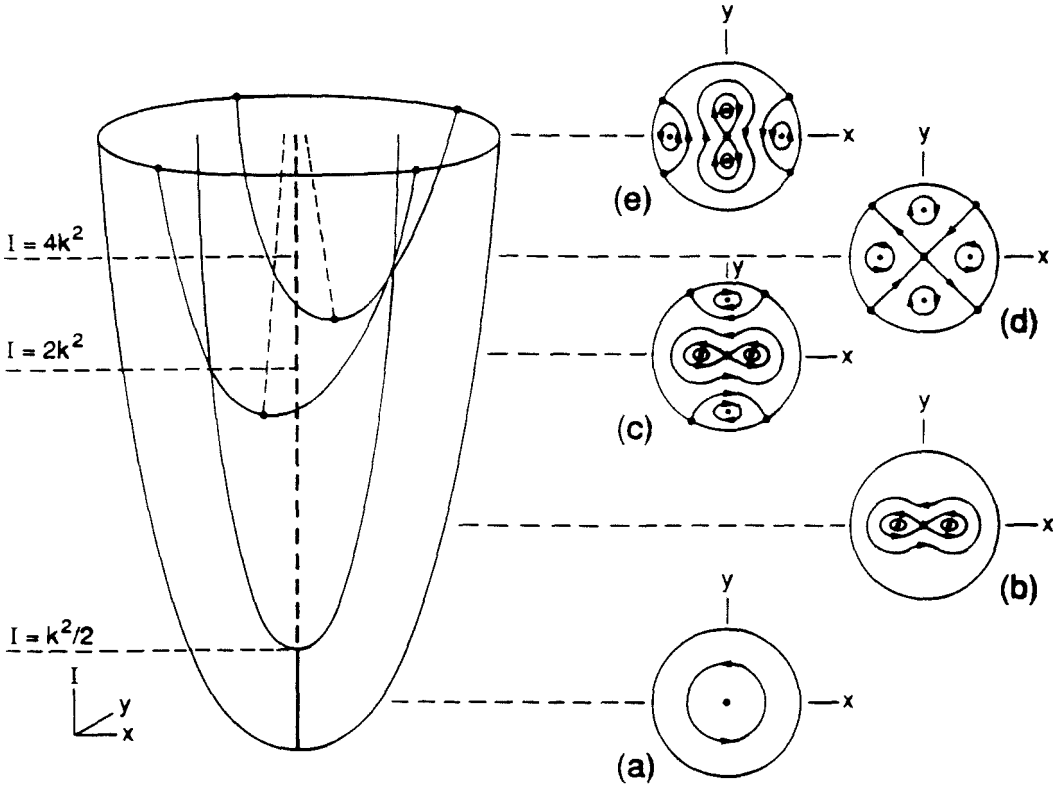

Figure 2.1. Paraboloid $x^2 + y^2 = 2I$ with the fixed points of (2.6)and the phase portraits on different levels of I:

(a) $0 < I < k^2/2$, (b) $k^2/2 < I < 2k^2$,

(c) $2k^2 < I < 4k^2$, (d) $I = 4k^2$,

(e) $I > 4k^2$.

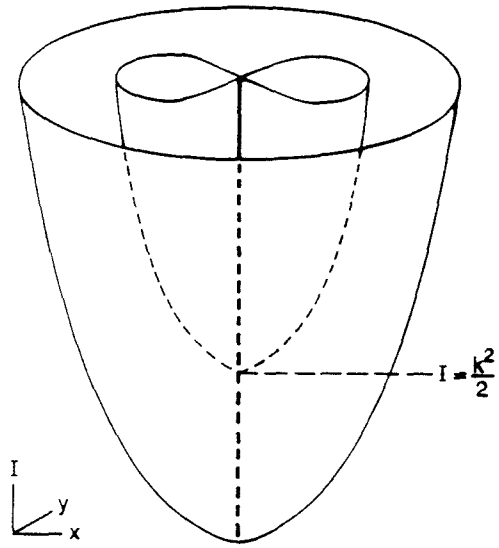

Figure 2.2. The homoclinic manifold for the annulus with $k^2/2 < I < 4k^2$ in Π_c.

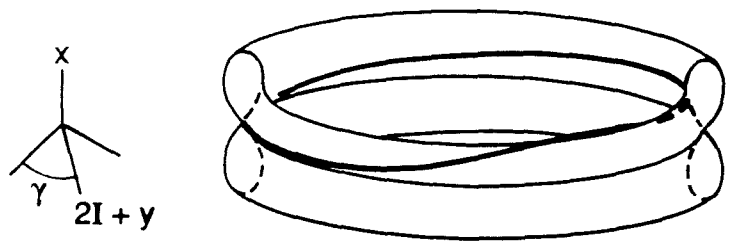

Figure 2.3. A "Pinched" torus of orbits homoclinic to a periodic orbit at a fixed I value.

CHAPTER 3

INVARIANT MANIFOLDS IN THE PERTURBED SYSTEM

3.1. Introduction. In this and all the later sections of this work we will investigate the perturbed equations

$$\dot{c} = -i\left(\tfrac{1}{2}|c|^2 + \tfrac{1}{2}|b|^2 - 1\right)c - \tfrac{i}{2}\left(cb^* + bc^*\right)b - \epsilon\alpha c - \epsilon\Gamma \qquad (3.1a)_\epsilon$$

$$\dot{b} = -i\left(\tfrac{1}{2}|c|^2 + \tfrac{3}{4}|b|^2 - \left(1 + k^2\right)\right)b - \tfrac{i}{2}\left(cb^* + c^*b\right)c - \epsilon\alpha b\,. \qquad (3.1b)_\epsilon$$

We will begin by showing that the plane Π_c, which is invariant under the flow of the unperturbed system, is also invariant under the flow of the perturbed system. Then restricting ourselves to a compact annulus $\mathcal{A}$ in Π_c, we will prove the existence of three-dimensional manifolds, $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$, which are spanned by orbits, and which are close to the unperturbed stable and unstable manifolds, $W^s_0(\mathcal{A})$ and $W^u_0(\mathcal{A})$, of the annulus $\mathcal{A}$. The role $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ play in the perturbed problem is similar to the role of $W^s_0(\mathcal{A})$ and $W^u_0(\mathcal{A})$ in the unperturbed problem. However, their interpretation as stable and unstable manifolds requires careful explanation since $\mathcal{A}$ is only locally invariant under the perturbed dynamics. Using the Melnikov technique, we will investigate the conditions for $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ to intersect and some special features of such intersections, such as their dimension and the fate of the orbits contained in them. Most of the subsequent work will be

devoted to determining the conditions under which such orbits will be biasymptotic to a fixed point or a periodic orbit in $\mathcal{A}$. In some cases, the existence of such orbits will imply the existence of nearby chaotic invariant sets.

We will consider two different situations: the nondissipative case when α is zero and only the forcing Γ is nonzero, and the dissipative case when both parameters α and Γ are nonzero. The main difference between the two cases is that in the nondissipative case, the system $(3.1)_\epsilon$ is still Hamiltonian with the Hamiltonian function

$$H_\epsilon = \tfrac{1}{8}|c|^4 + \tfrac{1}{2}|b|^2|c|^2 + \tfrac{3}{16}|b|^4 - \tfrac{1}{2}\left(1+k^2\right)|b|^2 - \tfrac{1}{2}|c|^2 + \tfrac{1}{8}\left(b^2c^{*2} + b^{*2}c^2\right) - \epsilon\Gamma c_2\,, \qquad (3.2)_\epsilon$$

whereas it is not in the dissipative case. We remark that our discussion in the first paragraph of this section applies to both of these cases.

In the nondissipative case the plane Π_c is still almost completely foliated by periodic orbits which are close to the unperturbed ones, except for a resonance layer around what used to be the circle of fixed points in the unperturbed system. This is to be expected due to the Hamiltonian nature of the nondissipative case. We will calculate these periodic orbits directly as the level sets of the restriction of the Hamiltonian $(3.2)_\epsilon$ to Π_c. However, it is interesting to note that we could prove their existence in a purely geometric fashion as well. We only need to observe that, away from the resonance layer, the three-dimensional level sets of the unperturbed Hamiltonian intersect Π_c transversely along the unperturbed periodic orbits. Since

transverse intersections persist under small perturbations, and since the periodic orbits away from the resonance have nonzero frequencies in the unperturbed case, they persist under perturbation.

As a simple consequence of our investigation of the intersections of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ and some transversality considerations, we will be able to prove the existence of transverse homoclinic orbits connecting periodic orbits in certain annuli $\mathcal{A}$ in Π_c, not containing the resonance, to themselves.

When the dissipation α is nonzero, we will show that no closed orbits remain in Π_c. Hence the orbits, found in the nondissipative case, which are homoclinic to periodic orbits outside the resonance have no counterpart in the dissipative case. All the interesting dynamics in this case, therefore, have their origin in the resonance band.

Due to its central importance, the resonance band will receive our special attention. It is born out of the circle of fixed points as a consequence of the perturbation. In the nondissipative case, the resonance band contains two fixed points, one of them connected to itself by two separatrices lying in Π_c and enclosing a family of periodic orbits nested around the other fixed point (see Figure 3.2). In the dissipative case, when α is not too large compared with Γ, the two fixed points persist, but the periodic orbits and the separatrices vanish. All these structures are very different from those that existed in the unperturbed problem. Above the resonance band off Π_c, the perturbed orbits in the intersections of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ for

I near the resonant value $I = 1$ also differ drastically from the unperturbed ones which are just the heteroclinic orbits connecting pairs of points along the circle of fixed points. Therefore, standard Melnikov analysis is not a sufficient tool for the investigation of the orbits homoclinic to the fixed points or the periodic orbits in the resonance band. Standard Melnikov analysis can only be applied in the case when in the unperturbed system we have an invariant set which persists for the perturbed system, and is, in the absence of the perturbation, connected to itself by a family of homoclinic orbits. The Melnikov function then provides a criterion for the persistence of some of those homoclinic orbits. Our analysis must therefore proceed in a manner different from the usual one. The technique we have developed consists of two stages.

In the first stage we use ramifications of the results about the intersections of the manifolds $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ to make sure that a certain orbit is contained in the intersection. In the second stage we use geometric singular perturbation theory in the vicinity of the resonance to describe that particular orbit with the help of orbits in Π_c. Loosely speaking, the perturbed orbit will be constructed from two parts, one part being close to the unperturbed heteroclinic orbit and the other part close to an orbit in Π_c. The two stages of our analysis correspond to the two parts in this construction with the end result being the connection of these two parts into one orbit of the perturbed system.

The main result we derive in the dissipative case is a proof of the existence

of orbits homoclinic to one of the two fixed points inside the resonance using the two stage technique that we have developed. The same kind of orbit appears to be responsible for the chaos found in the numerical simulations (see Bishop, Flesch, Forest, McLaughlin, and Overman [1989]) of the model. We note, however, that the orbits whose existence we will prove are not necessarily the ones observed numerically.

With slight modifications, we could obtain similar results for Π_b and orbits homoclinic to nondissipative periodic orbits in it but not for the phenomena in the resonance band which are the most interesting. Furthermore, since the dynamics originating near Π_b are not relevant for modeling the nonlinear Schroedinger and Sine-Gordon equations, we omit mentioning them in all further discussion.

Before concluding this section, we rewrite system $(3.1)_\epsilon$ in the coordinates which we will use for the remainder of our analysis, briefly comment on the class of problems to which some aspects of our problem belong, and present a short outline of the remainder of this chapter.

The coordinates we will use in all of our analysis are I, γ, x, and y. The perturbed equations $(3.1)_\epsilon$ in these coordinates assume the form

$$\dot{x} = -k^2 y + \tfrac{1}{4}y^3 - \tfrac{3}{4}x^2 y + \epsilon \left[\Gamma \frac{y}{\sqrt{2I - x^2 - y^2}} \sin\gamma - \alpha x \right] \qquad (3.3a)_\epsilon$$

$$\dot{y} = \left(k^2 - 2I\right) x + \tfrac{7}{4}x^3 + \tfrac{3}{4}xy^2 - \epsilon \left[\Gamma \frac{x}{\sqrt{2I - x^2 - y^2}} \sin\gamma + \alpha y \right] \qquad (3.3b)_\epsilon$$

$$\dot{I} = -\epsilon \left[\Gamma \sqrt{2I - x^2 - y^2} \cos\gamma + 2\alpha I \right] \qquad (3.3c)_\epsilon$$

$$\dot{\gamma} = 1 - I - x^2 + \epsilon\Gamma \frac{1}{\sqrt{2I - x^2 - y^2}} \sin\gamma \,. \qquad (3.3d)_\epsilon$$

The Hamiltonian for the nondissipative case in the I, γ, x, and y coordinates is

$$H_\epsilon = \tfrac{1}{2}I^2 - I - \tfrac{7}{16}x^4 - \tfrac{3}{8}x^2y^2 + \tfrac{1}{16}y^4 + (I - \tfrac{1}{2}k^2)x^2 - \tfrac{1}{2}k^2y^2 - \epsilon\Gamma\sqrt{2I - x^2 - y^2}\sin\gamma \,. \qquad (3.4)_\epsilon$$

Using the unperturbed Hamiltonian $H = H_{\epsilon=0}$ we can write $(3.3)_\epsilon$ in the more compact form

$$\dot{x} = \frac{\partial H}{\partial y}(x, y, I; k) + \epsilon g^x(x, y, I, \gamma; \alpha, \Gamma) \qquad (3.5a)_\epsilon$$

$$\dot{y} = -\frac{\partial H}{\partial x}(x, y, I; k) + \epsilon g^y(x, y, I, \gamma; \alpha, \Gamma) \qquad (3.5b)_\epsilon$$

$$\dot{I} = \epsilon g^I(x, y, I, \gamma; \alpha, \Gamma) \qquad (3.5c)_\epsilon$$

$$\dot{\gamma} = -\frac{\partial H}{\partial I}(x, y, I; k) + \epsilon g^\gamma(x, y, I, \gamma; \alpha, \Gamma) \qquad (3.5d)_\epsilon$$

Besides being convenient, this notation points to the fact that there is a whole class of systems, possibly in higher dimensions, having similar features as ours. In particular, if a system of the form $(3.5)_\epsilon$ possesses invariant manifolds, such as those parts of Π_c and Π_b which are connected to themselves by homoclinic manifolds, then it shares many common geometric properties with ours. A theory based on their geometry has been developed to analyze certain properties of such systems. An exposition of this theory can be found in Wiggins [1988]. Parts of it include the treatment of invariant manifolds to be presented in the next section and the Melnikov function analysis to be presented in sections 3.4 and 3.5, the results of which can immediately be extended to any two degree of freedom system of the

form $(3.5)_\epsilon$ possessing in its unperturbed form the same type of invariant manifolds as our unperturbed system $(3.5)_0$.

The remainder of Section 3 is organized as follows: In Section 3.2 we present a detailed discussion of the dynamics in the plane Π_c. In Section 3.3 we discuss the stable and unstable manifolds of annuli in Π_c. We establish criteria for these manifolds to intersect in Section 3.4. Finally, in Section 3.5, we examine orbits homoclinic to periodic orbits in Π_c away from the resonance band in the nondissipative case. Our analysis of the dynamics originating from the resonance band will be presented in Chapter 4.

3.2. Dynamics in Π_c. Since we will be interested in orbits homoclinic to fixed points and periodic orbits in Π_c, we now study the dynamics on this plane in detail. As we have remarked in Section 3.1, the plane Π_c is invariant under the flow of $(3.1)_\epsilon$. This is established by observing that in $(3.1b)_\epsilon$ we have $\dot{b} = 0$ for $b = 0$. Hence, the dynamics of orbits on Π_c are described by equation $(3.1a)_\epsilon$ restricted to Π_c which is

$$\dot{c} = -i\left(\tfrac{1}{2}|c|^2 - 1\right)c - \epsilon(\Gamma + \alpha c)\,. \qquad (3.6)_\epsilon$$

As we have shown in Section 2, we understand the dynamics of $(3.6)_\epsilon$ completely when $\epsilon = 0$. In particular, we have shown in Section 2.3 that Π_c is foliated by circular periodic orbits $c = \sqrt{2I}e^{i\gamma}$, where $\gamma = (1 - I)t + \gamma_0$. The orbit at $I = 1$ is a circle of fixed points and there is also a center-type fixed point at the origin, as is shown in Figure 3.1.

In the nondissipative case we can derive equation $(3.6)_\epsilon$ from the Hamiltonian

$$H_\epsilon = \tfrac{1}{8}|c|^4 - \tfrac{1}{2}|c|^2 - \epsilon\Gamma c_2 \,. \qquad (3.7)_\epsilon$$

We can again completely analyze the phase portrait of $(3.6)_\epsilon$ in this case. Since all orbits on Π_c are level sets of the Hamiltonian $(3.7)_\epsilon$, we can express c_1 from this equation in terms of c_2 and H_ϵ and hence find explicit expressions for the orbits. We present the phase portrait thus obtained in Figure 3.2. We see that periodic orbits away from $\frac{1}{2}|c| = I = 1$ persist and that their shape is only slightly distorted by the forcing. The fixed point near the origin also persists and is an $\mathcal{O}(\epsilon)$ distance away from its original position. This $\mathcal{O}(\epsilon)$ estimate follows by applying the implicit function theorem. The main changes that forcing introduces take place around $I = 1$ where we find a resonance band instead of the circle of fixed points. From Figure 3.2, it should be clear that only two fixed points from this circle survive the perturbation. One of them becomes a center and one a saddle, the saddle being connected to itself by two homoclinic orbits encircling the center. To find the accurate position of these points we must solve $(3.6)_\epsilon$ with $\alpha = 0$ and zero left hand side. The solutions of this equation are given by $c_1 = 0$ and c_2 assuming the values of the three real roots of the the cubic equation $\left(\frac{1}{2}c_2^2 - 1\right) c_2 - \epsilon\Gamma = 0$. We will not solve this cubic equation since for the purposes of our analysis we only care about the existence and approximate location of the fixed points and not about their exact coordinates.

In the dissipative case, orbits cannot be calculated so easily any more. By more

careful analysis, though, we can still recover most of the relevant dynamics.

We begin by discussing the fixed points. We can again use the implicit function theorem to show that the fixed point near the origin persists $\mathcal{O}(\epsilon)$ close to where it was in the unperturbed system. We will show shortly that the resonance band is still present when α is not too large compared with Γ and that there are two surviving fixed points in it. We could show their existence directly by again deriving a cubic equation for their precise location, but since we will show their existence later by a more refined method, we omit this step. We remark that the system $(3.6)_\epsilon$ is dissipative for positive $\epsilon\alpha$, hence its fixed points can only be saddles or sinks. This fact holds for the full four-dimensional equations $(3.1)_\epsilon$ as well. The eigenvalues in all four dimensions of the unperturbed center at the origin are easily calculated to be $\pm i$ and $\pm ik^2$. For small enough ϵ, the eigenvalues will pick up small negative real parts due to the dissipation. Therefore, this point can only become a sink.

We also need to address the question whether any closed orbits remain in Π_c in the dissipative case. The answer is given by the Bendixson criterion for planar systems (see Guckenheimer and Holmes [1983], p. 44, or Wiggins [1989]). This criterion states that in a convex region in the plane where the divergence of the vector field does not change sign, there cannot be any closed orbits. The divergence of the right hand side of equations $(3.6)_\epsilon$ is $-2\epsilon\alpha$, hence no closed orbits exist in Π_c. This excludes the existence of either periodic orbits or separatrix loops and leaves the three fixed points as the only possible limit sets in Π_c.

To obtain more detailed information about the dissipative dynamics on Π_c, we must separately examine what happens away from $\frac{1}{2}|c|^2 = I = 1$ and what happens near $\frac{1}{2}|c|^2 = I = 1$. The reason for separating the analysis in this way is that even in the dissipative case we expect to see resonant behavior near $\frac{1}{2}|c|^2 = I = 1$ since the unperturbed terms are comparable in size to the perturbation there.

In what is to follow, we first want to show that for positive $\epsilon\alpha$, all the orbits away from the resonance spiral inwards, either towards the resonance when $\frac{1}{2}|c|^2 = I > 1$ or towards the attracting fixed point near the origin when $\frac{1}{2}|c|^2 = I < 1$. We will show this by using the nonresonant periodic orbits of the Hamiltonian $(3.7)_\epsilon$ as the closed curves with the vector field $(3.6)_\epsilon$ everywhere on them pointing into the regions they enclose. This will imply, for instance, that the basin of attraction of the sink near the origin will contain all the points away from the resonance with $\frac{1}{2}|c|^2 = I < 1$. We will then examine the resonance in detail. In particular, we will look at the remaining two fixed points and determine as accurately as possible the locations of the basins of attraction of the two attracting fixed points in Π_c.

Let us now first look at the situation outside the resonance. We want to show that on every nonresonant orbit of the Hamiltonian H_ϵ, the vector field points into the region enclosed by it. To prove this, note that the outward pointing normal at any point of such an orbit is

$$\begin{aligned} \mathbf{n} &= \left(\frac{\partial H_\epsilon}{\partial c_1}, \frac{\partial H_\epsilon}{\partial c_2} \right) \\ &= \left(\left(\tfrac{1}{2}|c|^2 - 1 \right) c_1, \left(\tfrac{1}{2}|c|^2 - 1 \right) c_2 \right) \end{aligned}$$

when $\frac{1}{2}|c|^2 = I > 1$ and

$$\begin{aligned}\mathbf{n} &= \left(-\frac{\partial H_\epsilon}{\partial c_1}, -\frac{\partial H_\epsilon}{\partial c_2}\right) \\ &= \left(-\left(\tfrac{1}{2}|c|^2 - 1\right) c_1, -\left(\tfrac{1}{2}|c|^2 - 1\right) c_2\right)\end{aligned}$$

when $\frac{1}{2}|c|^2 = I < 1$. The scalar product between the normal $\mathbf{n}$ and the vector field $(3.6)_\epsilon$ is equal to

$$\langle \mathbf{n}, \dot{c} \rangle = \epsilon\alpha \left[-\left(\tfrac{1}{2}|c|^2 - 1\right)|c|^2 + \epsilon\Gamma c_2\right]$$

for $\frac{1}{2}|c|^2 = I > 1$ and

$$\langle \mathbf{n}, \dot{c} \rangle = \epsilon\alpha \left[\left(\tfrac{1}{2}|c|^2 - 1\right)|c|^2 - \epsilon\Gamma c_2\right]$$

for $\frac{1}{2}|c|^2 = I < 1$. On a nonresonant orbit, $\frac{1}{2}|c|^2 = I$ varies by an $\mathcal{O}(\epsilon)$ amount only, and since it is not equal to 1 or 0 (we must exclude orbits too close to the origin, too), either of the above two expressions, which holds on that orbit, is negative and our conclusion follows. Therefore, indeed, orbits spiral inwards when they are away from the resonance.

We will now take a closer look at the resonance band. The idea is to magnify it by first transforming equations $(3.6)_\epsilon$ to the action-angle variables which in this case happen to be I and γ, and then rescale the I coordinate so that the thickness of the band will be finite for all values of ϵ. We also want the limit as $\epsilon \to 0$ of the system we obtain from this rescaling to have a simple form, a property which will follow automatically from our transformations. The above described procedure is standard in the theory of averaging. A good general description of it can be

found in Arnold, [1988], p. 159, or Wiggins [1989], and an application is given in Guckenheimer and Holmes, [1983], pp. 205-211.

Equations $(3.6)_\epsilon$ transformed to I and γ coordinates are

$$\dot{I} = -\epsilon\left(\Gamma\sqrt{2I}\cos\gamma + 2\alpha I\right) \qquad (3.8a)_\epsilon$$

$$\dot{\gamma} = 1 - I + \frac{\epsilon\Gamma}{\sqrt{2I}}\sin\gamma\,. \qquad (3.8b)_\epsilon$$

We take a new variable h defined by $I = 1 + \sqrt{\epsilon}h$, rescale the time $\tau = \sqrt{\epsilon}t$ and denote differentiation with respect to τ by $'$. Thus we transform $(3.8)_\epsilon$ to the new set of equations

$$h' = -\Gamma\sqrt{2(1+\sqrt{\epsilon}h)}\cos\gamma - 2\alpha(1+\sqrt{\epsilon}h) \qquad (3.9a)_\epsilon$$

$$\gamma' = -h + \frac{\sqrt{\epsilon}\Gamma}{\sqrt{2(1+\sqrt{\epsilon}h)}}\sin\gamma\,. \qquad (3.9b)_\epsilon$$

In the limit as $\epsilon \to 0$, they become the equations of the constantly forced pendulum

$$h' = -\Gamma\sqrt{2}\cos\gamma - 2\alpha \qquad (3.9a)_0$$

$$\gamma' = -h \qquad (3.9b)_0$$

which possess the Hamiltonian

$$\mathcal{H} = \tfrac{1}{2}h^2 - \Gamma\sqrt{2}\sin\gamma - 2\alpha\gamma\,. \qquad (3.10)$$

This limit also shows that the original size of the resonance band was $\mathcal{O}(\sqrt{\epsilon})$, which will be important to know in the later sections when we need such estimates in

order to couple the dynamics in the resonance with the dynamics off Π_c near the resonance.

We will obtain information about the dynamics of the equations $(3.9)_\epsilon$ by first analyzing their limit $(3.9)_0$ and then applying perturbation techniques. Due to their Hamiltonian nature, equations $(3.9)_0$ can easily be studied.

Before taking advantage of this fact, though, we first investigate their equilibria. The fixed points of $(3.9)_0$ have coordinates (h, γ) which satisfy the equations

$$h = 0$$

$$\cos\gamma = -\sqrt{2}\frac{\alpha}{\Gamma}$$

We see that these equations give two fixed points, we will call them p_0 and q_0, when $\alpha < \frac{\Gamma}{\sqrt{2}}$. Since the angle $\cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}$ must lie between 0 and $\frac{\pi}{2}$ for positive α and Γ, we find their h-γ coordinates to be

$$p_0 = \left(0, \pi - \cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}\right) \tag{3.11a}$$

$$q_0 = \left(0, \pi + \cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}\right) . \tag{3.11b}$$

At $\alpha = \frac{\Gamma}{\sqrt{2}}$, they coalesce in a saddle-node bifurcation. Note that

$$\sin\gamma = \sqrt{1 - 2(\tfrac{\alpha}{\Gamma})^2}$$

at p_0 and

$$\sin\gamma = -\sqrt{1 - 2(\tfrac{\alpha}{\Gamma})^2}$$

at q_0. The matrix associated with the linearization of the system $(3.9)_\epsilon$ is given by

$$\begin{pmatrix} -\frac{\sqrt{\epsilon}\Gamma}{\sqrt{2(1+\sqrt{\epsilon}h)}}\cos\gamma - 2\sqrt{\epsilon}\alpha & \Gamma\sqrt{2(1+\sqrt{\epsilon}h)}\sin\gamma \\ -1 - \frac{\epsilon\Gamma}{[2(1+\sqrt{\epsilon}h)]^{\frac{3}{2}}}\sin\gamma & \frac{\sqrt{\epsilon}\Gamma}{\sqrt{2(1+\sqrt{\epsilon}h)}}\cos\gamma \end{pmatrix}.$$

The characteristc polynomial of this matrix is

$$\lambda^2 + 2\sqrt{\epsilon}\alpha\lambda + \Gamma\sqrt{2(1+\sqrt{\epsilon}h)}\sin\gamma - \epsilon\left(\frac{2\alpha\Gamma}{\sqrt{2(1+\sqrt{\epsilon}h)}}\cos\gamma + \frac{\Gamma^2}{2(1+\sqrt{\epsilon}h)}\cos 2\gamma\right), \tag{3.12}$$

which at $\epsilon = 0$ reduces to

$$\lambda^2 + \sqrt{2}\Gamma\sin\gamma\,.$$

From this equation and (3.11) one immediately obtains the eigenvalues at p_0 and q_0. At p_0 they are

$$\pm i\sqrt{\Gamma}\left[2(1-2(\tfrac{\alpha}{\Gamma})^2)\right]^{\frac{1}{4}},$$

and at q_0 they are

$$\pm\sqrt{\Gamma}\left[2(1-2(\tfrac{\alpha}{\Gamma})^2)\right]^{\frac{1}{4}}.$$

Therefore, p_0 is a center and q_0 a saddle.

Having found all the information about the fixed points of $(3.9)_0$, we now calculate its orbits. They are the level sets of the Hamiltonian (3.10), and can be obtained explicitly by expressing h in terms of $\mathcal{H}$ and γ from equation (3.10). The resulting phase portrait is presented in Figure 3.3. In particular, note that the stable and unstable manifolds of q_0, which we denote by $W_0^s(q_0)$ and $W_0^u(q_0)$, intersect along a separatrix loop, which we denote by $W(q_0)$. This loop divides the phase space in two regions, the inside of the loop $W(q_0)$ being one and the outside

the other. Note that all the orbits inside $\mathcal{W}(q_0)$ are periodic. They encircle the fixed point p_0. Outside of $\mathcal{W}(q_0)$ no orbits are periodic; they descend from $h = +\infty$, pass by $\mathcal{W}(q_0)$ and go off to $h = -\infty$. This agrees with the fact we established earlier in the c coordinates that in the perturbed problem orbits away from the resonance spiral inwards (meaning on average decreasing their h coordinate).

A special situation occurs in the nondissipative case, *i.e.*, when $\alpha = 0$. Equations $(3.9)_0$ then become the equations of the simple pendulum. The point q_0 is now connected to itself by two separatrix loops rather than one, and all the other orbits are periodic, as can be expected from the previously described direct analysis on the nondissipative case (see Figure 3.2). Therefore it comes as no surprise that for $\alpha = 0$ the equations $(3.9)_\epsilon$ are also Hamiltonian with the rescaled Hamiltonian

$$\hat{H}_\epsilon = \tfrac{1}{\epsilon}(H_\epsilon + \tfrac{1}{2}) = \tfrac{1}{2}h^2 - \Gamma\sqrt{2(1+\sqrt{\epsilon}h)}\sin\gamma$$

for all positive ϵ. Here $\hat{H}_\epsilon$ is the Hamiltonian $(3.7)_\epsilon$ first recast in the I-γ coordinates, which is

$$H_\epsilon = \tfrac{1}{2}I^2 - I - \epsilon\Gamma\sqrt{2I}\sin\gamma\,, \qquad (3.13)_\epsilon$$

and then transformed to the h-γ coordinates.

Having completed the analysis of equations $(3.9)_0$, we now turn to the perturbed equations $(3.9)_\epsilon$. We first analyze their equilibria with the help of knowledge about the equilibria of equations $(3.9)_0$. For small nonzero ϵ, the implicit function theorem states that there exist fixed points p_ϵ and q_ϵ which are $\mathcal{O}(\sqrt{\epsilon})$ away from p_0 and q_0, respectively, and whose coordinates are analytic functions of α, Γ, and $\sqrt{\epsilon}$.

The $\mathcal{O}(\sqrt{\epsilon})$, rather than an $\mathcal{O}(\epsilon)$ estimate, arises from the fact that equations $(3.9)_\epsilon$ contain $\sqrt{\epsilon}$ as the small parameter rather than ϵ. The eigenvalues of the stability matrix of $(3.9)_\epsilon$ evaluated at p_ϵ, can then be obtained from (3.12) and are

$$-\sqrt{\epsilon}\alpha \pm i\left\{\sqrt{\Gamma}\left[2(1-2(\tfrac{\alpha}{\Gamma})^2)\right]^{\frac{1}{4}} + \mathcal{O}(\sqrt{\epsilon})\right\}.$$

Similarly, the eigenvalues at q_ϵ are

$$-\sqrt{\epsilon}\alpha \pm \left\{\sqrt{\Gamma}\left[2(1-2(\tfrac{\alpha}{\Gamma})^2)\right]^{\frac{1}{4}} + \mathcal{O}(\sqrt{\epsilon})\right\}.$$

We see that q_ϵ remains a saddle for sufficiently small ϵ, and that p_ϵ turns into a sink for α and $\sqrt{\epsilon}$ positive.

We now analyze the global dynamics of system $(3.9)_\epsilon$. From this point on, we will always assume that α and $\sqrt{\epsilon}$ are nonnegative, except in Lemma 3.1. Our particular interest lies in determining as precisely as possible the relative locations of the basins of attraction of the two sinks, p_ϵ and the sink near the origin. As a result of our analysis, we will establish that the phase portrait of $(3.9)_\epsilon$ is as shown in Figure 3.4. The stable and unstable manifolds, $W^s_\epsilon(q_\epsilon)$ and $W^u_\epsilon(q_\epsilon)$, of the fixed point q_ϵ do not intersect as they did in system $(3.9)_0$. Both branches of $W^s_\epsilon(q_\epsilon)$ come from $h = +\infty$ as opposed to only one branch of $W^s_0(q_0)$. These two branches separate the plane into the basins of attraction of p_ϵ and the sink near the origin. Furthermore, the branch of $W^u_\epsilon(q_\epsilon)$ which in the unperturbed case coincided with a branch of $W^s_0(q_0)$ to form the loop $W(q_0)$, now falls into p_ϵ.

In order to prove these statements, we first return to the unperturbed problem and define two regions, $\mathcal{R}_0$ and $\mathcal{S}_0$, which will, under perturbation, deform into

trapping regions, $\mathcal{R}_\epsilon$, containing p_ϵ, and $\mathcal{S}_\epsilon$ containing the origin (see Figure 3.5 and Figure 3.7). The region $\mathcal{R}_0$ will be the interior of the loop $\mathcal{W}_0^s(q_0)$. In order to define $\mathcal{S}_0$, we first denote by Σ the vertical half ray

$$\Sigma = \left\{(h,\gamma) \,|\, 0 < h < \infty, \gamma = \pi - \cos^{-1}\sqrt{2}\tfrac{\alpha}{\Gamma}\right\}$$

coming out of p_0 and going off to $h = \infty$. The vector field $(3.9)_0$ on Σ is perpendicular to Σ. Let a_0^1 and a_0^2 be the two points of intersection of each branch of $\mathcal{W}_0^s(q_0)$ with Σ, closest to q_0 in terms of the arclength. Let $\mathcal{L}_0$ be the union of the pieces of $\mathcal{W}_0^s(q_0)$ between q_0 and a_0^1 and between q_0 and a_0^2, respectively, and the segment of Σ connecting a_0^1 and a_0^2 (see Figure 3.5). The curve $\mathcal{L}_0$ divides the h-γ plane into two regions; $\mathcal{S}_0$ will be the one of these two regions which contains large negative values of h (see Figure 3.5). For small positive η we define two more regions, $\mathcal{R}_{0,\eta}$ and $\mathcal{S}_{0,\eta}$, which we obtain by excluding from $\mathcal{R}_0$ and $\mathcal{S}_0$, respectively, all the points whose distance from the loop $\mathcal{W}(q_0)$ is less than η. We will need these two regions in Section 4.6 when we derive conditions for the existence of an orbit homoclinic to p_ϵ.

Before defining the perturbed counterparts, $\mathcal{R}_\epsilon$ and $\mathcal{S}_\epsilon$, of the regions $\mathcal{R}_0$ and $\mathcal{S}_0$, respectively, we briefly discuss $\mathcal{W}_\epsilon^s(q_\epsilon)$ and $\mathcal{W}_\epsilon^u(q_\epsilon)$. We want to point out two of their aspects. First, these two manifolds will not intersect for nonzero α and ϵ. The reason for this is that they could only intersect along a separatrix loop whose existence is ruled out since we have proven that no closed orbits can exist in Π_c for nonzero α and ϵ. Second, the pieces of $\mathcal{W}_\epsilon^s(q_\epsilon)$ and $\mathcal{W}_\epsilon^u(q_\epsilon)$ between q_ϵ and their first

intersection points with Σ, call them a_ϵ^1, a_ϵ^2 and a_ϵ^3, stay uniformly $\mathcal{O}(\sqrt{\epsilon})$ close to their unperturbed counterparts. The precise mathematical details of this statement are stated in the two lemmas presented below.

In the first lemma we show that the local stable and unstable manifolds of the point q_ϵ stay close to those of q_0. The result will hold in the disk $\mathcal{D}_\delta$ around q_0 with radius δ. We formulate this statement in

Lemma 3.1. *For any small enough fixed positive δ, there exists a positive ϵ_0 such that for $0 \leq \epsilon \leq \epsilon_0$, the point q_ϵ is contained in $\mathcal{D}_\delta$. Moreover, the local stable and unstable manifolds, $W^s_{\epsilon,loc}(q_\epsilon)$ and $W^u_{\epsilon,loc}(q_\epsilon)$ of the point q_ϵ vary analytically with $\sqrt{\epsilon}$, and therefore, stay uniformly $\mathcal{O}(\sqrt{\epsilon})$ close to the local stable and unstable manifolds, $W^s_{0,loc}(q_0)$ and $W^u_{0,loc}(q_0)$, of the point q_0 inside $\mathcal{D}_\delta$ (see Figure 3.6).*

PROOF: The proof is very similar to the proof of Proposition 4.1 in Section 4.2; therefore we omit it.

We now want to show that we can extend this estimate outside $\mathcal{D}_\delta$ to show the $\mathcal{O}(\sqrt{\epsilon})$ closeness of the global stable and unstable manifolds of q_0 and q_ϵ as well. We make the argument only for the stable manifold. The argument for the unstable manifolds is exactly the same.

For this purpose, take any two points, say b_0 and b_ϵ, lying on $W^s_{0,loc}(q_0)$ and $W^s_{\epsilon,loc}(q_\epsilon)$, respectively, whose distance from each other is of $\mathcal{O}(\sqrt{\epsilon})$. Let $(h_0(\tau, b_0), \gamma_0(\tau, b_0))$ denote the solution of $(3.9)_0$ passing through b_0 at $\tau = 0$, and let $(h_\epsilon(\tau, b_\epsilon), \gamma_\epsilon(\tau, b_\epsilon))$ denote the solution of $(3.9)_\epsilon$ going through b_ϵ at $\tau = 0$. We

apply a Gronwall-type estimate (see Appendix C) to these two solutions to prove

Lemma 3.2. *For any finite positive T, there exists a positive constant $K(T)$, such that for every τ with $-T \leq \tau \leq 0$, the inequality*

$$\| (h_0(\tau, b_0), \gamma_0(\tau, b_0)) - (h_\epsilon(\tau, b_\epsilon), \gamma_\epsilon(\tau, b_\epsilon)) \| \leq K(T)\sqrt{\epsilon}$$

holds.

From this lemma, our statement about the uniform $\mathcal{O}(\sqrt{\epsilon})$ proximity of the pieces of $\mathcal{W}^s_\epsilon(q_\epsilon)$ and $\mathcal{W}^u_\epsilon(q_\epsilon)$ between q_ϵ and a^1_0, a^2_0 and a^3_0, respectively, to their unperturbed counterparts follows quite easily. We only need to take T on each of the orbits so long that both the perturbed and the unperturbed solution pass through Σ; in particular, this shows that a^1_ϵ is $\mathcal{O}(\sqrt{\epsilon})$ away from a^1_0 and that both a^2_ϵ and a^3_ϵ are $\mathcal{O}(\sqrt{\epsilon})$ away from a^2_0 (see Figure 3.7).

At this point we are ready to construct the regions $\mathcal{R}_\epsilon$ and $\mathcal{S}_\epsilon$. To construct $\mathcal{R}_\epsilon$, take $\mathcal{K}_\epsilon$ to be the union of the piece of $\mathcal{W}^s_\epsilon(q_\epsilon)$ between q_ϵ and a^2_ϵ, the piece of $\mathcal{W}^u_\epsilon(q_\epsilon)$ between q_ϵ and a^3_ϵ and the piece of Σ between a^2_ϵ and a^3_ϵ. The region $\mathcal{R}_\epsilon$ is the interior of the curve $\mathcal{K}_\epsilon$. To construct $\mathcal{S}_\epsilon$ take $\mathcal{L}_\epsilon$ to be the union of the pieces of $\mathcal{W}^s_\epsilon(q_\epsilon)$ between q_ϵ and a^1_ϵ and a^2_ϵ, respectively, and the part of Σ between a^1_ϵ and a^2_ϵ; $\mathcal{S}_\epsilon$ will be the one of the two regions, into which $\mathcal{L}_\epsilon$ divides the phase space, which contains large negative h (see Figure 3.7).

We are now in a position to determine the relative locations of the basins of attraction of the two sinks, p_ϵ and the sink near the origin. We do this in

Proposition 3.3. *For any positive η, there exists a small enough positive ϵ_0, such that for $0 < \epsilon < \epsilon_0$ and $\sqrt{\epsilon}\alpha$ positive the following two statements hold:*

(a) *$\mathcal{R}_\epsilon$ is a trapping region containing $\mathcal{R}_{0,\eta}$, which in turn contains p_ϵ. The point p_ϵ is the forward limit point of all the orbits in $\mathcal{R}_\epsilon$.*

(b) *$\mathcal{S}_\epsilon$ is a trapping region containing $\mathcal{S}_{0,\eta}$. All the orbits in $\mathcal{S}_\epsilon$ are attracted to the sink near the origin.*

PROOF: First note that the vector field $(3.9)_0$ is transverse to Σ between a_0^1 and a_0^2, and hence $(3.9)_\epsilon$ is transverse to the segments of Σ between a_ϵ^1 and a_ϵ^2 and a_ϵ^2 and a_ϵ^3, respectively, for small ϵ. To prove (a), note that the separatrix $\mathcal{W}(q_0)$ could split in two ways, either so that the vector field on the part of Σ between a_ϵ^2 and a_ϵ^3 points out of $\mathcal{R}_\epsilon$ or else it points into it. We will show that the first situation is impossible. Then we will show that the second situation implies (a).

We now show that the first situation is impossible since it violates the Poincaré-Bendixson theorem. We only need to reverse time in our problem and take any point in the interior of $\Sigma \cap \mathcal{K}_\epsilon$ as the initial condition for $(3.9)_\epsilon$. The solution will trace a semiorbit, which will be all contained in $\mathcal{R}_\epsilon$ since the time reversed vector field points everywhere into $\mathcal{R}_\epsilon$ on $\mathcal{K}_\epsilon$ or else is tangent to $\mathcal{K}_\epsilon$. The Poincaré-Bendixson theorem (see Coddington and Levinson, [1955], chapter 16) then states that this semiorbit must have a forward limit set inside $\mathcal{R}_\epsilon$, which must be either a fixed point or a closed orbit. Since there are no closed orbits in Π_c and no other fixed point inside $\mathcal{R}_\epsilon$ than p_ϵ, the only candidate must be p_ϵ. However, p_ϵ cannot be

the forward limit point of any semiorbit because it is a source for the time-reversed vector field. This shows that the first situation is impossible.

The second possibility is that the vector field points into $\mathcal{R}_\epsilon$ on $\Sigma \cap \mathcal{K}_\epsilon$. In this case, an argument similar to the previous one shows that all orbits in $\mathcal{R}_\epsilon$ approach p_ϵ in forward time, and every point in $\mathcal{R}_\epsilon$ will be attracted to p_ϵ. By Lemma 3.2, we also see that for every small positive η we can find a small enough positive ϵ_0 so that for all positive $\epsilon \leq \epsilon_0$, the curve $\mathcal{K}_\epsilon$ is contained in an η neighborhood of $\mathcal{W}(q_0)$. This concludes the proof of (a).

To prove (b), note that on $\mathcal{L}_0 \cap \Sigma$ the vector field $(3.9)_0$ points into $\mathcal{S}_0$. Therefore for small ϵ, the vector field still points into $\mathcal{S}_\epsilon$ on $\mathcal{L}_\epsilon \cap \Sigma$ and hence $\mathcal{S}_\epsilon$ is a trapping region. By Lemma 3.2, $\mathcal{L}_\epsilon$ is uniformly $\mathcal{O}(\sqrt{\epsilon})$ close to $\mathcal{L}_0$, hence for any given small positive η, there exists a small enough positive ϵ_0 (we may have to decrease ϵ_0 from part (a)) such that for $0 \leq \epsilon \leq \epsilon_0$, the region $\mathcal{S}_{0,\eta}$ is contained in $\mathcal{S}_\epsilon$. Since the only other attracting set in Π_c is the sink near the origin, (b) is proven.

Using the above described geometry and the conclusions of Proposition 3.3 one can easily show that both branches of $\mathcal{W}^s_\epsilon(q_\epsilon)$ now come from $h = +\infty$. To prove this just note that no solution on $\mathcal{W}^s_\epsilon(q_\epsilon)$ can be contained in either $\mathcal{R}_\epsilon$ or $\mathcal{S}_\epsilon$ as $\tau \to \infty$ since neither of these regions contains any sources and that all the orbits outside the resonance with $\frac{1}{2}|c|^2 = I > 1$ must spiral towards the resonance.

We remark that even though we treated α and Γ as fixed parameters, it is easy to see that all the features of the above described dynamics in Π_c change smoothly

as α and Γ are varied.

Having analyzed the dynamics in Π_c in detail, we briefly indicate the possible invariant sets on Π_c, which could in the full four-dimensional phase space possess orbits homoclinic to themselves, giving rise to chaotic behavior.

In the nondissipative case, the selection is very wide. The periodic orbits both outside and inside the resonance are likely candidates; so are both fixed points in the resonance, p_ϵ and q_ϵ.

In the dissipative case the choice narrows considerably; only p_ϵ and q_ϵ remain as the candidates. In the course of our work we will only investigate two of the above possibilities: the periodic orbits outside the resonance in the nondissipative case and the fixed point p_ϵ in the dissipative case in Sections 3.5 and 4.6, respectively.

3.3. Persistence of the Invariant Manifolds in the Phase Space. In this section we discuss the stable and unstable manifolds of annuli in Π_c, their persistence and smoothness. The intersections of these manifolds will play a crucial role in the subsequent analysis of homoclinic phenomena to be developed in later sections.

We first state the essential features of the unperturbed problem which we will need in this section. We will often refer back to the geometry established in Section 2.3 and the solutions of the homoclinic orbits calculated in Section 2.4. To begin with, we state the precise definition of the annulus $\mathcal{A}$ in Π_c. It is defined to be

$$\mathcal{A} = \{(x, y, I, \gamma) \,|\, x = y = 0, 0 < I_1 < I < I_2, 0 \leq \gamma \leq 2\pi\}\,.$$

Orbits in $\mathcal{A}$ are given by I = constant and $\gamma = (1-I)t + \gamma_0$. We see that $\mathcal{A}$ is invariant under the unperturbed flow of the equations $(3.3)_0$ since $\dot{I} = 0$. The annulus $\mathcal{A}$ is a manifold with boundary, the boundary, denoted $\partial\mathcal{A}$, being the two circles around the origin in Π_c with radii $I = I_1$ and $I = I_2$. Another way of saying that $\mathcal{A}$ is invariant is that the vector field $(3.3)_0$ is everywhere tangent to $\partial\mathcal{A}$. This situation is different from that of the closed boundaryless invariant manifolds, such as tori, and as a result we may expect to encounter problems at the boundary $\partial\mathcal{A}$ when we perturb our system.

Now we define local stable and unstable manifolds of the annulus $\mathcal{A}$. As we have shown in Section 2.3 (see Figures 2.1 and 2.2) we must have $k^2/2 < I_1 < I_2$ in order to guarantee that for every orbit in $\mathcal{A}$ there exist orbits outside Π_c approaching it exponentially in either forward or backward time. Having imposed this restriction, we choose a small positive δ and a neighborhood

$$U_\delta = \{(x, y, I, \gamma) \,|\, x^2 + y^2 < \delta^2, I_1 < I < I_2, 0 \leq \gamma \leq 2\pi\}$$

of the annulus $\mathcal{A}$ and define the *local stable manifold* $W^s_{0,loc}(\mathcal{A})$ (*local unstable manifold* $W^u_{0,loc}(\mathcal{A})$) of the annulus $\mathcal{A}$ as the sets of all the points in U_δ which approach $\mathcal{A}$ in forward (backward) time (see Figure 3.8). Note that $W^s_{0,loc}(\mathcal{A})$ and $W^u_{0,loc}(\mathcal{A})$ are only *locally invariant*, *i.e.*, they are spanned by orbits, but the orbits enter $W^s_{0,loc}(\mathcal{A})$ (leave $W^u_{0,loc}(\mathcal{A})$) through the part of its boundary at $x^2 + y^2 = \delta^2$.

We now extend the local stable and unstable manifolds of $\mathcal{A}$ to global ones. This can only be done if the radii I_1 and I_2 of the annulus $\mathcal{A}$ satisfy either $k^2/2 <$

$I_1 < I_2 < 4k^2$ or $4k^2 < I_1 < I_2$. The reason for this restriction is the heteroclinic connection between Π_c and Π_b at $I = 4k^2$, discussed in Section 2.3 (see Figure 2.1). If $\mathcal{A}$ included points with $I = 4k^2$, this heteroclinic connection would introduce an unwanted singularity in its stable and unstable manifolds; therefore, we exclude this possibility. For $\mathcal{A}$ satisfying the above restriction, we define its *stable manifold* $W_0^s(\mathcal{A})$ (*unstable manifold* $W_0^u(\mathcal{A})$) to be the set of all the points which approach $\mathcal{A}$ in forward (backward) time (see Figure 3.9). Obviously, $W_0^s(\mathcal{A})$ contains $W_{0,loc}^s(\mathcal{A})$ ($W_0^u(\mathcal{A})$ contains $W_{0,loc}^u(\mathcal{A})$). Note that $W_0^s(\mathcal{A})$ and $W_0^u(\mathcal{A})$ coincide and are equal to the homoclinic manifold $W(\mathcal{A})$ of $\mathcal{A}$. We recall from Section 2.3 that $W(\mathcal{A})$ is a three-dimensional manifold implicitly defined by the equation (2.10), which is $H(x, y, I) - H(0, 0, I) = 0$, with I restricted to the interval $I_1 < I < I_2$. Moreover, we can parametrize it explicitly by

$$W(\mathcal{A}) = \{(x(-t_0, I), y(-t_0, I), I, \gamma(-t_0, I, \gamma_0)) \,|\, t_0 \in \mathbf{R}, I_1 < I < I_2, 0 \leq \gamma \leq 2\pi\} \tag{3.16}$$

where $x(t, I)$, $y(t, I)$, and $\gamma(t, I, \gamma_0)$ are calculated using the transformations (2.12) and (2.19) from the formulas (2.24) when $k^2/2 < I_1 < I_2 < 4k^2$ and (2.23) when $4k^2 < I_1 < I_2$.

We want to make some comments concerning this parametrization. The variables I and γ_0 are used to pick out a particular homoclinic orbit in $W(\mathcal{A})$. Once I and γ_0 are fixed this orbit is parametrized by t_0. As we have shown in Section 2.4, every solution on the homoclinic orbit labeled by I and γ_0 is of the form

$(x(t-t_0,I), y(t-t_0,I), I, \gamma(t-t_0,I,\gamma_0))$ for some t_0. Therefore setting $t=0$ and varying t_0 in $\mathbf{R}$ exhausts all the possible initial conditions, *i.e.*, all the possible points on that orbit. Every point in $W(\mathcal{A})$ can hence be uniquely described by the parametrization (3.16). We also remark that this parametrization is different from the more standard one used in Wiggins [1988]; the differences will be pointed out in Appendix E where we discuss the advantages and shortcomings of both parametrizations.

The homoclinic manifold $W(\mathcal{A})$ is again invariant under the flow of $(3.3)_0$ for the same reason as $\mathcal{A}$ is, *i.e.*, the vector field $(3.3)_0$ is everywhere tangent to its boundary $\partial\mathcal{A}$ which is described by either the equations

$$H(x,y,I) - H(0,0,I) = 0$$

$$I = I_1 \text{ or } I = I_2$$

or by the parametrization

$$\partial W(\mathcal{A})$$

$$= \{(x(-t_0,I), y(-t_0,I), I, \gamma(-t_0,I,\gamma_0)) \mid t_0 \in \mathbf{R}, I = I_1 \text{ or } I = I_2, 0 \leq \gamma_0 \leq 2\pi\} .$$

We can also see from the formulas (2.23) and (2.24) that x and y coordinates on any orbit starting in $W^s_{0,loc}(\mathcal{A})$ ($W^u_{0,loc}(\mathcal{A})$) at $t=0$ shrink to zero at the exponential rate $e^{-k\sqrt{2I-k^2}|t|}$ for positive (negative) time, while the I coordinate remains unchanged.

Having reviewed the features of the unperturbed problem essential to the analysis of the perturbed problem, we now perform this latter task. In the perturbed problem, the time evolution of the coordinates I and γ on $\mathcal{A}$ is given by the equations

$(3.8)_\epsilon$, which have been extensively studied in the previous section. The annulus $\mathcal{A}$ is still spanned by orbits, since it is a subset of the invariant plane Π_c. However, it may no longer be invariant since orbits may enter or leave it through its boundary $\partial\mathcal{A}$ and is hence only locally invariant. Typically, we would expect a point in $\mathcal{A}$ which is $\mathcal{O}(1)$ away from $\partial\mathcal{A}$ to have entered $\mathcal{A}$ or to leave $\mathcal{A}$ on a time scale of $\mathcal{O}(\frac{1}{\epsilon})$ if it does not remain in $\mathcal{A}$ forever.

The fact that $\mathcal{A}$ is only locally invariant causes some difficulties with defining the perturbed stable and unstable manifolds of $\mathcal{A}$. Usually, stable and unstable manifolds of invariant sets are defined as the sets of points which approach the invariant set for all forward and backward times, respectively. But in our problem, $\mathcal{A}$ is not invariant and hence an orbit in $\mathcal{A}$ off of Π_c may well approach an orbit which ultimately leaves $\mathcal{A}$. The question which then arises is whether to leave such orbits out of our definition of the stable and unstable manifolds of $\mathcal{A}$ or include them in it. The answer to this question is that it is more natural to include them. The reason for this choice is that orbits of precisely this type will span the three-dimensional manifolds $W^s_{\epsilon,loc}(\mathcal{A})$ and $W^u_{\epsilon,loc}(\mathcal{A})$ (see Figure 3.9) which, as we will show shortly, will play an important role in our analysis. We state their existence and properties rigorously in Proposition 3.4. These manifolds are graphs over $W^s_{0,loc}(\mathcal{A})$ and $W^u_{0,loc}(\mathcal{A})$, respectively, which vary smoothly with ϵ and other parameters, and also intersect along $\mathcal{A}$.

The boundaries $\partial W^s_{\epsilon,loc}(\mathcal{A})$ and $\partial W^u_{\epsilon,loc}(\mathcal{A})$ of these two manifolds consist of

three parts: the *bottom* at $I = I_1$, the *top* at $I = I_2$ and the *sides* at $x^2 + y^2 = \delta^2$. The common feature of $W^s_{\epsilon,loc}(\mathcal{A})$ and $W^s_{0,loc}(\mathcal{A})$ ($W^u_{\epsilon,loc}(\mathcal{A})$ and $W^u_{0,loc}(\mathcal{A})$) is that the x-y dynamics for small ϵ remain almost the same as in the unperturbed case. In particular, orbits still only enter $W^s_{\epsilon,loc}(\mathcal{A})$ (leave $W^u_{\epsilon,loc}(\mathcal{A})$) through its sides and the x and y coordinates contract at an exponential rate faster than $e^{-\kappa|t|}$ with $\kappa = k\sqrt{2I_1 - k^2} + \mathcal{O}(\epsilon)$, for positive (negative) times as long as their orbit stays in $W^s_{\epsilon,loc}(\mathcal{A})$ ($W^u_{\epsilon,loc}(\mathcal{A})$). However, the I-γ dynamics may change the orbits drastically; in particular, they may either enter or leave $W^s_{\epsilon,loc}(\mathcal{A})$ and $W^u_{\epsilon,loc}(\mathcal{A})$ through both their top and and their bottom (see Figure 3.8). If an orbit does stay in $W^s_{\epsilon,loc}(\mathcal{A})$ for all positive times (in $W^u_{\epsilon,loc}(\mathcal{A})$ for all negative times), though, it must in forward (backward) time approach an invariant set in $\mathcal{A}$ because of the exponential contraction of its x and y coordinates.

We now define $W^s_\epsilon(\mathcal{A})$ ($W^u_\epsilon(\mathcal{A})$) as the manifolds obtained by letting all the points on the above defined sides of $W^s_{\epsilon,loc}(\mathcal{A})$ ($W^u_{\epsilon,loc}(\mathcal{A})$) evolve under the flow of $(3.3)_\epsilon$ in backward (forward) time, and then taking the union with $W^s_{\epsilon,loc}(\mathcal{A})$ ($W^u_{\epsilon,loc}(\mathcal{A})$). This situation is shown in Figure 3.9. The manifolds $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ need not be invariant; indeed, orbits may enter and leave them through the above defined top and bottom of $W^s_{\epsilon,loc}(\mathcal{A})$ and $W^u_{\epsilon,loc}(\mathcal{A})$, respectively. Therefore, they need not be the stable and unstable manifolds of any object, in the usual sense of the terminology. But they will serve a good purpose in our problem, hence we will call $W^s_\epsilon(\mathcal{A})$ the *stable manifold* ($W^u_\epsilon(\mathcal{A})$ the *unstable manifold*) of the annulus

$\mathcal{A}$, and $W^s_{\epsilon,loc}(\mathcal{A})$ the *local stable manifold* ($W^u_{\epsilon,loc}(\mathcal{A})$ the *local unstable manifold*) of $\mathcal{A}$. With this discussion behind us, we only need to state the precise existence theorem for $W^s_{\epsilon,loc}(\mathcal{A})$ and $W^u_{\epsilon,loc}(\mathcal{A})$. It goes as follows.

Proposition 3.4. *There exist small positive ϵ_0 and δ_0 such that for $-\epsilon_0 \leq \epsilon \leq \epsilon_0$ and $0 < \delta \leq \delta_0$ there exist locally invariant manifolds $W^s_{\epsilon,loc}(\mathcal{A})$ and $W^u_{\epsilon,loc}(\mathcal{A})$ in U_δ, possessing the following properties:*

(1) *They vary in a C^∞ manner with ϵ and the other parameters in the problem.*

(2) *For $\epsilon = 0$, they coincide with the unperturbed local stable and unstable manifolds $W^s_{0,loc}(\mathcal{A})$ and $W^u_{0,loc}(\mathcal{A})$, respectively.*

(3) *They intersect along $\mathcal{A}$.*

(4) *Any point which starts at $t = 0$ on the sides of $W^s_{\epsilon,loc}(\mathcal{A})$ ($W^u_{\epsilon,loc}(\mathcal{A})$) will approach $\mathcal{A}$ in forward (backward) time at an exponential rate at least as fast as $e^{-\kappa|t|}$, where $\kappa = k\sqrt{2I_1 - k^2} + \mathcal{O}(\epsilon)$, as long as it stays in $W^s_{\epsilon,loc}(\mathcal{A})$ ($W^u_{\epsilon,loc}(\mathcal{A})$).*

PROOF: The proof of the first three statements is the same as in Wiggins [1988], p. 354. The last proposition follows from Fenichel [1974], Theorem 3, using very similar techniques as Wiggins [1988], p. 354, does in order to deduce his theorem from Fenichel [1971]. The details of the proof are long, tedious and involve concepts which will not be illuminating for the rest of our discussion; hence we do not present them here.

Following this proposition, let us make two comments. First note that the exponential rate $e^{-k\sqrt{2I_1-k^2}|t|}$ is the slowest one at which orbits in $W^s_{0,loc}(\mathcal{A})$ ($W^u_{0,loc}(\mathcal{A})$) approach $\mathcal{A}$ in forward (backward) time. Second, note that point (3) implies $W^s_{\epsilon,loc}(\mathcal{A})$ and $W^s_{0,loc}(\mathcal{A})$ ($W^u_{\epsilon,loc}(\mathcal{A})$ and $W^u_{0,loc}(\mathcal{A})$) to be uniformly $\mathcal{O}(\epsilon)$ close to each other in C^r topology for any r, *i.e.*, they and all their derivatives up to and including r-th are uniformly $\mathcal{O}(\epsilon)$ close.

In the next few sections we will learn the use of the manifolds $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$. In particular, we will see that their intersections (if they intersect at all) contain all the orbits asymptotic to any invariant set in $\mathcal{A}$ in both forward and backward time.

3.4. Intersections of the Manifolds $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ and the Nature of the Orbits Contained in Them. In this section we want to develop a computable distance measurement for the distance between $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$. Particularly interesting for us will be the points at which this distance vanishes, *i.e.*, where $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ intersect. If such a point exists, the invariance of the two manifolds forces the orbit on which this point lies to be contained in the intersection as well. The study of this type of orbit will occupy the remaining part of this work. The interesting feature of orbits lying in the intersection of the manifolds $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ is that all orbits homoclinic to fixed points or periodic orbits in $\mathcal{A}$ are among them. And, in turn, it is these orbits which are responsible for the possible existence of nearby chaotic invariant sets.

We note, however, that not all the orbits contained in the intersections of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ need to be homoclinic. In fact, in the dissipative case, we actually expect most of them to enter $W^u_{\epsilon,loc}(\mathcal{A})$ through its top, pass through the intersection to $W^s_{\epsilon,loc}(\mathcal{A})$ and leave $W^s_{\epsilon,loc}(\mathcal{A})$ through its bottom (see Figure 3.10). We can speculate that such an orbit will, in general, prior to entering $W^u_{\epsilon,loc}(\mathcal{A})$, asymptote in backward time to an orbit in Π_c spiraling inwards from $I = +\infty$, and, after leaving $W^s_{\epsilon,loc}(\mathcal{A})$, drift off to the sink at the origin. Not being asymptotic to a recurrent set, such an orbit will clearly not give rise to any kind of chaos. We will therefore spend a lot of effort in the forthcoming sections to determine what precisely happens to an orbit when it is contained in an intersection of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$.

To get a better picture of the difficulties encountered in our problem, we want to contrast it with the usual case to which the Melnikov function calculation applies: the periodically perturbed planar Hamiltonian system with a separatrix loop connecting a saddle type fixed point to itself (see Wiggins [1989]).

We can regard the suspended phase space (with time as the extra variable) as a cartesian product of the plane $\mathbf{R}^2$ and a circle. The unperturbed fixed point is then a trivial hyperbolic periodic orbit, and the saddle connection loop a pinched invariant torus (See Figure 3.11) foliated by orbits homoclinic to the periodic orbit. The word hyperbolic in this case means that the periodic orbit is isolated, that at every point on the orbit its local stable and unstable manifolds intersect transversely, and that

points on these two manifolds approach the periodic orbit at an exponential rate in forward and backward times, respectively. Pieces of the periodic orbit's global stable and unstable manifolds coincide along the pinched invariant torus.

Under perturbation the trivial periodic orbit persists as a real periodic orbit, its stable and unstable manifolds do not coincide any more, and if they intersect, the situation looks like Figure 3.11. Two important features should be noted. The first is that the hyperbolic periodic orbit and its local stable and unstable manifolds persist, under perturbation, essentially unchanged. The second feature is that the orbits along which perturbed global stable and unstable manifolds of the periodic orbit intersect are the survivors of the unperturbed pinched torus of homoclinic orbits.

Neither of these features is present in our system. The plane Π_c is invariant, but does not, as a whole, possess a simple homoclinic manifold. The annulus $\mathcal{A}$ in Π_c, however, does possess a simple homoclinic manifold but may not be invariant in the perturbed system. The periodic orbits in the nondissipative case have both of the above properties, but are not isolated in the full four-dimensional phase space and hence not hyperbolic. However, we will see below that they become hyperbolic as soon as we restrict ourselves to the three-dimensional level sets of the Hamiltonian H_ϵ. And finally, the fixed points p_ϵ and q_ϵ, being associated with the most complex dynamics in our problem, in their unperturbed form, p_0 and q_0, usually do not even possess orbits homoclinic to themselves, but are connected via heteroclinic orbits

to other fixed points on the circle (see Section 2.5). Moreover, p_ϵ and q_ϵ lie in the middle of a resonance band, a structure created by the perturbation and drastically different from the unperturbed circle of fixed points which stood in its place.

This is why we consider it necessary to approach the problem in two stages, the first one being to look for intersections of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ using the Melnikov technique, and the second one to concentrate on specific orbits contained in those intersections and study their nature, usually using other techniques of a more geometric type. The particular technique we will use will depend on the situation we will be looking at.

In the case of nondissipative nonresonant periodic orbits, we essentially recover a situation similar to a periodically forced planar Hamiltonian system with a separatrix loop discussed above when we restrict ourselves to any one of the three-dimensional level surfaces of the Hamiltonian H_ϵ. The reason for this is that on every level surface of H_ϵ there is only one periodic orbit and it is then easy to show its hyperbolicity inside that level surface. In fact, in certain simple situations (but not always in ours), a method can be applied which explicitly reduces a perturbed two degree of freedom system with the same unperturbed geometry as ours to a periodically forced planar system using the fact that the phase space is foliated by the level surfaces of the Hamiltonian (see Holmes and Marsden [1982]). In the next section, we will instead apply more geometric methods somewhat similar to those in Wiggins [1988], which will show how the intersections of the manifolds $W^s_\epsilon(\mathcal{A})$

and $W^u_\epsilon(\mathcal{A})$ are foliated by the level sets of H_ϵ into families of orbits homoclinic to the family of periodic orbits in $\mathcal{A}$ to obtain a more elegant version of the same type of result.

The second situation we will deal with, the orbit homoclinic to p_ϵ in the dissipative case, will require completely different techniques which are of a singular perturbation nature. As we will see in Chapter 4, this orbit will have to be constructed from two pieces and we will only be able to calculate the existence of one of the pieces with the use of the Melnikov function.

We now begin by investigating the intersections of the stable and unstable manifolds, $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$, of the annulus $\mathcal{A}$. We have seen that in the unperturbed case the stable and unstable manifolds, $W^s_0(\mathcal{A})$ and $W^u_0(\mathcal{A})$, of $\mathcal{A}$ coincide and form the homoclinic manifold $W(\mathcal{A})$. In the perturbed case, $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ cannot be expected to coincide any more. Instead, we expect them to either intersect along two-dimensional surfaces or else not intersect at all. The number two comes from the dimension count for transversal intersections of manifolds: The dimension of the intersection equals the sum of the dimensions of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ minus the dimension of the ambient space $\mathbf{R}^4$. We will measure the distance between $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ by using the unperturbed homoclinic manifold $W(\mathcal{A})$ as a framework on which to develop a geometric perturbation theory. We therefore first need a parametrization of $W(\mathcal{A})$ in order to describe the coordinates of points on $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ near $W(\mathcal{A})$. We then need a coordinate system at each point of $W(\mathcal{A})$

in order to measure how $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ split near that point. We also want to determine the number of independent directions in which we need to measure the distance and finally develop a computable expression for the distance measurement.

We recall that $W(\mathcal{A})$ is parametrized as in (3.16) by

$$W(\mathcal{A}) = \{(x(-t_0, I), y(-t_0, I), I, \gamma(-t_0, I, \gamma_0)) \mid t_0 \in \mathbf{R}, I_1 < I < I_2, 0 \leq \gamma_0 \leq 2\pi\}$$

where x, y and γ are obtained from the solutions calculated in section 2.4. Various features of this parametrization have been discussed in the previous section.

At each point $a = (x(-t_0, I), y(-t_0, I), I, \gamma(-t_0, I, \gamma_0))$ in $W(\mathcal{A})$ we attach the following coordinate system: The tangent $\mathbf{t}_{t_0}$ to the orbits on $W(\mathcal{A})$ at a, the normal $\mathbf{n}_W$ to $W(\mathcal{A})$ at a, and the two vectors, $\hat{I}$ and $\hat{\gamma}$, the unit vectors in I and γ directions. In the x-y-I-γ coordinates, the components of these vectors are

$$\mathbf{t}_{t_0}(x, y, I) = \left(\frac{\partial H}{\partial y}(x, y, I), -\frac{\partial H}{\partial x}(x, y, I), 0, -\frac{\partial H}{\partial I}(x, y, I)\right) \tag{3.17}$$

$$\mathbf{n}_W(x, y, I) = \left(\frac{\partial H}{\partial x}(x, y, I), \frac{\partial H}{\partial y}(x, y, I), \frac{\partial H}{\partial I}(x, y, I) - \frac{\partial H}{\partial y}(0, 0, I), 0\right) \tag{3.18}$$

$$\hat{I} = (0, 0, 1, 0) \tag{3.19}$$

$$\hat{\gamma} = (0, 0, 0, 1), \tag{3.20}$$

where, for brevity, we omit the arguments $(-t_0, I)$ in the x and y coordinates. The determinant of the matrix composed of these vectors is easily seen to be $\left(\frac{\partial H}{\partial x}\right)^2 + \left(\frac{\partial H}{\partial y}\right)^2$ and since both of these derivatives vanish only on $W(\mathcal{A})$ for $x = y = 0$, *i.e.*, on Π_c we see that $\mathbf{t}_{t_0}$, $\mathbf{n}_W$, $\hat{I}$, and $\hat{\gamma}$ indeed form a coordinate system at every point

a in $W(\mathcal{A})$ except at Π_c. There we will not need it, since we do not measure the splitting distance there.

We claim that we only need to measure the distance between $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ in one direction. We will make this explicit as we set up the measurement.

The distance measurement is set up as follows. At each point $\mathcal{A}$ in $W(\mathcal{A})$ we will measure the distance between $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ near a along the normal $\mathbf{n}_W(x(-t_0, I), y(-t_0, I), I)$ to $W(\mathcal{A})$ at a. By its very definition, $\mathbf{n}_W$ intersects $W(\mathcal{A})$ transversely at a, and therefore it is also transverse to $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ for small enough ϵ. Hence, a line through a parallel to $\mathbf{n}_W(x(-t_0, I), y(-t_0, I), I)$ intersects $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ near a transversely in at least one point each.

However, $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ may fold back on themselves and therefore intersect $\mathbf{n}_W$ at more than one point. (See Figure 3.14. for a simplified situation.) We will denote the two points between which we want to measure the distance by a^s_ϵ and a^u_ϵ, where a^s_ϵ lies in $W^s_\epsilon(\mathcal{A})$ and a^u_ϵ in $W^u_\epsilon(\mathcal{A})$. We will choose a^s_ϵ to be the point of intersection of $\mathbf{n}_W$ and $W^s_\epsilon(\mathcal{A})$ whose trajectory needs the least amount of time to reach $W^s_{\epsilon,loc}(\mathcal{A})$. We choose a^u_ϵ in an analogous manner. This choice agrees with our intuition and can be shown to be the only mathematically correct one. (For a more complete description see Wiggins [1988] or Wiggins [1989].)

Summarizing the above, the line parallel to $\mathbf{n}_W(x(-t_0, I), y(-t_0, I), I)$ at each point $a = (x(-t_0, I), y(-t_0, I), I, \gamma(-t_0, I, \gamma_0))$ on $W(\mathcal{A})$ and going through a intersects $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ in two points, a^s_ϵ and a^u_ϵ, which can be picked in a unique

manner. If $a_\epsilon^s = a_\epsilon^u$, the manifolds $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$ will intersect at that point. It now follows from the previous sentence that at each point a a measurement in only one direction, that of $\mathbf{n}_W$, will determine whether $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$ intersect or not.

We now derive a computable expression for the distance between a_ϵ^s and a_ϵ^u. Since a_ϵ^s and a_ϵ^u lie on $\mathbf{n}_W$, the vector $a_\epsilon^s - a_\epsilon^u$ is parallel to $\mathbf{n}_W$, and hence

$$\left| \frac{\langle \mathbf{n}_W, a_\epsilon^u - a_\epsilon^s \rangle}{\|\mathbf{n}_W\|} \right| = \|a_\epsilon^u - a_\epsilon^s\| .$$

We therefore define the signed distance between a_ϵ^s and a_ϵ^u as

$$\begin{aligned} d(a;\epsilon) &= d(t_0, I, \gamma_0; k, \alpha, \Gamma; \epsilon) \\ &= \frac{\langle \mathbf{n}_W(x,y,I), a_\epsilon^u - a_\epsilon^s \rangle}{\|\mathbf{n}_W(x,y,I)\|} \end{aligned} \tag{3.17}$$

with $(x, y, I) = (x(-t_0, I), y(-t_0, I), I)$. The reason we have chosen this expression over $\|a_\epsilon^u - a_\epsilon^s\|$ is the fact that we can explicitly calculate the $\mathcal{O}(\epsilon)$ approximation to it and then justify that the distance vanishes near the points where the approximation vanishes. If for some $(t_0, I, \gamma_0; k, \alpha, \Gamma; \epsilon)$ we have $d(t_0, I, \gamma_0; k, \alpha, \Gamma; \epsilon) = 0$ then $a_\epsilon^s = a_\epsilon^u$ there and $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$ intersect.

We now Taylor expand $d(a;\epsilon)$ to get

$$\begin{aligned} d(t_0, I, \gamma_0; k, \alpha, \Gamma; \epsilon) = {} & d(t_0, I, \gamma_0; k, \alpha, \Gamma; 0) \\ & + \epsilon \frac{\partial d}{\partial \epsilon}(t_0, I, \gamma_0; k, \alpha, \Gamma; 0) + \mathcal{O}(\epsilon^2) \end{aligned}$$

where

$$d(t_0, I, \gamma_0; k, \alpha, \Gamma; 0) = 0$$

and

$$\frac{\partial d}{\partial \epsilon}(t_0, I, \gamma_0; k, \alpha, \Gamma; 0) = \frac{\langle \mathbf{n}_W(x(-t_0, I), y(-t_0, I), I), \frac{\partial a^u_\epsilon}{\partial \epsilon}|_{\epsilon=0} - \frac{\partial a^s_\epsilon}{\partial \epsilon}|_{\epsilon=0} \rangle}{\|\mathbf{n}_W(x(-t_0, I), y(-t_0, I), I)\|}.$$

The numerator of the last expression is called the *Melnikov function*, and we will denote it by $M(I, \gamma_0, k; \alpha, \Gamma)$. It turns out to be computable explicitly using only the unperturbed solutions and the perturbed vector field. In Appendix E we will show that it has the form

$$M(I, \gamma_0, k; \alpha, \Gamma) = \int_{-\infty}^{\infty} \langle \mathbf{n}_W(x(t, I), y(t, I), I), g(x(t, I), y(t, I), I, \gamma(t, I, \gamma_0), \alpha, \Gamma) \rangle \, dt \quad (3.18)$$

where $g = (g^x, g^y, g^I, g^\gamma)$ is the perturbation part of the vector field $(3.5)_\epsilon$ divided by ϵ.

Note that the Melnikov function does not contain t_0 explicitly and that, therefore, all the explicit dependence on t_0 of the $\mathcal{O}(\epsilon)$ of $d(a; \epsilon)$ comes from its denominator. This phenomenon occurs because of our choice of parametrization of $W(\mathcal{A})$ by orbits, and corresponds to the fact that the Melnikov function is constant along the unperturbed orbits. We will show this fact when we derive the Melnikov function in Appendix E. We will also elaborate more on its consequences for the intersections of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ after we state the theorem concerning the connection between zeros of the Melnikov function and these intersections.

We expect the zeros of the Melnikov function to approximate the zeros of the signed distance between $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$. The following proposition will confirm our expectations.

Proposition 3.5. *If for* $I = \bar{I}$, $\gamma_0 = \bar{\gamma}_0$, $k = \bar{k}$ $\alpha = \bar{\alpha}$, *and* $\Gamma = \bar{\Gamma}$ *the following statements are true:*

(1) $M(\bar{I}, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma}) = 0$, *and*

(2) *At least one of the first partial derivatives of* $M(I, \gamma_0, k; \alpha, \Gamma)$ *is nonzero at* $I = \bar{I}$, $\gamma_0 = \bar{\gamma}_0$, $k = \bar{k}$ $\alpha = \bar{\alpha}$, and $\Gamma = \bar{\Gamma}$,

then for ϵ *sufficiently small,* $W^s_\epsilon(\mathcal{A})$ *and* $W^u_\epsilon(\mathcal{A})$ *intersect near* $(t_0, \bar{I}, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma})$ *for every* $t_0 \in \mathbf{R}$.

Furthermore, if either $\frac{\partial M}{\partial I}(\bar{I}, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma})$ *or* $\frac{\partial M}{\partial \gamma_0}(\bar{I}, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma})$ *is nonzero, the intersection is transverse for every* t_0.

PROOF: Define

$$\tilde{d}(t_0, I, \gamma_0; k, \alpha, \Gamma; \epsilon) = \frac{1}{\epsilon} \|\mathbf{n}_W(x(-t_0, I), y(-t_0, I), I)\| \, d(t_0, I, \gamma_0; k, \alpha, \Gamma; \epsilon)$$

$$= M(I, \gamma_0, k; \alpha, \Gamma) + \mathcal{O}(\epsilon) .$$

We have $\tilde{d}(t_0, \bar{I}, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma}; 0) = 0$ and one of the derivatives, say for definiteness $\frac{\partial \tilde{d}}{\partial I}(t_0, \bar{I}, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma}; 0) = \frac{\partial M}{\partial I}(\bar{I}, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma})$ nonzero. Therefore, for small ϵ, by the implicit function theorem, we can express $I = I(t_0, \gamma_0; k, \alpha, \Gamma; \epsilon)$ with $I(t_0, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma}; 0) = \bar{I}$ and $\tilde{d}(t_0, I(t_0, \gamma_0; k, \alpha, \Gamma; \epsilon), \gamma_0; k, \alpha, \Gamma; \epsilon) = 0$. Since $\|\mathbf{n}_W\|$ is nonzero away from Π_c, this shows that $d(t_0, I(t_0, \gamma_0; k, \alpha, \Gamma; \epsilon), \gamma_0; k, \alpha, \Gamma; \epsilon) = 0$, hence $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ intersect there. If the derivative on another variable is nonzero, we repeat the same proof with I replaced by that variable.

To prove the transversality of the intersection, we note that as t_0, I and γ_0 run through any compact subset of $\mathbf{R} \times (I_1, I_2) \times [0, 2\pi]$, the corresponding point

a will describe a compact three-dimensional subset of $W(\mathcal{A})$, and a_ϵ^s and a_ϵ^u will describe compact three-dimensional subsets of $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$, respectively. In particular, this shows that t_0, I, and γ_0 can also be used to parametrize $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$ near the given point $a \in W(\mathcal{A})$. Therefore, the tangent spaces of $W_\epsilon^s(\mathcal{A})$ at a_ϵ^s and $W_\epsilon^u(\mathcal{A})$ at a_ϵ^u are spanned by the triples of vectors

$$\frac{\partial a_\epsilon^s}{\partial t_0}, \qquad \frac{\partial a_\epsilon^s}{\partial I}, \qquad \frac{\partial a_\epsilon^s}{\partial \gamma_0}$$

and

$$\frac{\partial a_\epsilon^u}{\partial t_0}, \qquad \frac{\partial a_\epsilon^u}{\partial I}, \qquad \frac{\partial a_\epsilon^u}{\partial \gamma_0},$$

respectively. For small ϵ, these tangent spaces are still transverse to $\mathbf{n}_W$. At an intersection point we have $a_\epsilon^s = a_\epsilon^u$. Therefore, we can construct the vector sum of the two tangent spaces and if it contains a vector which has a component in the direction of $\mathbf{n}_W$, it must contain $\mathbf{n}_W$, and is therefore four-dimensional, which implies that $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$ intersect transversely there.

If $\frac{\partial M}{\partial I}$ is nonzero at a, we look at $\langle \frac{\partial a_\epsilon^u}{\partial I} - \frac{\partial a_\epsilon^s}{\partial I}, \mathbf{n}_W \rangle$ at that point. We have

$$\begin{aligned}
&\left\langle \frac{\partial a_\epsilon^u}{\partial I} - \frac{\partial a_\epsilon^s}{\partial I}, \mathbf{n}_W \right\rangle \\
=&\left\langle \frac{\partial a_\epsilon^u}{\partial I} - \frac{\partial a_\epsilon^s}{\partial I}, \mathbf{n}_W \right\rangle + \left\langle a_\epsilon^u - a_\epsilon^s, \frac{\partial \mathbf{n}_W}{\partial I} \right\rangle \\
=&\frac{\partial}{\partial I}\langle a_\epsilon^u - a_\epsilon^s, \mathbf{n}_W \rangle ,
\end{aligned}$$

the second equality being the result of the fact that $a_\epsilon^u - a_\epsilon^s = 0$. But

$$\frac{\partial}{\partial I}\langle a_\epsilon^u - a_\epsilon^s, \mathbf{n}_W \rangle = \epsilon \frac{\partial M}{\partial I}(\bar{I}, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma}) + \mathcal{O}(\epsilon^2),$$

and for small enough values of ϵ this is nonzero. Hence the vector $\frac{\partial a^u_\epsilon}{\partial I} - \frac{\partial a^s_\epsilon}{\partial I}$ has a nonzero component in the direction of $\mathbf{n}_W$ and the transversality is proved.

If $\frac{\partial M}{\partial \gamma_0}$ is nonzero at $\mathcal{A}$, we find in the same way that

$$\left\langle \frac{\partial a^u_\epsilon}{\partial \gamma_0} - \frac{\partial a^s_\epsilon}{\partial \gamma_0}, \mathbf{n}_W \right\rangle = \epsilon \frac{\partial M}{\partial \gamma_0}(\bar{I}, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma}) + \mathcal{O}(\epsilon^2)$$

which again proves the transversality. This concludes the proof of Proposition 3.5.

If we have $M(\bar{I}, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma}) = 0$, and again for definiteness assume

$$\frac{\partial M}{\partial I}(\bar{I}, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma})$$

to be nonzero, we have by the implicit function theorem

$$I(t_0, \bar{\gamma}_0; \bar{k}, \bar{\alpha}, \bar{\Gamma}; \epsilon) = \bar{I} + \mathcal{O}(\epsilon) \tag{3.19}$$

for each fixed t_0. However, the error term need not be bounded as to $t_0 \to \pm\infty$. In fact, we have argued before that a typical intersection orbit in the perturbed case spirals inward from $I = +\infty$, staying close to Π_c until it hits the top of $W^u_{\epsilon,loc}(\mathcal{A})$. It then leaves $W^u_{\epsilon,loc}(\mathcal{A})$ through its side and follows an unperturbed orbit uniformly $\mathcal{O}(\epsilon)$ close until it enters $W^s_{\epsilon,loc}(\mathcal{A})$. It is near this piece of the orbit that (3.19) is uniformly valid. The orbit then typically leaves $W^s_{\epsilon,loc}(\mathcal{A})$ through its bottom.

This explains why the Melnikov function is independent of t_0: The intersection surface of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ can be parametrized by t_0, I and γ_0 where in the zeroth order in ϵ, I and γ_0 are connected by the equation $M(I, \gamma_0, k; \alpha, \Gamma) = 0$. This equation picks out the orbits on a two-dimensional submanifold of $W(\mathcal{A})$, which is

$\mathcal{O}(\epsilon)$ close to the intersection surface of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ outside of $W^s_{\epsilon,loc}(\mathcal{A})$ and $W^u_{\epsilon,loc}(\mathcal{A})$. The parameter to t_0 then parametrizes each orbit on that surface.

We must also remark that the intersection surfaces vary smoothly with k, α and γ.

We now present the concrete form of the Melnikov function in our problem. It is given by

$$M(I, \gamma_0, k; \alpha, \Gamma) = \Gamma F(I, k) \cos \gamma_0 + \alpha G(I, k) .$$

The computation is performed in Appendix F. We see that if $F(I, k)$ and $G(I, k)$ are nonzero and

$$\frac{\alpha}{\Gamma} \leq \frac{F(I, k)}{G(I, k)}$$

for some k and $I_1 \leq I \leq I_2$, then we find an intersection surface of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ near the surface parametrized by I and t_0, with

$$\gamma_0 = \cos^{-1} \frac{\alpha\, G(I, k)}{\Gamma\, F(I, k)} .$$

3.5. Orbits Homoclinic to Periodic Orbits in Π_c Away from the Resonance Band for Nondissipative Perturbations. We saw in the previous section how to compute when $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ intersect transversely at a point and what this implies for the orbit going through that intersection point. The situation we found especially interesting was when the orbit did not leave $W^s_{\epsilon,loc}(\mathcal{A})$ and $W^u_{\epsilon,loc}(\mathcal{A})$, since in that case it had to approach a fixed point or a closed invariant curve inside $\mathcal{A}$ in forward and possibly another one in backward time. In Section

3.2 we have shown that there are many closed, in fact periodic, orbits in Π_c when α is zero.

In this section we look at the intersections of the stable and unstable manifolds of the nonresonant periodic orbits in Π_c. We want to show that for the periodic orbits contained in $\mathcal{A}$, the existence of such intersections can easily be inferred from the existence of intersections of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$. All we need to do is to carefully examine the transversality of the intersections of the level sets of the Hamiltonian $H_\epsilon(x, y, I, \gamma)$ with $\mathcal{A}$, $W^s_\epsilon(\mathcal{A})$, and $W^u_\epsilon(\mathcal{A})$, and also with $W^s_\epsilon(\mathcal{A}) \cap W^u_\epsilon(\mathcal{A})$.

It will be a simple matter to show that all these intersections are transverse, and from this we will immediately deduce that when we restrict ourselves to a level set of $H_\epsilon(x, y, I, \gamma)$, transverse intersections of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ become transverse intersections of the stable and unstable manifolds of the periodic orbits in $\mathcal{A}$. We will therefore show that, under certain conditions which can easily be calculated, the periodic orbits in $\mathcal{A}$ possess transverse homoclinic orbits. In conclusion, we will discuss the implications of these transverse homoclinic orbits for the existence of nearby chaotic dynamics.

We start with a simple observation that outside the resonance band, each periodic orbit can be at least locally uniquely labeled by the value of the Hamiltonian H_ϵ on it. We can either calculate this fact directly by expressing c_1 in terms of c_2 and H_ϵ from the equation $(3.7)_\epsilon$, or else from the Hamiltonian $(3.13)_\epsilon$, which is $H_\epsilon = \frac{1}{2}I^2 - I - \epsilon\Gamma\sqrt{2I}\sin\gamma$. If we set $H_\epsilon = E =$ constant, we see that we can

express I in terms of γ and E if $\frac{\partial H_\epsilon}{\partial I} = I - 1$ is nonzero. If we denote the periodic orbit thus obtained by $O_\epsilon(E)$, we see that

Proposition 3.6. *The orbit $O_\epsilon(E)$ varies smoothly with ϵ. In particular, $O_\epsilon(E)$ and $O_0(E)$ are $\mathcal{O}(\epsilon)$ apart.*

For fixed E, we denote the three-dimensional level sets of the Hamiltonian by

$$S_0^E = \{(x, y, I, \gamma) \,|\, H(x, y, I) = E\}$$

$$S_\epsilon^E = \{(x, y, I, \gamma) \,|\, H_\epsilon(x, y, I, \gamma) = E\}\,.$$

We will call them energy manifolds.

In Section 3.3 we imposed certain restrictions on the radii I_1 and I_2 of the annulus $\mathcal{A}$ in order that the homoclinic manifold $W(\mathcal{A})$ of $\mathcal{A}$ exists. Here we assume in addition that either $I_1 < I_2 < 1$ or $1 < I_1 < I_2$ holds so that $\mathcal{A}$ does not contain the resonance. Under these restrictions we can prove the following facts about the nondissipative case of the perturbed problem: The annulus $\mathcal{A}$, its stable and unstable manifolds, $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$, respectively, and their intersection $W_\epsilon^s(\mathcal{A}) \cap W_\epsilon^u(\mathcal{A})$ all intersect the manifolds S_ϵ^E transversely (see Figure 3.16). For every periodic orbit $O_\epsilon(E)$, the surfaces $S_\epsilon^E \cap W_\epsilon^s(\mathcal{A})$ and $S_\epsilon^E \cap W_\epsilon^u(\mathcal{A})$ are its stable and unstable manifolds. Transversality of all of the above-mentioned intersections with S_ϵ^E guarantees that if $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$ intersect transversely in a surface, the I coordinate of whose points can take on all the possible values between I_1 and I_2, then transversality of the intersection of $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$ implies transversality

of the intersection of the stable and unstable manifolds of all of the periodic orbits $O_\epsilon(E)$ in $\mathcal{A}$.

Therefore, since we restrict ourselves to an energy manifold S_ϵ^E, we are back in the familiar situation of a hyperbolic periodic orbit whose stable and unstable manifolds intersect transversely. As we have mentioned before, another method exists for obtaining the same result. This alternative method consists of using γ as the independent variable for the dynamics restricted to S_ϵ^E (see Holmes and Marsden [1982], and Ercolani, McLaughlin, and Forest [1989]). Even though this method gives the same Melnikov function as ours in the end, it relies crucially on the assumption that $\dot{\gamma}$ is nonzero along the unperturbed homoclinic surface $W(\mathcal{A}) \cap S_0^E$. Therefore, it cannot be applied to all situations in our problem where $\dot{\gamma} = 1 - I - x^2$ which may well have isolated zeros.

We now present a few propositions which will prove the assertions we made above. We first prove

Proposition 3.7. *For small enough values of ϵ, the manifolds S_ϵ^E intersect $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$ transversely.*

PROOF: Since transverse intersections survive small perturbations, transversality of the intersection of S_0^E and $W(\mathcal{A})$ implies the result of our proposition. Hence, we only need to prove that the manifolds S_0^E and $W(\mathcal{A})$ intersect transversely. This proof proceeds as follows. To begin with, we recall that the unperturbed homoclinic manifold $W(\mathcal{A})$ is three-dimensional and is described by the equation (2.10), which

is $H(x,y,I) - H(0,0,I) = 0$. The normal to this manifold is

$$\mathbf{n}_W = \left(\frac{\partial H(x,y,I)}{\partial x}, \frac{\partial H(x,y,I)}{\partial y}, \frac{\partial H(x,y,I)}{\partial I} - \frac{\partial H(0,0,I)}{\partial I}, 0 \right) .$$

The normal to the three-dimensional manifold S_0^E is

$$\mathbf{n}_E = \left(\frac{\partial H(x,y,I)}{\partial x}, \frac{\partial H(x,y,I)}{\partial y}, \frac{\partial H(x,y,I)}{\partial I}, 0 \right) ,$$

so at each point of intersection, the normals differ by the vector

$$\left(0, 0, -\frac{\partial H(0,0,I)}{\partial I}, 0 \right) = (0, 0, 1 - I, 0) .$$

Since we have excluded the points with $I = 1$ from $\mathcal{A}$ and hence also from $W(\mathcal{A})$, this vector is nonzero on $\mathcal{A}$. Since the normals of the two manifolds S_ϵ^E and $W(\mathcal{A})$ do not coincide at the intersection point, neither do their tangent spaces and hence the manifolds intersect transversely in a two-dimensional surface. This concludes the proof of the proposition.

This proposition also shows that the manifolds S_ϵ^E foliate $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$ into two-dimensional energy "slices," $W_\epsilon^s(O_\epsilon(E)) = S_\epsilon^E \cap W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(O_\epsilon(E)) = S_\epsilon^E \cap W_\epsilon^u(\mathcal{A})$. We will now show that these energy "slices" are nothing but the stable and unstable manifolds of the orbits $O_\epsilon(E)$.

Proposition 3.8. *For every periodic orbit $O_\epsilon(E)$, the surface $W_\epsilon^s(O_\epsilon(E))$ is its stable manifold and the surface $W_\epsilon^u(O_\epsilon(E))$ its unstable manifold.*

PROOF: The proof of this proposition will consist of two parts. We will first prove that if we restrict ourselves to $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$, the manifolds $W_\epsilon^s(O_\epsilon(E))$ and

$W^u_\epsilon(O_\epsilon(E))$ intersect $\mathcal{A}$ transversely along $O_\epsilon(E)$ (see Figure 3.16). We will then show that orbits on $W^s_\epsilon(O_\epsilon(E))$ and $W^u_\epsilon(O_\epsilon(E))$ approach $O_\epsilon(E)$ in forward and backward time, respectively.

In the first part, we must start by proving that the energy surfaces S^E_ϵ intersect $\mathcal{A}$ transversely along the periodic orbits $O_\epsilon(E)$. Since transverse intersections persist under perturbations we only need this proof for $\epsilon = 0$. The latter consists of checking that the energy surfaces S^E_0 intersect $\mathcal{A}$ transversely. We start this by noting that the tangent space of Π_c at any point in $\mathcal{A}$ is spanned by the vectors

$$\hat{I} = (0, 0, 1, 0)$$

$$\hat{\gamma} = (0, 0, 0, 1)\,.$$

The normal to S^E_0 at any point in Π_c away from $I = 1$ is

$$\mathbf{n}_E = (0, 0, \frac{\partial H}{\partial I}(0, 0, I), 0) = (0, 0, I - 1, 0)$$

which is a multiple of $\hat{I}$. Therefore, the sum of the tangent spaces of Π_c and S^E_0 at a given point in Π_c equals the sum of the tangent space of S^E_0 and the normal to S^E_0, which is $\mathbf{R}^4$. Transversality of the intersection between S^E_ϵ and $\mathcal{A}$ is then proven.

The statement that $W^s_\epsilon(O_\epsilon(E))$ and $W^u_\epsilon(O_\epsilon(E))$ intersect $\mathcal{A}$ along the periodic orbits $O_\epsilon(E)$ transversely when we restrict ourselves to $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$, respectively, now follows quite easily from the fact that S^E_ϵ is transverse to both $\mathcal{A}$ and the manifolds $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$.

We now perform the second part of the proof. For brevity, we do it only for the stable manifolds; the argument for the unstable manifolds is the same. We begin by noting that every orbit starting in $W^s_\epsilon(O_\epsilon(E))$ will be $\mathcal{O}(\epsilon)$ close to an orbit in $W^s_0(O_0(E))$ for finite times because of the Gronwall estimates. This means that it will eventually come to $W^s_{\epsilon,loc}(\mathcal{A})$. Once in there, it cannot leave $W^s_{\epsilon,loc}(\mathcal{A})$, since it must stay on $W^s_{\epsilon,loc}(\mathcal{A}) \cap W^s_\epsilon(O_\epsilon(E))$ which is $\mathcal{O}(\epsilon)$ close to $W^s_{0,loc}(\mathcal{A}) \cap W^s_0(O_0(E))$. The I coordinate on the latter is a constant given by the equation $\frac{1}{2}I^2 - I = E$, and hence the I coordinate on $W^s_{\epsilon,loc}(\mathcal{A}) \cap W^s_\epsilon(O_\epsilon(E))$ can not vary by more than $\mathcal{O}(\epsilon)$, which is certainly not enough for any orbit in it to escape $W^s_{\epsilon,loc}(\mathcal{A})$. Therefore, the orbit must approach an invariant set in $\mathcal{A}$ because of Proposition 3.4 in Section 3.3. The only possible such set is $W^s_\epsilon(O_\epsilon(E)) \cap \mathcal{A} = O_\epsilon(E)$, which concludes our proof.

We remark that the situation in Proposition 3.8 is analogous to that of the usual planar Hamiltonian system with a saddle-type fixed point. By continuity, the value of the Hamiltonian at the fixed point and on its stable and unstable manifolds must be the same. In other words, in that case the fixed point, together with its stable and unstable manifolds, forms one level set of the Hamiltonian, just as $O_\epsilon(E)$ together with $W^s_\epsilon(O_\epsilon(E))$ and $W^u_\epsilon(O_\epsilon(E))$ does when restricted to the union of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ in our problem.

We now turn our attention to the intersections of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ and their implications for our situation. It is clear that if $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ intersect transversely along a two-dimensional surface $\mathcal{S}_\epsilon$, and if the energy manifold S^E_ϵ

intersects $\mathcal{S}_\epsilon$ transversely, then $W^s_\epsilon(O_\epsilon(E))$ and $W^u_\epsilon(O_\epsilon(E))$ intersect transversely inside S^E_ϵ along $S^E_\epsilon \cap \mathcal{S}_\epsilon$.

To check for transverse intersections of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$, we apply the Melnikov theory developed in the previous section. The Melnikov function in the nondissipative case reduces to

$$M(I, \gamma_0; k, 0, \Gamma) = \Gamma F(I, k) \cos \gamma_0 \, .$$

It has zeros at $\gamma_0 = \frac{(2n+1)\pi}{2}$ for all integer n. If $F(I,k)$ is nonzero for $I_1 \leq I \leq I_2$ at a given value of k, the partial derivative $\frac{\partial M}{\partial \gamma_0}(I, \gamma_0; k, 0, \Gamma) = (-1)^n \Gamma F(I,k)$ is nonzero and hence we can describe the intersection surface by $\gamma_0 = \frac{(2n+1)\pi}{2} + \mathcal{O}(\epsilon)$, where the $\mathcal{O}(\epsilon)$ terms depend on the other parameters in the Melnikov function; in particular I, and possibly t_0. Therefore, $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ intersect transversely along a surface $\mathcal{S}_\epsilon$ which is near the surface

$$S_0 = \left\{ (x(-t_0, I), y(-t_0, I), I, \gamma(-t_0, I, \gamma_0)) \,|\, t_0 \in R, I_1 < I < I_2, \gamma_0 = \tfrac{(2n+1)\pi}{2} \right\}$$

where $x(-t_0, I)$, $y(-t_0, I)$, and $\gamma(-t_0, I, \gamma_0)$ are the homoclinic coordinates described in the previous section.

We are now in a position to establish

Proposition 3.9. *The surface $\mathcal{S}_\epsilon$ intersects the manifolds S^E_ϵ transversely.*

PROOF: As in the above propositions, we only need the proof for $\epsilon = 0$. We first note that $\mathcal{S}_0$ is described by the two equations

$$H(x, y, I) - H(0, 0, I) = 0 \qquad (3.55a)$$

$$\gamma_0 = \frac{(2n+1)\pi}{2}\,. \tag{3.55b}$$

From the above parametrization of $\mathcal{S}_0$ we find the two independent tangents of $\mathcal{S}_0$ to be

$$\begin{aligned}\mathbf{t}_{t_0} &= \left(\frac{\partial x(-t_0,I)}{\partial t_0}, \frac{\partial y(-t_0,I)}{\partial t_0}, 0, \frac{\partial \gamma(-t_0,I,\gamma_0)}{\partial t_0}\right) \\ &= \left(-\frac{\partial H}{\partial y}(x,y,I), \frac{\partial H}{\partial x}(x,y,I), 0, \frac{\partial H}{\partial I}(x,y,I)\right)\end{aligned}$$

and

$$\mathbf{t}_I = \left(\frac{\partial x(-t_0,I)}{\partial I}, \frac{\partial y(-t_0,I)}{\partial I}, 0, \frac{\partial \gamma(-t_0,I,\gamma_0)}{\partial I}\right),$$

where, for less cumbersome notation, we suppress the argument $(-t_0, I)$ in x, y and γ whenever they appear in $H(x,y,I)$. Now, obviously, $\mathbf{t}_{t_0} \cdot \mathbf{n}_E = 0$, but

$$\begin{aligned}\mathbf{t}_I \cdot \mathbf{n}_E &= \frac{\partial H}{\partial x}(x,y,I)\frac{\partial x}{\partial I}(-t_0,I) + \frac{\partial H}{\partial y}(x,y,I)\frac{\partial y}{\partial I}(-t_0,I) + \frac{\partial H}{\partial I} \\ &= \frac{dH}{dI}(x(-t_0,I), y(-t_0,I), I)\end{aligned}$$

From (3.55a) we get

$$\frac{dH}{dI}(x(-t_0,I), y(-t_0,I), I) = \frac{\partial H}{\partial I}(0,0,I)\,,$$

, therefore,

$$\mathbf{t}_I \cdot \mathbf{n}_E = \frac{\partial H}{\partial I}(0,0,I) = I - 1$$

which is nonzero except at the resonance at $I = 1$, which we have excluded from $\mathcal{A}$ by hypothesis. This means that S and S_0^E intersect transversely, and by a dimensional count we see that they do so along a curve. This concludes our proof.

The intersection of $S_\epsilon^E \cap \mathcal{S}_\epsilon$ is precisely $W_\epsilon^s(O_\epsilon(E)) \cap W_\epsilon^u(O_\epsilon(E))$. Since $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$ also intersect transversely, it should be clear that $W_\epsilon^s(O_\epsilon(E))$ and $W_\epsilon^u(O_\epsilon(E))$ intersect transversely inside S_ϵ^E along $S_\epsilon^E \cap \mathcal{S}_\epsilon$. Thus we have proved

Proposition 3.10. *If $F(I,k)$ is nonzero on an interval $I_1 < I < I_2$ not containing $I = 1$, then for all small enough ϵ, the stable and unstable manifolds of all the periodic orbits contained in the annulus $\mathcal{A}$ intersect transversely on the energy manifolds S_ϵ^E. Therefore, all the periodic orbits on that annulus possess transverse homoclinic orbits.*

From the above discussion we conclude that the sufficient criterion for transverse intersections on the energy manifold of the stable and unstable manifolds of periodic orbits in Π_c contained inside an annulus in Π_c were

(a) that the Melnikov function had a zero for each value of I in that annulus, and

(b) that its partial derivative on the angle γ_0 was nonzero for all the values of I involved, and hence the zeros formed a curve which could be parametrized by I and that $\frac{\partial H}{\partial I}(0,0,I)$ was nonzero for all those I values.

This is important to note since the same conditions apply to the general situation.

Let us now see what the implications of transverse intersections of $W_\epsilon^s(O_\epsilon(E))$ and $W_\epsilon^u(O_\epsilon(E))$ are for the possible existence of a nearby chaotic invariant set. Assume for this that $O_\epsilon(E)$ lies near the circle in Π_c with $I = \bar{I}$. If in the unperturbed case $\dot{\gamma}$ is bounded away from zero along the homoclinic manifolds at $I = \bar{I}$, we can set up a Poincaré map by fixing γ at a particular value, say $\bar{\gamma}$, and monitoring x, y, and I as γ returns to $\bar{\gamma} + 2\pi$. We can then use the Smale-Birkhoff homoclinic theorem (see Wiggins [1988]) to show that this map possesses a chaotic invariant

Cantor set with the Poincaré map acting on it like a shift on two symbols. In fact, because the perturbed problem still possesses the symmetry with respect to the map $(x, y, I, \gamma) \mapsto (-x, -y, I, \gamma)$, each orbit $O_\epsilon(E)$ whose stable and unstable manifolds intersect transversely in S_ϵ^E possesses two transverse homoclinic orbits. These orbits imply the existence of another chaotic invariant Cantor set. The Poincaré map again acts on it like a shift on two symbols, but the two symbols now have a clear meaning. Each one of them corresponds to a passage of the solution near one of the two symmetric homoclinic orbits. Therefore, the dynamics of the solutions with the initial conditions in that invariant set exhibit random jumping from one of two symmetric loops to another, similar to that in the perturbed pendulum equation.

We will not give the rigorous details of the above discussion, since they are well documented in Wiggins [1988], and a similar case will be presented later in Chapter 4, and in Appendix G. We remark that if $\dot{\gamma}$ is not bounded away from zero, we should still be able to set up a Poincaré map by carefully choosing $\gamma = \bar{\gamma}$ and possibly waiting for more than one period for the γ coordinate of the orbits to return to $\bar{\gamma}$. Once the Poincaré map is set up, however, everything else should be the same as when $\dot{\gamma}$ is of one sign.

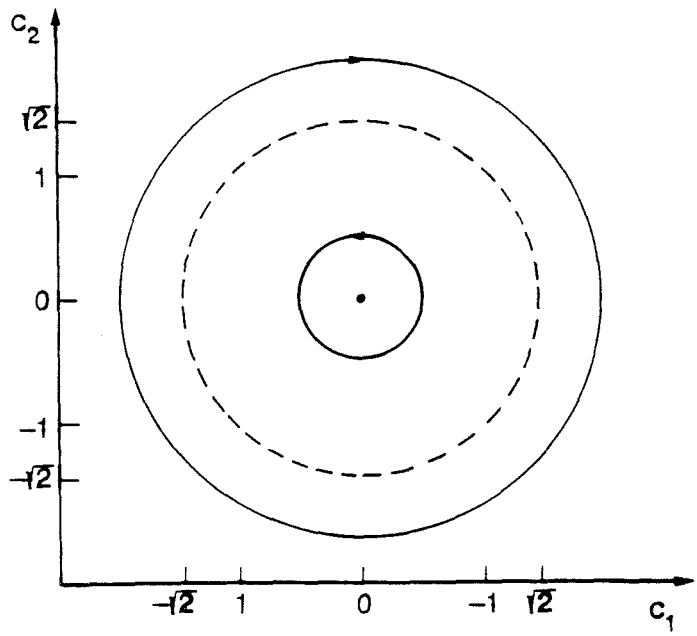

Figure 3.1. Dynamics in Π_c in the unperturbed case.

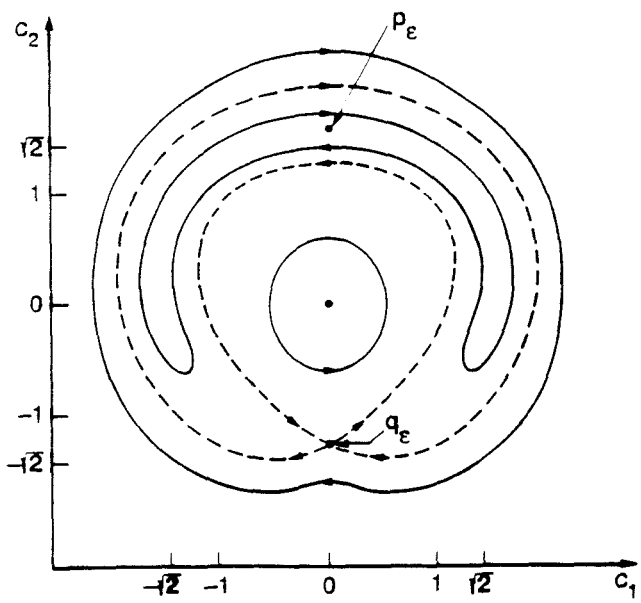

Figure 3.2. Dynamics in Π_c in the nondissipative case.

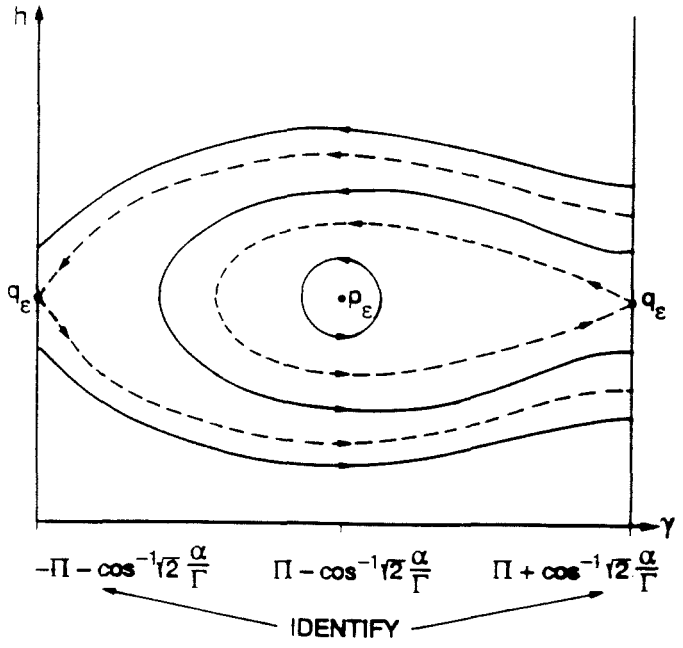

Figure 3.3. Phase portrait of the equations $(3.9)_0$.

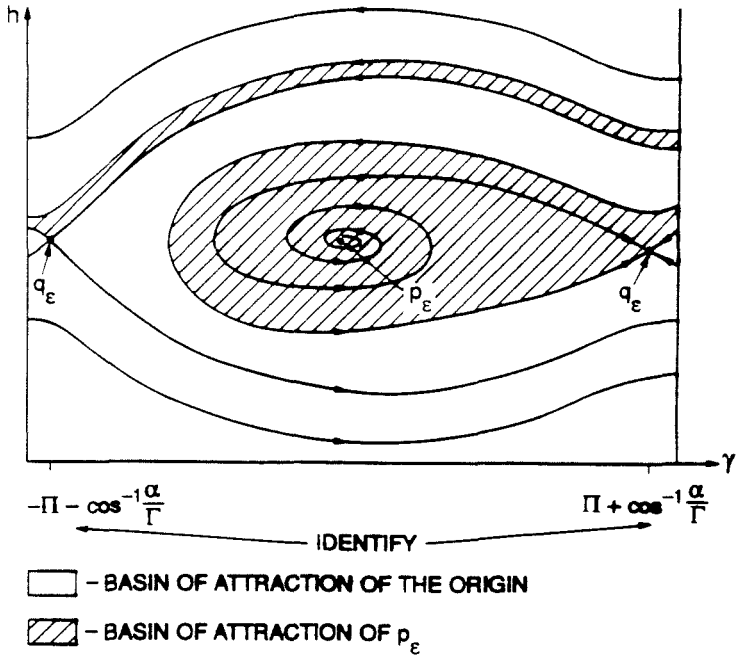

Figure 3.4. Phase portrait of the equations $(3.9)_\epsilon$.

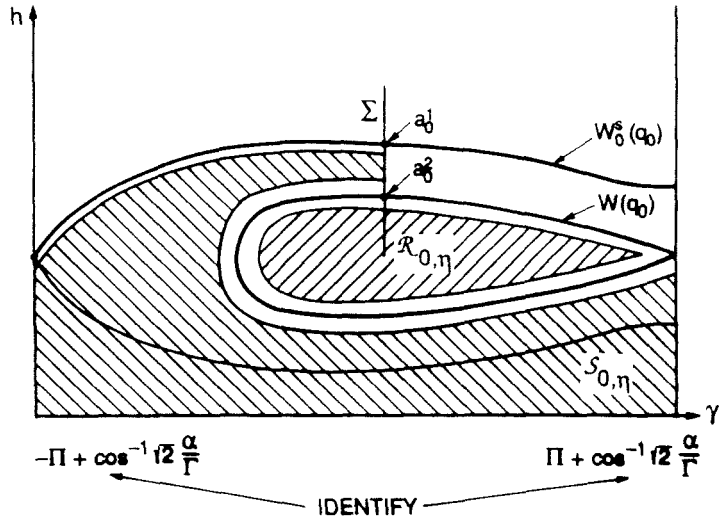

Figure 3.5. The regions $\mathcal{R}_{0,\eta}$ and $\mathcal{S}_{0,\eta}$.

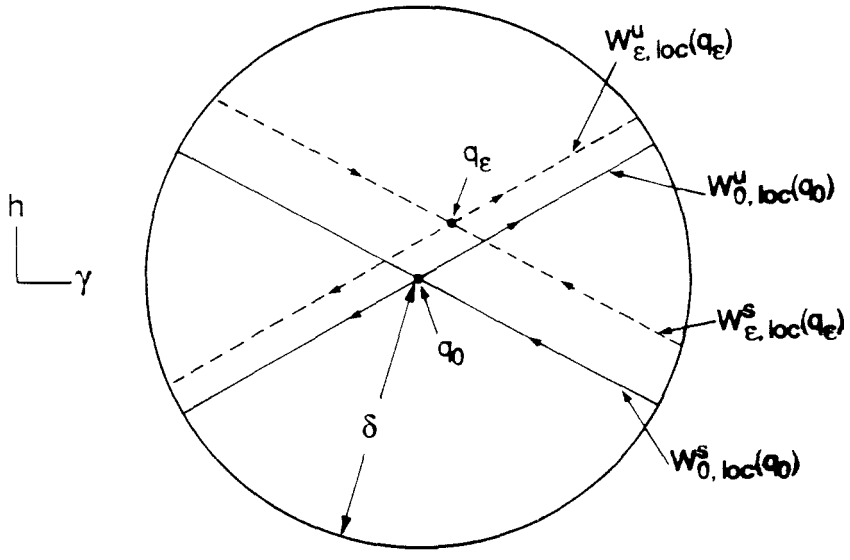

Figure 3.6. Local stable and unstable manifolds of q_0 and q_ϵ are close.

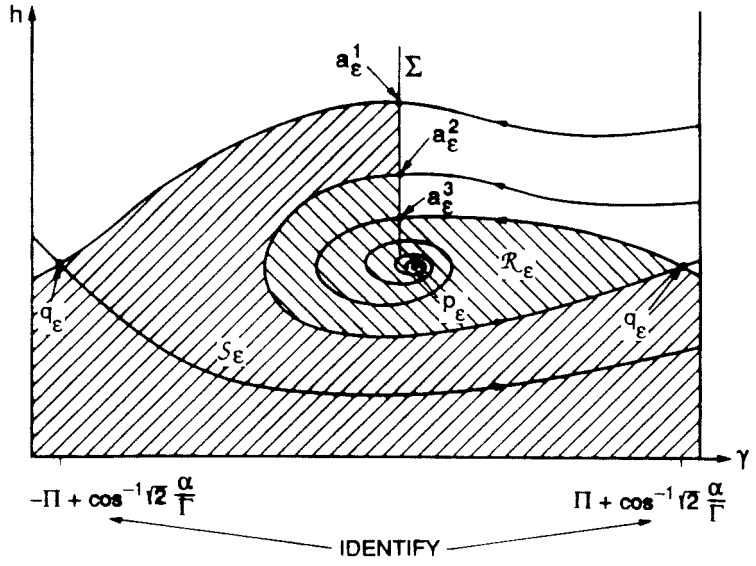

Figure 3.7. Trapping regions for the equations $(3.9)_\epsilon$.

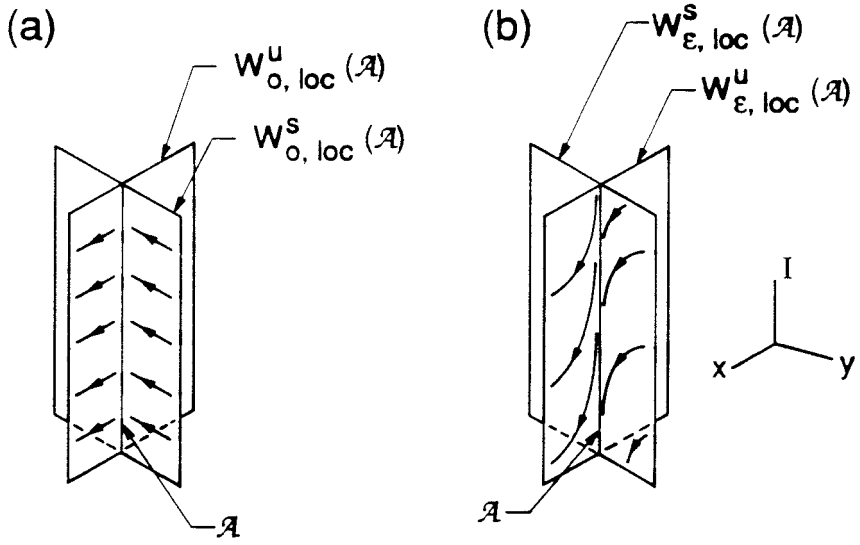

Figure 3.8. The local stable and unstable manifolds of $\mathcal{A}$ at a fixed value of γ:

(a) unperturbed, (b) perturbed.

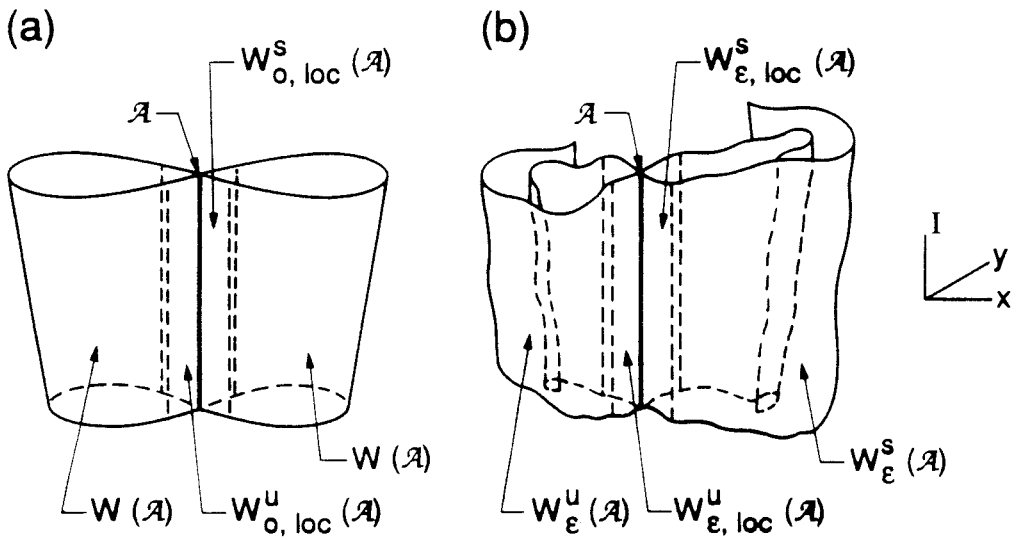

Figure 3.9. Global stable and unstable manifolds of $\mathcal{A}$ at a fixed value of γ:

(a) unperturbed, (b) perturbed.

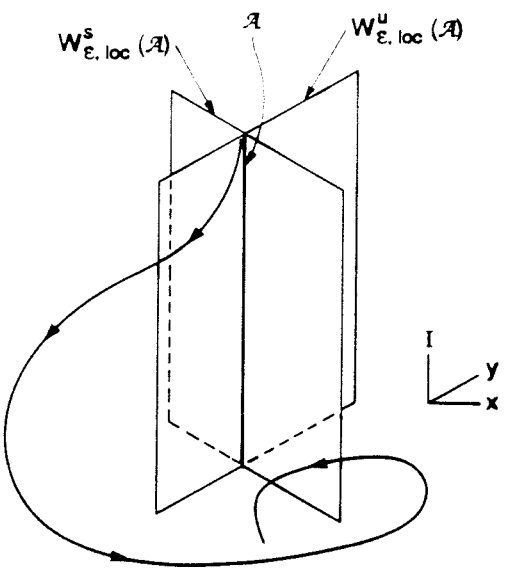

Figure 3.10. An intersection orbit which leaves $W^s_{\epsilon,loc}(\mathcal{A})$.

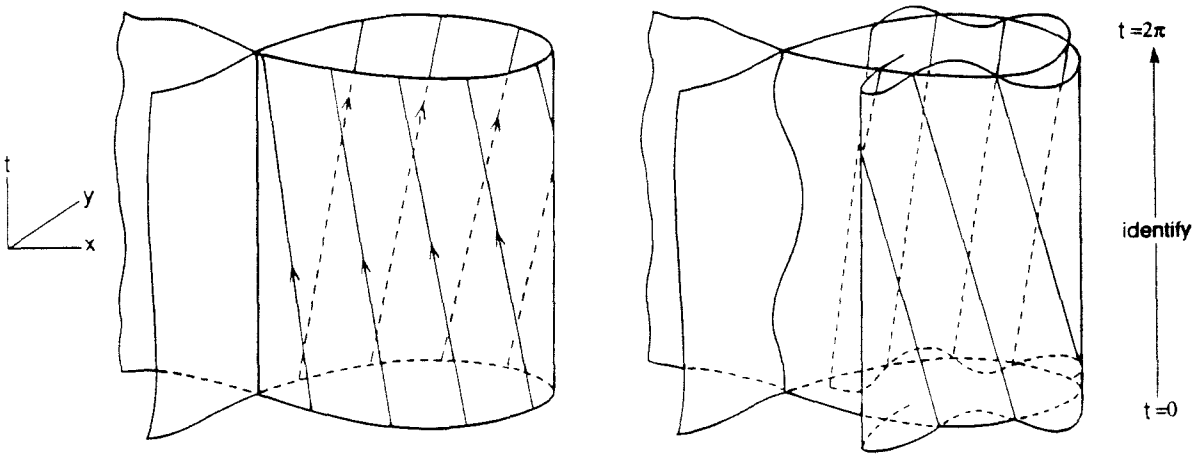

Figure 3.11. Periodically forced planar Hamiltonian system with a separatrix loop.

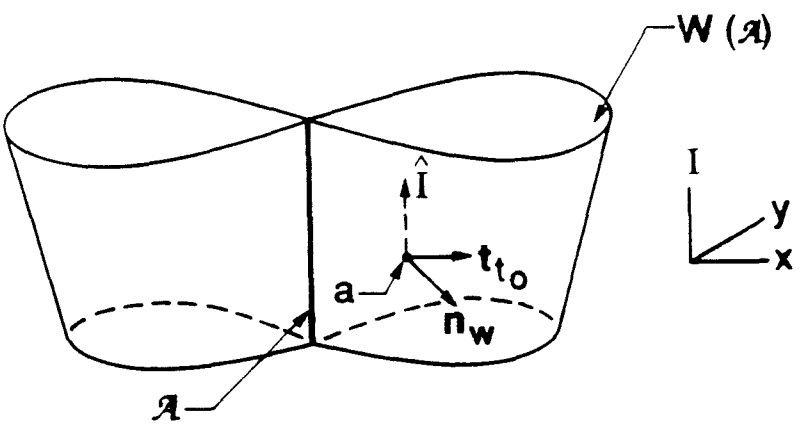

Figure 3.12. The homoclinic coordinate system.

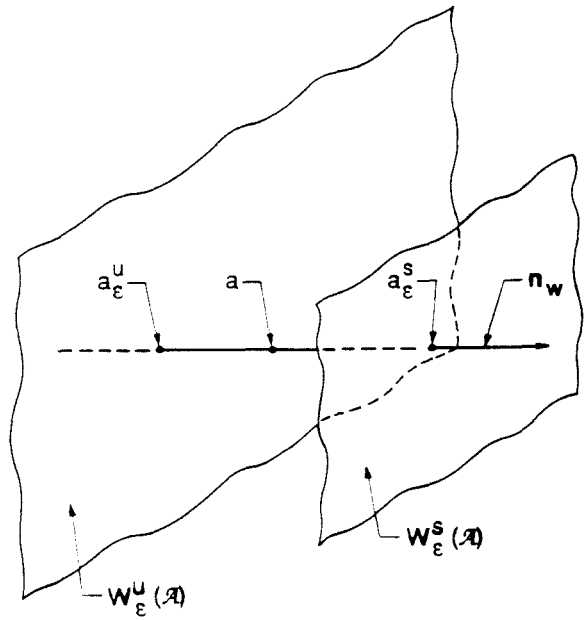

Figure 3.13. The normal $\mathbf{n}_W$ intersects $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$ transversely.

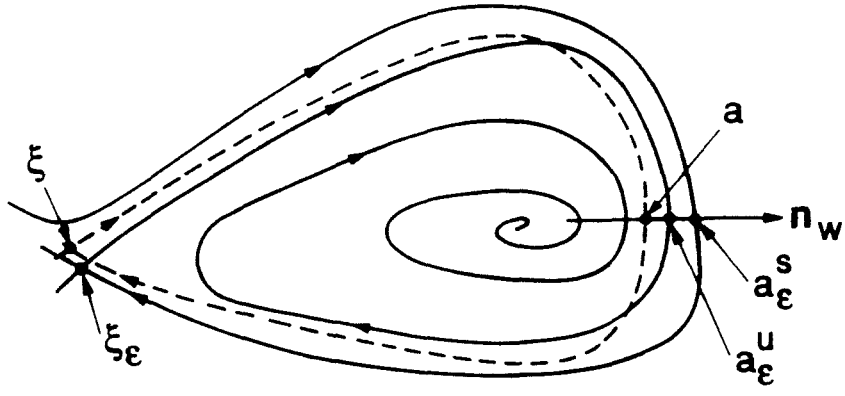

Figure 3.14. Points a^s_ϵ and a^u_ϵ, closest to the fixed point ξ_ϵ.

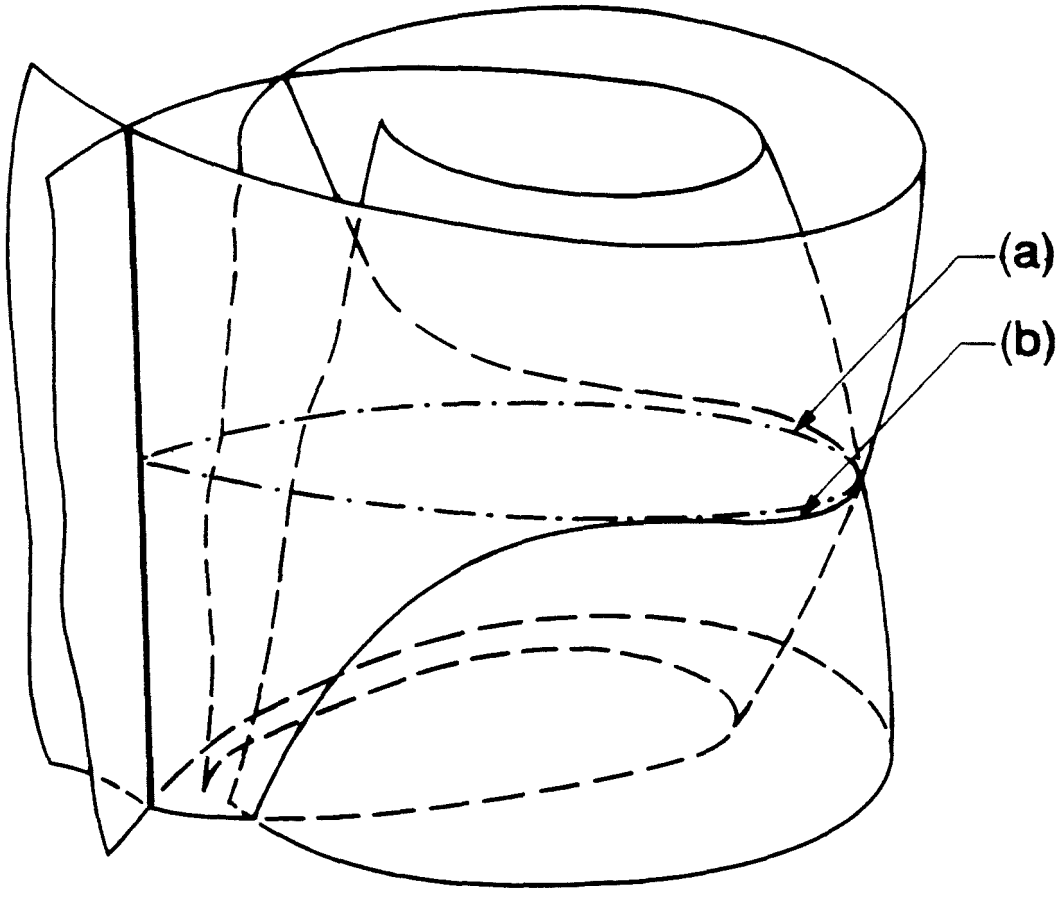

Figure 3.15. Intersections are close to surfaces of unperturbed orbits outside $W^{s}_{\epsilon,loc}(\mathcal{A})$ and $W^{u}_{\epsilon,loc}(\mathcal{A})$:

(a) an unperturbed orbit, (b) an intersection orbit.

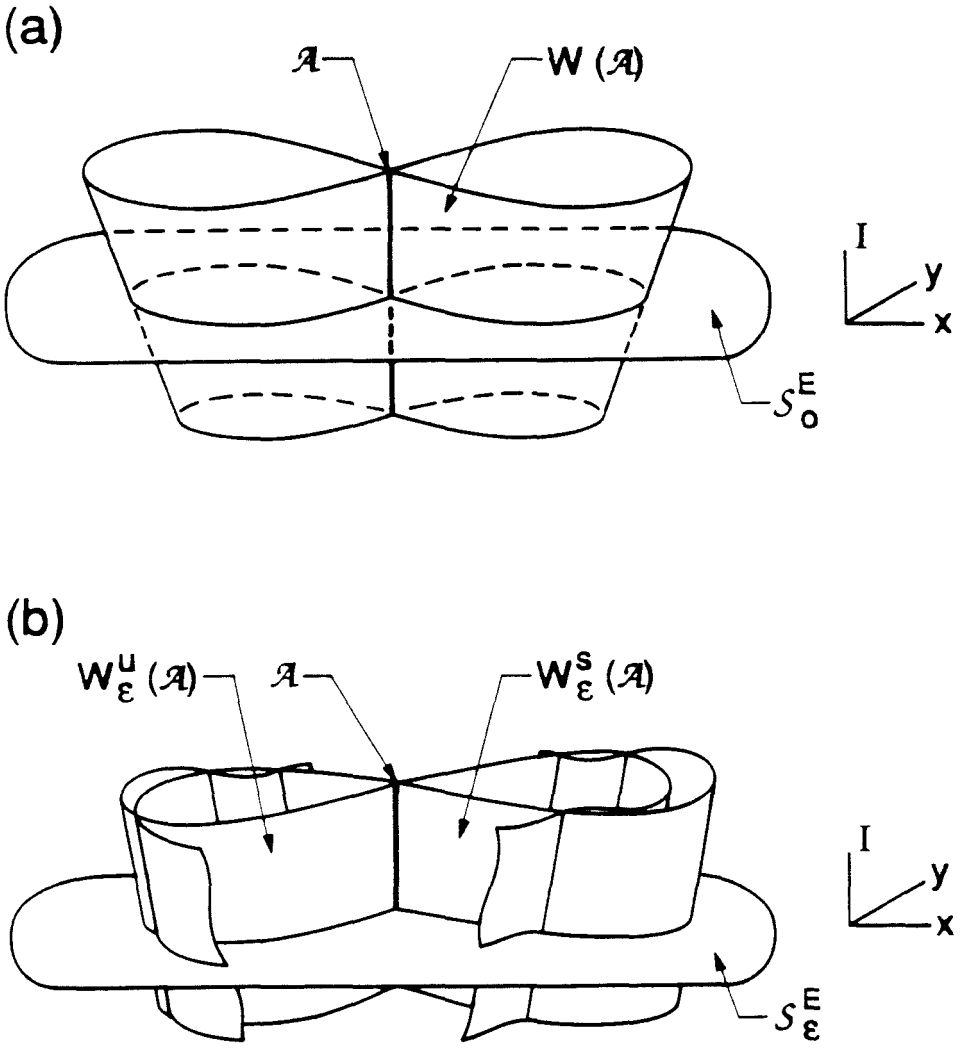

Figure 3.16. The energy manifolds intersect the stable and unstable manifolds of $\mathcal{A}$ transversely:

(a) unperturbed, (b) perturbed.

CHAPTER 4

A ŠILNIKOV ORBIT IN THE DISSIPATIVE CASE

4.1 Introduction. In this section, we discuss only the dissipative case. We investigate the existence of orbits homoclinic to the fixed point p_ϵ in Π_c and determine whether such orbit imply the existence of nearby chaotic invariant sets. As we have mentioned before, p_ϵ and q_ϵ are the only two invariant sets in Π_c capable of possessing homoclinic orbits in the dissipative case. We chose to study orbits homoclinic to p_ϵ since, as we will see, they mimic closely the mechanism for chaos observed in the numerical experiments performed by Ercolani, Forest, McLaughlin, and Overman, (see the references quoted in the introduction). Moreover, it appears that the results we have obtained will readily carry over to the infinite-dimensional phase space of the nonlinear Schroedinger equation.

All the difficulties we had to overcome in order to prove the existence of a Šilnikov homoclinic orbit can be attributed to one cause: We are studying orbits homoclinic to a resonance phenomenon born out of the perturbation. In Section 3.4 we have pointed out the shortcomings of the chief tool used so far for establishing the existence of chaotic Cantor sets in the phase space, the Melnikov theory. Its main use was as a persistence theory. It gives a sufficient criterion for determinig if certain orbits from the unperturbed phase spaces survive a given perturbation.

In our problem, orbits whose existence we are trying to prove do not exist in the unperturbed phase space; therefore, if we want to take advantage of the Melnikov theory at all, we have to find a new role for it. We have discussed a part of this new role in Section 3.4, where we have shown that the Melnikov function can be interpreted as the distance between the stable and the unstable manifolds of the annulus $\mathcal{A}$ in Π_c near a given point on the unperturbed homoclinic manifold.

The second part of the new role that the Melnikov theory will play in our work will be as a measure of the distance between a particular orbit, the unstable manifold, $W^u_\epsilon(p_\epsilon)$, of p_ϵ, and the three-dimensional stable manifold, $W^s_\epsilon(\mathcal{A})$, of the annulus $\mathcal{A}$. We develop this distance measurement in Section 4.3. This measurement is completely unlike any previous Melnikov technique in that we are measuring the distance between the stable manifold and unstable manifold of two different invariant sets having different dimensions. We will then show that the Melnikov function we need for this case is the same one we computed in Section 3.4 with $I = 1$ and γ_0 expressed in terms of k, α, and Γ via a condition which fixes the tail of $W^u_0(p_0)$ at p_0. The condition that the Melnikov function be equal to zero therefore gives us one equation for two unknowns, k and the ratio $\frac{\alpha}{\Gamma}$. In the space of the parameters k, α and Γ this equation defines a two dimensional surface, *i.e.*, a set with zero measure. This implies that the existence of a saddle connection at p_ϵ must be a rare event. We will discuss in Appendix G, however, how the existence of a saddle connection to p_ϵ at one point in the parameter space implies the existence

of a nearby chaotic invariant set in a neighborhood of that point. This means that the set of the points in the parameter space where a chaotic invariant set exists in the phase space will not have measure zero.

Also, showing persistence of the local unstable manifold, $W^u_{\epsilon,loc}(p_\epsilon)$, of p_ϵ and establishing its smoothness with respect to ϵ are nontrivial problems. They are nontrivial since the unperturbed point p_0 is not hyperbolic, *i.e.*, two of its eigenvalues are zero. Therefore, in Section 4.2 we will have to slightly modify the existing persistence theorem for the stable manifolds of a fixed point. Our theorem about the persistence of $W^u_\epsilon(p_\epsilon)$ and the closeness of $W^u_\epsilon(p_\epsilon)$ and $W^u_0(p_0)$ will first hold only locally in a neighborhood of p_0 and p_ϵ; *i.e.*, for $W^u_{\epsilon,loc}(p_\epsilon)$ and $W^u_{0,loc}(p_0)$. We will then extend $W^u_{\epsilon,loc}(p_\epsilon)$ to become the global unstable manifold, $W^u_\epsilon(p_\epsilon)$, of the fixed point p_ϵ and prove that two solutions starting on $W^u_{\epsilon,loc}(p_\epsilon)$ and $W^u_{0,loc}(p_0)$ a distance $\mathcal{O}(\epsilon)$ apart at $t = 0$ will stay $\mathcal{O}(\epsilon)$ close for all semi-infinite, positive t. This will extend the closeness of $W^u_\epsilon(p_\epsilon)$ and $W^u_0(p_0)$ to the beyond the small neighborhood of $\mathcal{A}$ in which the local manifolds are close. Closeness for the global manifolds will be important to know in Section 4.6. when we discuss the infinite time behavior of solutions on $W^u_\epsilon(p_\epsilon)$.

In sections 4.2 and 4.3 we will describe the behavior of orbits in $W^u_\epsilon(p_\epsilon)$ on the time intervals $(-\infty, t]$ for all finite t and state the conditions for the corresponding parts of $W^u_\epsilon(p_\epsilon)$ to be contained in $W^s_\epsilon(\mathcal{A})$, or equivalently, in $W^s_\epsilon(\mathcal{A}) \cap W^u_\epsilon(\mathcal{A})$. As we know, however, the solutions on $W^u_\epsilon(p_\epsilon)$ may still drift out of $W^s_{\epsilon,loc}(\mathcal{A})$ through

its bottom (defined in Section 3.3) as $t \to +\infty$. We will apply rescaling and special geometric singular perturbation techniques to follow the solution on $W^u_\epsilon(p_\epsilon)$ when it is close to $\mathcal{A}$ at very large values of t. In particular, we will show that the solution approaches a certain solution in $\mathcal{A}$ exponentially for large t, as long as it stays in $W^s_{\epsilon,loc}(\mathcal{A})$. We will then use the information about various basins of attraction in Π_c which we obtained in Section 3.2, in order to present the criteria for $W^u_\epsilon(p_\epsilon)$ to return to p_ϵ and establish the saddle connection. The discussion of the rescaling and the geometric singular perturbation techniques will be distributed among Sections 4.4, 4.5 and 4.6. In Section 4.4 we will discuss the rescaling, and in Section 4.5 we will present the necessary tool from the geometric singular perturbation theory, the stable (and unstable) fibers. We can compare the stable fibers to a shrinking leash connecting the solution on $W^u_\epsilon(p_\epsilon)$ to the corresponding solution in Π_c, making sure that they approach each other as time progresses. In Section 4.6, we then join all this information into an effective existence theorem for the orbit homoclinic to p_ϵ. It is in these last three sections that most of the hard work is hidden; in particular, Theorem 4.4 in Section 4.5 is a consequence of the hard analysis performed in four papers by N. Fenichel ([1971], [1974], [1977], [1979]). We stress again that the necessity to use these nontrivial tools stems from the fact that the resonant structures in our perturbed phase space do not persist under perturbation, but are rather created by it. Thus we are not looking for the persistence of an orbit homoclinic to p_0, but are instead constructing an orbit homoclinic to p_ϵ from a

heteroclinic orbit coming out of p_0 and orbits in Π_c.

The final result will be the criterion for the existence of two symmetric Šilnikov-type loops homoclinic to p_ϵ Section 4.6, and their implications for the existing chaotic dynamics. The dynamics near such a loop have first been theoretically described by Šilnikov [1965]. We will present a brief discussion of the relevant dynamics (with some of the details relegated to Appendix G) and point out the features they have in common with the dynamics of the Sine-Gordon equation.

4.2 Persistence and Smoothness of the Unstable Manifold $W^u_\epsilon(p_\epsilon)$ of the Fixed Point p_ϵ. We now turn to the first topic in our program, the unstable manifold $W^u_\epsilon(p_\epsilon)$ of p_ϵ. For positive damping $\epsilon\alpha$ the fixed point p_ϵ possesses a three-dimensional stable and a one-dimensional unstable manifold. This statement follows readily from the considerations in Section 3.2 and the fact that p_0 in the unperturbed case had a positive, a negative, and two zero eigenvalues. The unstable manifold, $W^u_\epsilon(p_\epsilon)$, of the point p_ϵ is our candidate for the homoclinic loop. However, even though its existence follows from standard theorems for every positive ϵ and α, its smooth dependence on ϵ down to and including $\epsilon = 0$ does not, and therefore we must address this question here.

As usual, we first examine the local situation and prove

Proposition 4.1. *For small enough ϵ and positive $\epsilon\alpha$, there exists a one-dimensional manifold $W^u_{\epsilon,loc}(p_\epsilon)$, the strong local unstable manifold of the fixed point p_ϵ. It depends analytically on ϵ and the parameters k, α, and Γ, and is uniformly $\mathcal{O}(\epsilon)$*

close to $W^u_{0,loc}(p_0)$.

PROOF: The proof is long and tedious, and since it is not crucial for the understanding of the rest of the material presented here, it will be relegated to Appendix D. Also, the subtleties of the analyticity of $W^u_\epsilon(p_\epsilon)$ will be discussed there.

We now want to extend this result outside a neighborhood of p_0 and p_ϵ. To this end, let $\hat{p}_0$ be a point on $W^u_{0,loc}(p_0)$ and let $\hat{p}_\epsilon$ be a point on $W^u_{\epsilon,loc}(p_\epsilon)$, a distance $\mathcal{O}(\epsilon)$ away from $\hat{p}_0$. Denote the solution of $(3.3)_0$ passing through $\hat{p}_0$ at $t = 0$ by $(\hat{x}_0(t,\hat{p}_0), \hat{y}_0(t,\hat{p}_0), 1, \hat{\gamma}_0(t,\hat{p}_0))$. (Note that it is just a time shift of either the solutions (2.23) or (2.24), depending on the value of k) and the solution of $(3.3)_\epsilon$ passing through $\hat{p}_\epsilon$ at $t = 0$ by $(\hat{x}_\epsilon(t,\hat{p}_\epsilon), \hat{y}_\epsilon(t,\hat{p}_\epsilon), \hat{I}_\epsilon(t,\hat{p}_\epsilon), \hat{\gamma}_\epsilon(t,\hat{p}_\epsilon))$. Using Gronwall's inequality (see Appendix C) we get

Proposition 4.2. *For any fixed positive time t, there exists a constant $N(T)$ such that for every t with $0 \le t \le T$, the inequality*

$$\|(\hat{x}_0(t,\hat{p}_0), \hat{y}_0(t,\hat{p}_0), 1, \hat{\gamma}_0(t,\hat{p}_0)) - (\hat{x}_\epsilon(t,\hat{p}_\epsilon), \hat{y}_\epsilon(t,\hat{p}_\epsilon), \hat{I}_\epsilon(t,\hat{p}_\epsilon), \hat{\gamma}_\epsilon(t,\hat{p}_\epsilon))\| \le \epsilon N(T)$$

holds.

Propositions 4.1 and 4.2 combined give us control over solutions on $W^u_\epsilon(p_\epsilon)$ on time intervals $(-\infty, t]$ for any finite t. The price we pay for it is having to decrease ϵ as we increase t. We will need the results from this section in Section 4.6, when we use the point $(\hat{x}_\epsilon(T,\hat{p}_\epsilon), \hat{y}_\epsilon(T,\hat{p}_\epsilon), \hat{I}_\epsilon(T,\hat{p}_\epsilon), \hat{\gamma}_\epsilon(T,\hat{p}_\epsilon))$ as the initial condition for a solution on the part of $W^u_\epsilon(p_\epsilon)$ which is no longer close to a solution on $W^u_0(p_0)$

but is instead close to a solution in Π_c.

4.3. The Distance Between $W^u_\epsilon(p_\epsilon)$ and $W^s_\epsilon(\mathcal{A})$. In this section we develop a calculation, analogous to the one in Section 3.4, which will enable us to calculate when $W^u_\epsilon(p_\epsilon)$ is contained in $W^s_\epsilon(\mathcal{A})$. The Melnikov function we will thus obtain will be the same as the one we computed in Section 3.4 with I and γ_0 fixed at some prescribed values; but the details of the geometry and the conditions which ensure that the zeros of the Melnikov function imply that $W^u_\epsilon(p_\epsilon)$ is contained in $W^s_\epsilon(\mathcal{A})$ are sufficiently different so that they deserve a special derivation.

We begin with a review of the relevant material from preceeding sections. We have seen in the previous section that for positive $\epsilon\alpha$ the unstable manifold $W^u_\epsilon(p_\epsilon)$ persists under perturbation as a one-dimensional manifold, a part of it being uniformly $\mathcal{O}(\epsilon)$ close to $W^u_0(p_0)$. Also, $W^u_\epsilon(p_\epsilon)$ depends smoothly on ϵ and all the other parameters in the problem.

We will assume that $I_1 < 1 < I_2$ so that the annulus $\mathcal{A}$ is contained in the resonance and set up the measurement of the distance between $W^u_\epsilon(p_\epsilon)$ and $W^s_\epsilon(\mathcal{A})$ using perturbation theory near $W^u_0(p_0)$ in a manner analogous to the one in Section 3.4. We will again need to follow the same steps as in Section 3.4. First, we will parametrize $W^u_0(p_0)$. Second, we will attach a coordinate system at each point of $W^u_0(p_0)$ in order to measure the distance between $W^u_\epsilon(p_\epsilon)$ and $W^s_\epsilon(\mathcal{A})$. Finally, we will determine the number of directions in which we need to measure the distance and develop the Melnikov function. Some of these steps will be the same as in

Section 3.4. We will develop the non-similar steps from the beginning, and contrast them with the corresponding steps in Section 3.4.

We first set up the parametrization of $W_0^u(p_0)$ by restricting the parametrization (3.16) of $W(\mathcal{A})$ to $W_0^u(p_0)$. In this parametrization, the choice of I and γ_0 singles out a particular orbit on $W(\mathcal{A})$. Since $W_0^u(p_0)$ is one of the orbits of the pinched homoclinic torus connecting the unperturbed circle of fixed points in Π_c at $I = 1$ to itself, the choice of I for our work with the resonance is clearly $I = 1$. The choice of γ_0 will follow from the definition of $W_0^u(p_0)$ as the orbit on which all the trajectories approach p_0 in backward time. The backward time limit of the γ component of a particular trajectory on $W_0^u(p_0)$ is given by formula (2.32a) for $0 < k < \frac{1}{2}$ and by formula (2.35a) for $\frac{1}{2} < k < \sqrt{2}$. But since all the other trajectories are obtained from this particular one by finite time translation, the backward time limit of their γ component is given by the same formulas (2.32a) and (2.35a) used for the particular trajectory.

Therefore, we can calculate $\gamma_0 = \bar{\gamma}_0(k, \alpha, \Gamma)$ which determines $W_0^u(p_0)$ from the γ component, $\pi - \cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}$, of p_0 and formulas (2.32a) and (2.35a), respectively. The result we obtain is

$$\bar{\gamma}_0(k, \alpha, \Gamma) = \frac{3\pi}{2} - \cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma} - \cot^{-1}\frac{k}{\sqrt{2-k^2}} - \frac{1}{\sqrt{7}}\tanh^{-1}\sqrt{\frac{7k^2}{2-k^2}} \qquad (4.1a)$$

for $k < \frac{1}{2}$ and

$$\bar{\gamma}_0(k, \alpha, \Gamma) = \pi - \cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma} - \tan^{-1}\frac{\sqrt{2-k^2}}{k} - \frac{1}{\sqrt{7}}\tanh^{-1}\sqrt{\frac{2-k^2}{7k^2}} \qquad (4.1b)$$

for $\frac{1}{2} < k < \sqrt{2}$. Our parametrization of $W_0^u(p_0)$ is therefore given by

$$W_0^u(p_0) = \{(x(-t_0, I), y(-t_0, I), I, \gamma(-t_0, I, \gamma_0)) \mid t_0 \in \mathbf{R}, I = 1, \gamma_0 = \bar{\gamma}_0(k, \alpha, \Gamma)\} .$$

At each point $a = (x(-t_0, 1), y(-t_0, 1), 1, \gamma(-t_0, 1, \bar{\gamma}_0(k, \alpha, \Gamma)))$ we then attach the moving coordinate frame spanned by the vectors $\mathbf{t}_{t_0}(x, y, I)$, $\mathbf{n}_W(x, y, I)$, $\hat{I}$ and $\hat{\gamma}$, exactly as in Section 3.4, where the components of these vectors are expressed by the formulas (3.17) through (3.20).

We next want to show that the number of independent directions along which we need to measure the distance between $W_\epsilon^u(p_\epsilon)$ and $W_\epsilon^s(\mathcal{A})$ is one and at the same time set up the measurement. Recall from Section 3.4 that we picked two points, $a_\epsilon^s \in W_\epsilon^s(\mathcal{A})$ and $a_\epsilon^u \in W_\epsilon^u(\mathcal{A})$, near $a \in W(\mathcal{A})$ in a unique way, and then measured the distance between a_ϵ^s and a_ϵ^u. If it happened that $a_\epsilon^s = a_\epsilon^u$ then $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$ intersected there.

We want to do the same thing here. However, we need to be careful since the way of obtaining a_ϵ^s and a_ϵ^u will now have to be different from that of Section 3.4. To point out the difference, let us first recall how we found a_ϵ^s and a_ϵ^u in Section 3.4. There we looked at the intersection points of the line through a, parallel to $\mathbf{n}_W$ at a, with $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$. The points a_ϵ^s and a_ϵ^u were the intersection points on $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$ whose trajectories took the least amount of time to reach $W_{\epsilon,loc}^s(\mathcal{A})$ and $W_{\epsilon,loc}^u(\mathcal{A})$, respectively. The key ingredient in the definition of a_ϵ^s and a_ϵ^u was the fact that $\mathbf{n}_W$ intersected $W(\mathcal{A})$ at a transversely in precisely the point a. This implied that the line through a parallel to $\mathbf{n}_W$ would also intersect $W_\epsilon^s(\mathcal{A})$

and $W^u_\epsilon(\mathcal{A})$ transversely, each one in at least one isolated point. This guaranteed the existence of a^s_ϵ and a^u_ϵ and, as we mentioned before, we could then choose them in a unique manner.

In this section we want a^u_ϵ to be on $W^u_\epsilon(p_\epsilon)$. However, at no point a on $W^u_0(p_0)$ can $W^u_0(p_0)$ and $\mathbf{n}_W$ intersect transversely; at best the sum of their tangent spaces is two-dimensional and for transversality it would have to be four. Because of this, $\mathbf{n}_W$ and $W^u_\epsilon(p_\epsilon)$ cannot be expected to intersect at all, and a^u_ϵ cannot be defined as one of the intersection points. In order to insure transversality, and thus the existence of a^u_ϵ, we must attach at a a subspace with precisely enough dimensions so that it will intersect $W^u_0(p_0)$ at a transversely in just the point a. Since $W^u_0(p_0)$ is one-dimensional, this subspace must be three-dimensional. The best possible choice for such a subspace is the subspace Σ_a spanned by the vectors $\mathbf{n}_W$, $\hat{I}$ and $\hat{\gamma}$ at a. This subspace is transverse to $W^u_0(p_0)$ at a where they intersect, since the tangent $\mathbf{t}_{t_0}$ to $W^u_0(p_0)$ at a and the above three vectors form a basis of $\mathbf{R}^4$, as was explained in Section 3.4. Therefore, Σ_a also intersects $W^u_\epsilon(p_\epsilon)$ transversely near a, and hence we can again define a^u_ϵ as the intersection point of Σ_a and $W^u_\epsilon(p_\epsilon)$ closest to $W^u_{\epsilon,loc}(\mathcal{A})$ in terms of the time the trajectory starting at a^u_ϵ takes to reach $W^u_{\epsilon,loc}(\mathcal{A})$.

From here on the situation is similar to that in Section 3.4. Since $\mathbf{n}_W$ is orthogonal to $W(\mathcal{A})$ at a, a line parallel to $\mathbf{n}_W$ through a^u_ϵ will be transverse to $W^s_\epsilon(\mathcal{A})$ and will hence intersect it in at least one point. We define a^s_ϵ as the

intersection point of the line parallel to $\mathbf{n}_W$ through a_ϵ^u and $W_\epsilon^s(\mathcal{A})$ which is closest to $W_{\epsilon,loc}^s(\mathcal{A})$, again in terms of the time the trajectory starting at a_ϵ^s takes to reach $W_{\epsilon,loc}^s(\mathcal{A})$. The signed distance $d(t_0, 1, \bar{\gamma}_0(k,\alpha,\Gamma); k, \alpha, \Gamma; \epsilon)$, and the derivation and form of the Melnikov function $M(1, \bar{\gamma}_0(k,\alpha,\Gamma); k, \alpha, \Gamma)$ are the same as in Section 3.4, except for the fact that the I and γ_0 variables are not free any more but instead assume the specific values $I = 1$ and $\gamma_0 = \bar{\gamma}_0(k,\alpha,\Gamma)$, which determine the unperturbed orbit $W_0^u(p_0)$. The main difference, again, lies in the geometry, since it is meaningless to talk about transversality of an intersection of $W_\epsilon^u(p_\epsilon)$ and $W_\epsilon^s(\mathcal{A})$. The reason for this is that $W_\epsilon^u(p_\epsilon)$ is an orbit and hence if it intersects $W_\epsilon^s(\mathcal{A})$ at one point, it must, because of the local invariance of $W_\epsilon^s(\mathcal{A})$, be contained in $W_\epsilon^s(\mathcal{A})$ until it leaves $W_\epsilon^s(\mathcal{A})$ through either the top or the bottom of $W_{\epsilon,loc}^s(\mathcal{A})$. We will thus say that in such a case $W_\epsilon^u(p_\epsilon)$ is contained in $W_\epsilon^s(\mathcal{A})$ *at least locally.* In particular, if $W_\epsilon^u(p_\epsilon)$ and $W_\epsilon^s(\mathcal{A})$ intersect, every solution on $W_\epsilon^u(p_\epsilon)$ must be contained in $W_\epsilon^s(\mathcal{A})$ for all large negative times. From the above discussion, we therefore see that $W_\epsilon^u(p_\epsilon)$ is either at least locally contained in $W_\epsilon^s(\mathcal{A})$ or else they have no points other than p_ϵ in common.

We remark that $W_0^u(p_0)$ and $W_\epsilon^u(p_\epsilon)$, the distance $d(t_0, 1, \bar{\gamma}_0(k,\alpha,\Gamma); k, \alpha, \Gamma; \epsilon)$ and the Melnikov function $M(1, \bar{\gamma}_0(k,\alpha,\Gamma); k, \alpha, \Gamma)$ all vary smoothly with the parameters k, α and Γ.

We now state the connection between zeros of the Melnikov function and intersections of $W_\epsilon^u(p_\epsilon)$ and $W_\epsilon^s(\mathcal{A})$ in

Proposition 4.3. *Suppose that for some* $k = \bar{k}$, $\alpha = \bar{\alpha}$ *and* $\Gamma = \bar{\Gamma}$ *the following statements are true:*

(1) $M(1, \bar{\gamma}_0(\bar{k}, \bar{\alpha}, \bar{\Gamma}); \bar{k}, \bar{\alpha}, \bar{\Gamma}) = 0.$

(2) *At least one of the first partial derivatives of* $M(1, \bar{\gamma}_0(k, \alpha, \Gamma); k, \alpha, \Gamma)$ *with respect to* k, α *or* Γ *is nonzero at* $k = \bar{k}$, $\alpha = \bar{\alpha}$ *and* $\Gamma = \bar{\Gamma}$.

Then for ϵ *sufficiently small,* $W^u_\epsilon(p_\epsilon)$ *is locally contained in* $W^s_\epsilon(\mathcal{A})$ *for some* k, α, and Γ *near* $k = \bar{k}$, $\alpha = \bar{\alpha}$ and $\Gamma = \bar{\Gamma}$.

PROOF: Fix t_0 and define

$$\begin{aligned}
&\tilde{d}(t_0, 1, \bar{\gamma}_0(k, \alpha, \Gamma); k, \alpha, \Gamma; \epsilon) \\
&= \frac{1}{\epsilon} \|\mathbf{n}_W(x(-t_0, I), y(-t_0, I), I)\| d(t_0, 1, \bar{\gamma}_0(k, \alpha, \Gamma); k, \alpha, \Gamma; \epsilon) \\
&= M(1, \bar{\gamma}_0(k, \alpha, \Gamma); k, \alpha, \Gamma) + \mathcal{O}(\epsilon)
\end{aligned}$$

as in Section 3.4. We have $\tilde{d}(t_0, 1, \bar{\gamma}_0(\bar{k}, \bar{\alpha}, \bar{\Gamma}); \bar{k}, \bar{\alpha}, \bar{\Gamma}; 0) = 0$ and one of the derivatives, say for definiteness $\frac{\partial \tilde{d}}{\partial \alpha}(t_0, 1, \bar{\gamma}_0(\bar{k}, \bar{\alpha}, \bar{\Gamma}); \bar{k}, \bar{\alpha}, \bar{\Gamma}; 0) = \frac{\partial M}{\partial \alpha}(1, \bar{\gamma}_0(\bar{k}, \bar{\alpha}, \bar{\Gamma}); \bar{k}, \bar{\alpha}, \bar{\Gamma})$ nonzero. Therefore, for small ϵ, we can express by the implicit function theorem $\alpha = \alpha(t_0, k, \Gamma; \epsilon)$ with $\alpha = \alpha(t_0, \bar{k}, \bar{\Gamma}; 0) = \bar{\alpha}$ and

$$\tilde{d}(t_0, 1, \gamma_0(k, \alpha(t_0, k, \Gamma; \epsilon), \Gamma); k, \alpha(t_0, k, \Gamma; \epsilon), \Gamma; \epsilon) = 0\,,$$

which implies

$$d(t_0, 1, \gamma_0(k, \alpha(t_0, k, \Gamma; \epsilon), \Gamma); k, \alpha(t_0, k, \Gamma; \epsilon), \Gamma; \epsilon) = 0\,.$$

This means that for all ϵ small enough $a^s_\epsilon = a^u_\epsilon$ near the point

$$a = (x(-t_0, 1), y(-t_0, 1), 1, \gamma(-t_0, 1, \bar{\gamma}_0(k, \alpha(t_0, k, \Gamma; \epsilon), \Gamma))) \tag{4.2}$$

on $W_0^u(p_0)$. (We will comment on the curious appearance of ϵ in (4.2) shortly.) But if this is true for one point a in $W_0^u(p_0)$, it must be true for all the other points in $W_0^u(p_0)$ with the same values of k, α, and Γ as well, since $W_\epsilon^u(p_\epsilon)$ is either locally contained in $W_\epsilon^s(\mathcal{A})$ or else they do not intersect at all. This shows that $\alpha(t_0, k, \Gamma; \epsilon)$ is independent of t_0, *i.e.*, $\alpha = \alpha(k, \Gamma; \epsilon)$ only, and concludes our proof.

We now want to make a comment concerning the expression (4.2). We note that the appearance of ϵ in (4.2) at first seems a bit mysterious, because a lies on $W_0^u(p_0)$, an unperturbed heteroclinic orbit, and ϵ is the perturbation parameter. However, we recall that the position of $W_0^u(p_0)$ depends on the parameters α and Γ, which also only enter in our equations through the perturbation. Moreover, in the above proof the parameter α must be a function of ϵ in order that $a_\epsilon^s = a_\epsilon^u$ for all small enough ϵ. This forces the choice of $W_0^u(p_0)$, and hence also that of p_0, to depend on ϵ in our proof.

To conclude this section, note that Proposition 4.3 only provides a criterion for $W_\epsilon^u(p_\epsilon)$ to be contained in $W_\epsilon^s(\mathcal{A})$ locally. This means that solutions on $W_\epsilon^u(p_\epsilon)$ leave $W_{\epsilon,loc}^s(\mathcal{A})$ as $t \to \infty$. In order to bring the forward asymptotic behavior of $W_\epsilon^u(p_\epsilon)$ under control, we must apply different methods which we will develop in the next few sections.

4.4. The Three Systems Needed to Establish the Existence of the Saddle Connection at p_ϵ. Having given in Proposition 4.3 the conditions on the values of α and Γ at which $W_\epsilon^u(p_\epsilon)$ is at least locally contained in $W_\epsilon^s(\mathcal{A})$, we

now have to investigate whether $W^u_\epsilon(p_\epsilon)$, in the case when these conditions are met, returns to p_ϵ as $t \to \infty$ or else leaves a neighborhood of $\mathcal{A}$. Intuitively, one would tend to believe that the sufficient condition for $W^u_\epsilon(p_\epsilon)$ to return to p_ϵ will be if the forward-time limit point of $W^u_0(p_0)$ is contained in the resonance. The reason for this belief is that we have shown on the one hand that the width of the resonance to be of order $\mathcal{O}(\sqrt{\epsilon})$ and on the other hand that $W^u_\epsilon(p_\epsilon)$ returns to any prescribed neighborhood of $\mathcal{A}$ at a distance of $\mathcal{O}(\epsilon)$ of where $W^u_0(p_0)$ does. Inside this neighborhood, the two orbits differ drastically. We may, therefore, expect the dynamics on $W^u_0(p_0)$ to bear no relationship with those on $W^u_\epsilon(p_\epsilon)$. However, all the orbits in the resonance tend to p_ϵ and it is reasonable to expect that so will all the orbits in its immediate neighborhood.

If one tries to prove these conjectures rigorously, though, one encounters some serious difficulties. The reason for these difficulties lies in the impreciseness of the statement that $W^u_\epsilon(p_\epsilon)$ returns to any prescribed neighborhood of $\mathcal{A}$ a distance $\mathcal{O}(\epsilon)$ away from where $W^u_0(p_0)$ does. The precise statements are as follows: Given any fixed positive δ, there will exist a large enough $t = T$ such that the x and y coordinates of any two trajectories on $W^u_0(p_0)$ and $W^u_\epsilon(p_\epsilon)$, which started a distance $\mathcal{O}(\epsilon)$ apart at $t = 0$, will satisfy $x^2 + y^2 < \delta$ at $t = T$, while the trajectories will still be a distance $\mathcal{O}(\epsilon)$ apart. This δ is completely independent of ϵ, and in fact it is incorrect to assume that we could make it shrink to zero together with ϵ. Therefore, as ϵ tends to zero, the width of the resonance becomes very small compared to δ

and there is no guarantee that $W^u_\epsilon(p_\epsilon)$ does not miss it.

After some thought we recognize two major obstacles to our attempted argument. The first obstacle is the fact that the width of the resonance shrinks to zero with ϵ. The second one is that there exists no *à-priori* reason for the orbits in $W^s_\epsilon(\mathcal{A})$ to behave in the same way as those in $\mathcal{A}$, no matter how close to $\mathcal{A}$ they may be. The problem with the width of the resonance band is, as we will see shortly, easily overcome by rescaling. The second obstacle is more difficult to overcome, however, and requires for its final resolution strong general results obtained by hard analysis, similar to those used in proving persistence of invariant manifolds.

In this section we resolve the first of the two obstacles discussed above. Using the appropriate rescaling we will magnify the resonance band and obtain two important sets of equations. Knowledge of these two sets of equations will allow us to understand the dynamics in and near the resonance. We will use one set of equations when we study the dynamics in the resonance which become slower and slower as $\epsilon \to 0$, and we will use the second set of equations when we study the exponential approach of trajectories in $W^s_\epsilon(\mathcal{A})$ to $\mathcal{A}$.

We make the same transformation $I = 1 + \sqrt{\epsilon}h$ as in Section 3.2 to arrive at the system

$$\dot{x} = -k^2 y + \tfrac{1}{4}y^3 - \tfrac{3}{4}x^2 y + \epsilon \left[\frac{\Gamma y}{\sqrt{2(1+\sqrt{\epsilon}h) - x^2 - y^2}} \sin\gamma - \alpha x \right] \qquad (4.3a)_\epsilon$$

$$\dot{y} = \left(k^2 - 2\right) x + \tfrac{7}{4}x^3 + \tfrac{3}{4}xy^2 - 2\sqrt{\epsilon}hx - \epsilon \left[\frac{\Gamma x}{\sqrt{2(1+\sqrt{\epsilon}h) - x^2 - y^2}} \sin\gamma + \alpha y \right] \qquad (4.3b)_\epsilon$$

$$\dot{h} = -\sqrt{\epsilon}\left[\Gamma\sqrt{2(1+\sqrt{\epsilon}h) - x^2 - y^2}\cos\gamma + 2\alpha(1+\sqrt{\epsilon}h)\right] \qquad (4.3c)_\epsilon$$

$$\dot{\gamma} = -x^2 - \sqrt{\epsilon}h + \epsilon\frac{\Gamma}{\sqrt{2(1+\sqrt{\epsilon}h) - x^2 - y^2}}\sin\gamma\,. \qquad (4.3d)_\epsilon$$

Putting $x = y = 0$, we can rescale time $\tau = \sqrt{\epsilon}t$ and again obtain the familiar equations

$$h' = -\Gamma\sqrt{2(1+\sqrt{\epsilon}h)}\cos\gamma - 2\alpha(1+\sqrt{\epsilon}h) \qquad (3.9a)_\epsilon$$

$$\gamma' = -h + \frac{\sqrt{\epsilon}\Gamma}{\sqrt{2(1+\sqrt{\epsilon}h)}}\sin\gamma \qquad (3.9b)_\epsilon$$

with $'$ representing the derivative on τ, which we have extensively studied in Section 3.2. Recall, in particular, that the width of the resonance band in the x, y, h and γ coordinates is $\mathcal{O}(1)$ and that its interior is approximately the interior of a separatrix $\mathcal{W}(q_0)$ when ϵ tends to zero. We note the two limiting sets of these equations which we get as $\epsilon \to 0$. They are

$$\dot{x} = -k^2 y + \tfrac{1}{4}y^3 - \tfrac{3}{4}x^2 y \qquad (4.3a)_0$$

$$\dot{y} = \left(k^2 - 2\right)x + \tfrac{7}{4}x^3 + \tfrac{3}{4}xy^2 \qquad (4.3b)_0$$

$$\dot{h} = 0 \qquad (4.3c)_0$$

$$\dot{\gamma} = -x^2 \qquad (4.3d)_0$$

and

$$h' = -\Gamma\sqrt{2}\cos\gamma - 2\alpha \qquad (3.9a)_0$$

$$\gamma' = -h \qquad (3.9b)_0\,.$$

Note that in the system $(4.3)_0$, the equations for the x, y and γ are the same as the equations for the x, y and γ in the system (2.6) at $I = 1$, and hence the solutions of these equations are also the same. In particular, the homoclinic orbit solutions from Section 2.4 with $I = 1$ hold in this case. The main difference between the systems $(4.3)_0$ and (2.6) is that the circle of fixed points has been magnified to cover the whole h-γ plane, which we will call $\hat{\Pi}_c$. (Note that for nonzero ϵ, $\hat{\Pi}_c$ and Π_c are identical.) We can indeed see directly that $\hat{\Pi}_c$ is a plane of fixed points for $(4.3)_0$. Also, the previously mentioned homoclinic solutions, together with the equation $h =$ constant, describe heteroclinic orbits connecting pairs of points in $\hat{\Pi}_c$.

We will again investigate the dynamics associated with annuli in $\hat{\Pi}_c$. Coordinates on such an annulus $\hat{\mathcal{A}}$ will be given by

$$\hat{\mathcal{A}} = \{(x, y, h, \gamma) \,|\, x = y = 0, h_1 < h < h_2, 0 \leq \gamma \leq 2\pi\} .$$

The system $(4.3)_\epsilon$ and its limits $(4.3)_0$ and $(3.9)_0$ will provide all the information we need to complete our analysis. Note that $(4.3)_\epsilon$ and $(3.3)_\epsilon$ are the same equations in different coordinates, and therefore, all the statements about one of them also hold for the other. The fact that we need two limits is typical of singular perturbation problems. As we have mentioned before, the so called inner limit, $(4.3)_0$, will provide the estimate for the rates of fast exponential approach of the orbits in $W^s_\epsilon(\hat{\mathcal{A}})$ to $\hat{\mathcal{A}}$. The outer equations $(3.9)_\epsilon$ and their limit $(3.9)_0$ will in turn provide the slow dynamics along $\hat{\mathcal{A}}$. The purpose of the next few sections will be to discuss how these two systems are coupled in $(4.3)_\epsilon$ and thus show that the dynamics in $W^s_\epsilon(\hat{\mathcal{A}})$

close to $\hat{\mathcal{A}}$ really are very similar to those in $\hat{\mathcal{A}}$.

4.5. Stable and Unstable Fibers. In the previous section we have rescaled the resonance; it is now $\mathcal{O}(1)$ wide and due to rescaling, $W^u_\epsilon(p_\epsilon)$ returns to any given neighborhood of $\hat{\Pi}_c$ a distance $\mathcal{O}(\sqrt{\epsilon})$ away from $W^u_0(p_0)$. Restricting ourselves to the annulus $\hat{\mathcal{A}}$ in $\hat{\Pi}_c$ given by

$$\hat{\mathcal{A}} = \{(x, y, h, \gamma) \,|\, x = y = 0, h_1 < h < h_2, 0 \leq \gamma \leq 2\pi\}$$

with h_1 and h_2 large enough, so that $\hat{\mathcal{A}}$ contains the resonance, we now want to show that once they are in a neighborhood of $\hat{\mathcal{A}}$, the solutions on $W^u_\epsilon(p_\epsilon)$ have the same forward time asymptotic behavior as solutions on an orbit in $\hat{\mathcal{A}}$ passing through an $\mathcal{O}(\sqrt{\epsilon})$ small neighborhood of the forward limit point of $W^u_0(p_0)$. In fact, any point on $W^u_\epsilon(p_\epsilon)$ will approach a corresponding point moving on that planar orbit exponentially in forward time. It is in this way that we are combining the dynamics of $(4.3)_0$ and $(3.9)_\epsilon$. The equations $(3.9)_\epsilon$ will give the orbit in $\hat{\mathcal{A}}$ and $(4.3)_0$ will provide the exponential rates of attraction towards $\hat{\mathcal{A}}$ transverse to $\hat{\mathcal{A}}$ along that orbit.

We begin this section with a simple example in $\mathbf{R}^2$

$$\dot{\xi} = -\epsilon\xi \qquad (4.4a)_\epsilon$$

$$\dot{\eta} = -\eta \qquad (4.4b)_\epsilon$$

which in a naive way models the behavior in $W^s_\epsilon(\hat{\mathcal{A}})$ in a small neighborhood of p_ϵ.

In particular, ξ corresponds to h and γ, and η corresponds to the distance from $\hat{A}$ in $(4.3)_\epsilon$.

For zero ϵ, the solution of $(4.4)_0$ is

$$\xi = \xi_0$$

$$\eta = \eta_0 e^{-t}.$$

We see that the ξ axis is a line of fixed points for $(4.4)_0$. The orbits of $(4.4)_0$ are straight lines of constant $\xi = \xi_0$ and all the points on them approach the base point $(\xi, \eta) = (\xi_0, 0)$ at an exponential rate e^{-t}. More specifically, if we only allow η_0 to have its absolute value smaller than, say, a positive number M, the distance between any point starting at $t = 0$ with one of the allowed values of η_0 and the base point will be less than Me^{-t}.

An analogous situation occurs in our system $(4.3)_0$. There, the plane $\hat{\Pi}_c$ is a plane of fixed points. The one-dimensional local stable manifolds of the fixed points foliate the local stable manifold of $\hat{\Pi}_c$. Solutions on these local stable manifolds approach the fixed points at their base at the exponential rate $e^{-k\sqrt{2-k^2}t}$ (this was discussed in Section 3.3). We thus see that the local stable manifolds of the fixed points in $\hat{\Pi}_c$ play the same role in the dynamics of the equations $(4.3)_0$ as the local stable manifolds of the fixed points on the ξ axis do in our example.

For nonzero ϵ, the solution of $(4.4)_\epsilon$ is

$$\xi = \xi_0 e^{-\epsilon t}$$

$$\eta = \eta_0 e^{-t},$$

and the orbits are now a one-parameter family of parabolas $\left\{\eta = C\xi^{\frac{1}{\epsilon}} \mid C \in R\right\}$. We see that the orbits differ drastically from the ones in the case of zero ϵ, but we can easily convince ourselves that vertical lines still bear a special meaning; they are mapped onto one another by the time flow, and all the points starting on the vertical segment $\mathcal{F}^s_\epsilon(\xi_0, 0) = \{(\xi, \eta) \mid \xi = \xi_0, -M < \eta < M\}$ will be mapped onto the segment $\{(\xi, \eta) \mid \xi = \xi_0 e^{-\epsilon t}, -Me^{-\epsilon t} < \eta < Me^{-\epsilon t}\}$ which is shrinking with the same exponential rate e^{-t} as it was for $\epsilon = 0$. (This notation will become clear shortly.) From the standpoint of orbits, any trajectory on any orbit being contained in $\mathcal{F}^s_\epsilon(\xi_0, 0)$ at $t = 0$ will approach the trajectory $(\xi_0 e^{-\epsilon t}, 0)$ at an exponential rate at least as fast as Me^{-t}. It is true that trajectories starting on different vertical lines approach each other exponentially as well; in fact, they all approach the origin, but the rate of approach is much slower, only $e^{-\epsilon t}$.

The vertical segments in our example problem $(4.4)_\epsilon$ are important enough to deserve a special name; we will call them the *stable fibers*. We note three important facts about them. First, stable fibers need not be orbits, yet they can be very well used to describe the dynamics of $(4.4)_\epsilon$.

Second, they form a *positively invariant family*, which means that the image of the fiber $\mathcal{F}^s_\epsilon(\xi_0, 0)$ under the flow of $(4.4)_\epsilon$ is, at time t, contained in the fiber $\mathcal{F}^s_\epsilon(\xi_0 e^{-\epsilon t}, 0)$.

Finally, even though orbits may change drastically under perturbation, fibers

persist. Moreover, we see that in our example, every fiber $\mathcal{F}^s_\epsilon(\xi_0, 0)$ can be represented as a graph over the ξ_0 and η variables by the equation $\xi = \xi(\xi_0, \eta) = \xi_0$.

We now want to explain how the features in our example follow from a general theory which can be applied to the phase space of $(4.3)\epsilon$. We want to show that $W^s_{\epsilon,loc}(\hat{\mathcal{A}})$ is foliated by the stable fibers of the points in a compact neighborhood of $\hat{\mathcal{A}}$. These fibers are smooth curves, vary smoothly with $\sqrt{\epsilon}$, and coincide with the local stable manifolds of the fixed points in $\hat{\mathcal{A}}$ at $\epsilon = 0$. Trajectories with the initial conditions on the same fiber will approach each other at the exponential rate $e^{-\kappa t}$ with $\kappa = k\sqrt{2 - k^2} + \mathcal{O}(\sqrt{\epsilon})$ as long as they stay in $W^s_{\epsilon,loc}(\hat{\mathcal{A}})$.

We first look at the unperturbed case, the equations $(4.3)_0$. There, $\hat{\mathcal{A}}$ is a manifold of fixed points with one-dimensional stable, one-dimensional unstable, and two-dimensional center directions. Note that the center directions are those associated with the two zero eigenvalues of each fixed point in $\hat{\mathcal{A}}$, *i.e.*, they are a consequence of the fact that $\hat{\mathcal{A}}$ is an annulus of fixed points. No zero eigenvalues are associated with the directions transverse to $\hat{\mathcal{A}}$; the eigendirections transverse to $\hat{\mathcal{A}}$ therefore correspond to either exponential contraction or exponential expansion. This property characterizes $\hat{\mathcal{A}}$ as being *normally hyperbolic* and is a key factor in the persistence of the stable fibers to be discussed shortly.

The second important property of the unperturbed case is that the local stable manifolds of the fixed points in $\hat{\mathcal{A}}$ can be expressed as graphs with y, h, and γ being functions of x and the coordinates h_c and γ_c of the fixed point in $\hat{\mathcal{A}}$ whose

stable manifold this graph describes. The form of the functions y, h, and γ can be calculated as follows: From $(4.3)_0$ we see that $h = h_c$. We can express y in terms of x from the Hamiltonian (2.5) with $I = 1$. Then we express θ in terms of x and $y(x)$ from (2.12), t in terms of $\theta(x)$ from either (2.23b) or (2.24b) with $I = 1$, from which one of the two depending on k, β in terms of $t(\theta(x))$ and γ_0 from either (2.23c) or (2.24c) with $I = 1$, again with the choice of the form depending on k, and finally γ in terms of β and θ from (2.19). The angle γ_0 can then be expressed from the formula $\gamma_c = \gamma(\infty)$ with $\gamma(\infty)$ given by either (2.32b) or (2.35b).

Obviously, the algebra is prohibitive, but for us it suffices to note that the representation of the stable manifolds of the fixed points in $\hat{\mathcal{A}}$ as graphs over x, h_c, and γ_c is possible.

We also note that, given any point p on $W^s_{\epsilon,loc}(\hat{\mathcal{A}})$ there exists precisely one fixed point in $\hat{\mathcal{A}}$ on whose stable manifold p lies. Therefore, we can express the h_c and γ_c coordinates of that fixed point in terms of, say, x, h, and γ coordinates of p.

We now turn our attention to the perturbed system. We want to show:

Theorem 4.4. *For every integer $r \geq 2$ and small positive number d we can find small enough δ and $\sqrt{\epsilon_0}$ such that for every $\sqrt{\epsilon}$ with $-\sqrt{\epsilon_0} < \sqrt{\epsilon} < \sqrt{\epsilon_0}$, the local stable manifold $W^s_{\epsilon,loc}(\hat{\mathcal{A}})$ of the annulus $\hat{\mathcal{A}}$ is foliated by fibers $\mathcal{F}^s_\epsilon(0, 0, h_c, \gamma_c)$ which are represented by the functions*

$$y = y(x, h_c, \gamma_c; k, \alpha, \Gamma; \sqrt{\epsilon}) \tag{4.5a}$$

$$h = h(x, h_c, \gamma_c; k, \alpha, \Gamma; \sqrt{\epsilon}) \tag{4.5b}$$

$$\gamma = \gamma(x, h_c, \gamma_c; k, \alpha, \Gamma; \sqrt{\epsilon}) , \tag{4.5c}$$

these functions being defined for x, h_c, γ_c *and* $\sqrt{\epsilon}$ *satisfying* $-\delta < x < \delta$, $h_1 - d < h_c < h_2 + d$, $0 \leq \gamma_c \leq 2\pi$ *and* $-\sqrt{\epsilon_0} < \sqrt{\epsilon} < \sqrt{\epsilon_0}$. *The fibers possess the following properties:*

(1) The functions y, h *and* γ *are* C^∞ *in* x *and* C^r *in* h_c, γ_c, k, α, Γ *and* $\sqrt{\epsilon}$ *including at* $\sqrt{\epsilon} = 0$.

(2) We have

$$y(0, h_c, \gamma_c; k, \alpha, \Gamma; \sqrt{\epsilon}) = 0 \tag{4.6a}$$

$$h(0, h_c, \gamma_c; k, \alpha, \Gamma; \sqrt{\epsilon}) = 0 \tag{4.6b}$$

$$\gamma(0, h_c, \gamma_c; k, \alpha, \Gamma; \sqrt{\epsilon}) = 0 , \tag{4.6c}$$

(3) At $\epsilon = 0$, *the fibers coincide with the stable manifolds of the fixed points in* $\hat{\mathcal{A}}$.

(4) The functions y, h *and* γ *are one-to-one on a neighborhood of* $\hat{\mathcal{A}}$ *in* $\hat{\Pi}_c$. *In particular, we can express* h_c *and* γ_c *in terms of* x, h, *and* γ *from (4.5). This means that a fiber is uniquely determined by any one of its points.*

(5) Let p *be a point in* $W^s_{\epsilon,loc}(\hat{\mathcal{A}})$, *whose* y, h, *and* γ *coordinates are (because of(4)) uniquely determined by its* x *coordinate and the coordinates* h_c *and* γ_c *of a point* p_c *in* $\hat{\Pi}_c$. *Let* $p^t = (x(t,p), y(t,p), h(t,p), \gamma(t,p))$ *denote the trajectory passing through* p *at* $t = 0$. *Likewise, let* $p_c^t = (0, 0, h_c(t, p_c), \gamma_c(t, p_c))$ *denote*

the trajectory passing through p_c at $t = 0$. Then, for positive t, we have

$$y(t,p) = y(x(t,p), h_c(t,p_c), \gamma_c(t,p_c); k, \alpha, \Gamma; \sqrt{\epsilon}) \tag{4.7a}$$

$$h(t,p) = h(x(t,p), h_c(t,p_c), \gamma_c(t,p_c); k, \alpha, \Gamma; \sqrt{\epsilon}) \tag{4.7b}$$

$$\gamma(t,p) = \gamma(x(t,p), h_c(t,p_c), \gamma_c(t,p_c); k, \alpha, \Gamma; \sqrt{\epsilon}) \tag{4.7c}$$

for as long as p^t stays in $W^s_{\epsilon,loc}(\hat{\mathcal{A}})$. This means that the flow of the equations $(4.3)_\epsilon$ maps the fiber $\mathcal{F}^s_\epsilon(p) = \mathcal{F}^s_\epsilon(p_c)$ inside the fiber $\mathcal{F}^s_\epsilon(p^t) = \mathcal{F}^s_\epsilon(p^t_c)$, i.e., fibers are a locally positively invariant family.

(6) Let $0 < \kappa < k\sqrt{2-k^2}$. Then there exists a constant K, such that for all $\sqrt{\epsilon}$ with $-\sqrt{\epsilon_0} < \sqrt{\epsilon} < \sqrt{\epsilon_0}$ we have $|x(t,p)| < Ke^{-\kappa t}$ and therefore also $|y(t,p)| < Ke^{-\kappa t}$ for all t such that p^t is in $W^s_{\epsilon,loc}(\hat{\mathcal{A}})$. In particular, this shows that all the trajectories passing through a point on the fiber $\mathcal{F}^s_\epsilon(p_c)$ will approach p^t_c, and hence each other, exponentially in forward time at the rate $e^{-\kappa t}$ as long as they stay in $W^s_{\epsilon,loc}(\hat{\mathcal{A}})$. Moreover, if p^t stays in $W^s_{\epsilon,loc}(\hat{\mathcal{A}})$ for all positive times, the fiber $\mathcal{F}^s_\epsilon(p)$ consists of precisely the points which approach p^t at least at the rate $e^{-\kappa t}$ in forward time.

(7) If $\{\mathcal{G}^s_\epsilon(p)\}$ is another family of fibers represented by the graphs

$$y = \tilde{y}(x, h_c, \gamma_c; k, \alpha, \Gamma; \sqrt{\epsilon})$$

$$h = \tilde{h}(x, h_c, \gamma_c; k, \alpha, \Gamma; \sqrt{\epsilon})$$

$$\gamma = \tilde{\gamma}(x, h_c, \gamma_c; k, \alpha, \Gamma; \sqrt{\epsilon})$$

and if the functions y, h, *and* γ *and* $\tilde{y}$, $\tilde{h}$ *and* $\tilde{\gamma}$ *have a tangency at the point* $(\bar{x}, \bar{h}_c, \bar{\gamma}_c; k, \alpha, \Gamma; \sqrt{\epsilon})$ *then that tangency is of order* r.

PROOF: The proof of this theorem follows easily from Theorem 9.1 in Fenichel, [1979]. We point out that the previously discussed normal hyperbolicity of $\hat{\mathcal{A}}$ in the unperturbed case plays a crucial role in the proof of Fenichel's theorem.

Before concluding this section, we want to make two comments. The first one is that the roles of the variables x and y can be interchanged in all of the above discussion. The second one is that we could define unstable fibers for $W^u_{\epsilon,loc}(\hat{\mathcal{A}})$ in precisely the same manner as we defined the stable ones. Fibers will be the main tool in our discussion of the conditions for $W^u_\epsilon(p_\epsilon)$ to return to p_ϵ. These conditions will be presented in the next section.

4.6 Orbits Homoclinic to p_ϵ. In this section we prove our main result: the existence of orbits homoclinic to p_ϵ. Most of the technicalities are behind us, and we have all the necessary tools at our disposal. The Melnikov calculation developed in Section 4.3 will enable us to check whether $W^u_\epsilon(p_\epsilon)$ is at least locally contained in $W^s_\epsilon(\mathcal{A})$ of some annulus $\mathcal{A}$ in Π_c. In case the answer to this question is affirmative, the estimates from Section 4.2 will bring two solutions starting at points $\hat{p}_0 \in W^u_{0,loc}(p_0)$ and $\hat{p}_\epsilon \in W^u_{\epsilon,loc}(p_\epsilon)$ back near $\mathcal{A}$ to points $p_0^T \in W^s_{0,loc}(\mathcal{A})$ and $p_\epsilon^T \in W^s_{\epsilon,loc}(\mathcal{A})$, and make sure that the x-y-h-γ coordinates of the points p_0^T and p_ϵ^T are at most $\mathcal{O}(\sqrt{\epsilon})$ apart. We will then limit ourselves to a smaller annulus

$\hat{\mathcal{A}}$ in the x-y-h-γ coordinates and show that the points p_0^∞ and p_ϵ^∞, the intersection points with $\hat{\mathcal{A}}$ of the fibers $\mathcal{F}_0^s(p_0^\infty)$ and $\mathcal{F}_\epsilon^s(p_\epsilon^\infty)$, respectively, are also $\mathcal{O}(\sqrt{\epsilon})$ close, which will enable us to locate p_ϵ^∞ to within an $\mathcal{O}(\sqrt{\epsilon})$ distance. By the results of the previous section, the solutions having p_ϵ^T and p_ϵ^∞, respectively, as initial conditions, will approach each other exponentially for as long as they stay in $W_{\epsilon,loc}^s(\hat{\mathcal{A}})$. We, therefore, only need to find the forward asymptotic behavior of the orbit passing through p_ϵ^∞ to determine what happens to $W_\epsilon^u(p_\epsilon)$, and this we can do by using the results studied in Section 3.2. In particular, if p_ϵ^∞ is attracted to p_ϵ, a homoclinic loop connecting p_ϵ to itself exists.

Having laid out the strategy, the remaining task is to provide the detailed propositions and proofs. We will first assume $W_\epsilon^u(p_\epsilon)$ to be locally contained in $W_\epsilon^s(\mathcal{A})$ and explore the consequences of this assumption in view of Sections 4.2 and 4.5. Recall that we proved that $W_{\epsilon,loc}^u(p_\epsilon)$ and $W_{0,loc}^u(p_0)$ are $\mathcal{O}(\epsilon)$ close in (x, y, I, γ) coordinates. We then chose two points, $\hat{p}_0$ and $\hat{p}_\epsilon$ on those two manifolds, $\mathcal{O}(\epsilon)$ away from each other and showed that the solutions starting at $\hat{p}_0$ and $\hat{p}_\epsilon$, respectively, remain uniformly $\mathcal{O}(\epsilon)$ away from each other for finite positive times. In particular, let T be the time it takes the solution starting at $\hat{p}_0$ to return to $x = \frac{\delta}{2}$, denote that point by p_0^T, and the corresponding point on $W_\epsilon^u(p_\epsilon)$, which is $\mathcal{O}(\epsilon)$ away, by p_ϵ^T. In the x, y ,h and γ coordinates, p_0^T and p_ϵ^T are $\mathcal{O}(\sqrt{\epsilon})$ apart. Let p_0^∞ and p_ϵ^∞ denote the unique intersection points of the fibers $\mathcal{F}_0^s(p_0^T)$ and $\mathcal{F}_\epsilon^s(p_\epsilon^T)$ with $\hat{\Pi}_c$. Then we have

Lemma 4.5. *The points p_0^∞ and p_ϵ^∞ are $\mathcal{O}(\sqrt{\epsilon})$ apart.*

PROOF: Let the x,h and γ coordinates of p_0^T and p_ϵ^T be

$$(x_0^T, h_0^T, \gamma_0^T) \qquad \text{and} \qquad (x_\epsilon^T, h_\epsilon^T, \gamma_\epsilon^T),$$

respectively, and let (h_{0c}, γ_{0c}) and $(h_{\epsilon c}, \gamma_{\epsilon c})$ be the h-γ coordinates of the points p_0^∞ and p_ϵ^∞ (which lie in Π_c). Then

$$|h_{\epsilon c} - h_{0c}| \leq M_x|x_\epsilon^T - x_0^T| + M_h|h_\epsilon^T - h_0^T| + M_\gamma|\gamma_\epsilon^T - \gamma_0^T| + M_\epsilon\sqrt{\epsilon}$$

where

$$M_x = \max\left\{ \frac{\partial h_c}{\partial x} \middle| -\delta < x < \delta, h_1 < h < h_1, 0 \leq \gamma \leq 2\pi, -\sqrt{\epsilon_0} < \sqrt{\epsilon} < \sqrt{\epsilon_0} \right\}$$

$$M_h = \max\left\{ \frac{\partial h_c}{\partial h} \middle| -\delta < x < \delta, h_1 < h < h_1, 0 \leq \gamma \leq 2\pi, -\sqrt{\epsilon_0} < \sqrt{\epsilon} < \sqrt{\epsilon_0} \right\}$$

$$M_\gamma = \max\left\{ \frac{\partial h_c}{\partial \gamma} \middle| -\delta < x < \delta, h_1 < h < h_1, 0 \leq \gamma \leq 2\pi, -\sqrt{\epsilon_0} < \sqrt{\epsilon} < \sqrt{\epsilon_0} \right\}$$

$$M_\epsilon = \max\left\{ \frac{\partial h_c}{\partial \sqrt{\epsilon}} \middle| -\delta < x < \delta, h_1 < h < h_1, 0 \leq \gamma \leq 2\pi, -\sqrt{\epsilon_0} < \sqrt{\epsilon} < \sqrt{\epsilon_0} \right\},$$

where $h_c = h_c(x, h, \gamma; k, \alpha, \Gamma; \sqrt{\epsilon})$. By the above discussion, the quantities $|x_\epsilon^T - x_0^T|$, $|h_\epsilon^T - h_0^T|$, and $|\gamma_\epsilon^T - \gamma_0^T|$ are all $\mathcal{O}(\sqrt{\epsilon})$, and therefore so is $|h_{\epsilon c} - h_{0c}|$. The same type of an inequality is valid for $|\gamma_{\epsilon c} - \gamma_{0c}|$ and the lemma follows.

Before stating our main result, we will find the position of the point p_0^∞ and investigate whether, for given k,α and Γ, it lies inside the region $\mathcal{R}_0$ enclosed by the separatrix $\mathcal{W}(q_0)$ or in the region $\mathcal{S}_0$, which were both defined in Section 3.2. The point p_ϵ^∞ will then lie in $\mathcal{R}_{0,\eta}$ or $\mathcal{S}_{0,\eta}$, respectively, (see Section 3.2.) for some

small η. The orbit with p_ϵ^∞ as the initial condition, will then, by Proposition 3.3, be either attracted to p_ϵ or else leave a neighborhood of $\hat{\mathcal{A}}$. The same will happen to the orbit starting at p_ϵ^T since p_ϵ^T and p_ϵ^∞ lie on the same fiber, and hence, orbits starting at p_ϵ^T and p_ϵ^∞ approach each other exponentially fast in forward time. The same type of argument shows that if p_0^∞ is contained in $\mathcal{S}_0$, the orbit starting at p_ϵ^T will drift away from $\hat{\mathcal{A}}$, most probably to the sink near the origin. The h-γ coordinates of the point p_0^∞ are

$$p_0^\infty = (0, \gamma^\infty(k, \alpha, \Gamma))$$

with

$$\gamma^\infty(k,\alpha,\Gamma) = 2\pi - \cos^{-1}\sqrt{2}\tfrac{\alpha}{\Gamma} - 2\cot^{-1}\frac{k}{\sqrt{2-k^2}} - \frac{2}{\sqrt{7}}\tanh^{-1}\sqrt{\frac{7k^2}{2-k^2}}$$

for $k < \frac{1}{2}$ and

$$\gamma^\infty(k,\alpha,\Gamma) = \pi - \cos^{-1}\sqrt{2}\tfrac{\alpha}{\Gamma} - 2\tan^{-1}\frac{\sqrt{2-k^2}}{k} - \frac{2}{\sqrt{7}}\tanh^{-1}\sqrt{\frac{2-k^2}{7k^2}}$$

for $\frac{1}{2} < k < \sqrt{2}$.

These formulas can be calculated from the γ coordinate of p_0, which is $\pi - \cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}$, and the formulas (2.33) for $0 < k < \frac{1}{2}$ and (2.36) for $\frac{1}{2} < k < \sqrt{2}$.

To determine whether p_0^∞ is contained in $\mathcal{R}_0$ or $\mathcal{S}_0$, we must first calculate the γ components of the two intersection points of $\mathcal{W}(q_0)$ with the γ axis. One of these two points is, of course, q_0. The position of the second one can be obtained by noting that the Hamiltonian (3.10) has the same value on the whole of $\mathcal{W}(q_0)$ as

it does at q_0. If we denote the γ coordinate of that second intersection point by $\tilde{\gamma}(k,\alpha,\Gamma)$, we thus see that $\tilde{\gamma}(k,\alpha,\Gamma)$ must satisfy the equation

$$\Gamma\sqrt{2}\sin\tilde{\gamma}(k,\alpha,\Gamma)+2\alpha\tilde{\gamma}(k,\alpha,\Gamma)+\Gamma\sqrt{2\left(1-2\left(\tfrac{\alpha}{\Gamma}\right)^2\right)}-2\alpha\left(\pi+\cos^{-1}\sqrt{2}\tfrac{\alpha}{\Gamma}\right)=0\,.$$

Since the second intersection point must lie between p_0 and q_0, its γ component must lie between the γ components of these two points. This requirement gives the inequality $\pi-\cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}<\tilde{\gamma}(k,\alpha,\Gamma)<\cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}+\pi$ (see Figure 3.3).

The above discussion has proved

Lemma 4.6. *For given k, α and Γ, the point p_0^∞ will be:*

(1) *contained in $\mathcal{R}_0$ if for some integer n the inequality*

$$\tilde{\gamma}(k,\alpha,\Gamma)<\gamma^\infty(k,\alpha,\Gamma)+2n\pi<\pi+\cos^{-1}\sqrt{2}\tfrac{\alpha}{\Gamma},$$

holds, and will be

(2) *contained in $\mathcal{S}_0$ if for some integer n the inequality*

$$\cos^{-1}\sqrt{2}\tfrac{\alpha}{\Gamma}-\pi<\gamma^\infty(k,\alpha,\Gamma)+2n\pi<\gamma^\infty(k,\alpha,\Gamma)$$

holds.

We are now ready to state our main result in

Theorem 4.7. *If for some $k=\bar{k}$, $\alpha=\bar{\alpha}$ and $\Gamma=\bar{\Gamma}$ we have*

$$M(1,\bar{\gamma}_0(\bar{k},\bar{\alpha},\bar{\Gamma});\bar{k},\bar{\alpha},\bar{\Gamma})=0$$

with at least one of the partial derivatives of M with respect to α or Γ being nonzero. Then

(1) *If for $k = \bar{k}$, $\alpha = \bar{\alpha}$ and $\Gamma = \bar{\Gamma}$ the inequality $\tilde{\gamma}(k, \alpha, \Gamma) < \gamma^{\infty}(\bar{k}, \bar{\alpha}, \bar{\Gamma}) + 2n\pi < \pi + \cos^{-1}\sqrt{2}\frac{\bar{\alpha}}{\bar{\Gamma}}$ holds, then for $k = \bar{k}$ and some α and Γ near $\alpha = \bar{\alpha}$ and $\Gamma = \bar{\Gamma}$ a homoclinic orbit connecting p_ϵ to itself exists for small ϵ.*

(2) *If for $k = \bar{k}$, $\alpha = \bar{\alpha}$ and $\Gamma = \bar{\Gamma}$ the inequality $\cos^{-1}\sqrt{2}\frac{\bar{\alpha}}{\bar{\Gamma}} - \pi < \gamma^{\infty}(\bar{k}, \bar{\alpha}, \bar{\Gamma}) + 2n\pi < \tilde{\gamma}(\bar{k}, \bar{\alpha}, \bar{\Gamma})$ holds, then for small ϵ and $k = \bar{k}$ no homoclinic orbit connecting p_ϵ to itself can exist for α near $\alpha = \bar{\alpha}$ and Γ near $\Gamma = \bar{\Gamma}$.*

PROOF: Assume, for definiteness, that $\frac{\partial M}{\partial \alpha}(1, \bar{\gamma}_0(\bar{k}, \bar{\alpha}, \bar{\Gamma}); \bar{k}, \bar{\alpha}, \bar{\Gamma})$ is nonzero. Then, as in the proof of Proposition 4.3, we find that there exists a function $\alpha = \alpha(k, \Gamma, \epsilon)$ with $\alpha(\bar{k}, \bar{\Gamma}, 0) = \bar{\alpha}$ such that $W^u_\epsilon(p_\epsilon)$ is at least locally contained in $W^s_\epsilon(\mathcal{A})$ for k, $\alpha = \alpha(k, \Gamma, \epsilon)$ Γ and ϵ near $k = \bar{k}$, $\alpha = \bar{\alpha}$, $\Gamma = \bar{\Gamma}$.

If the situation (1) arises, then by our previous discussion, and in particular, Lemma 4.6, a solution on $W^u_\epsilon(p_\epsilon)$ tends to p_ϵ in forward time and the existence of a homoclinic loop is proved. If situation (2) arises, then by the same arguments, $W^u_\epsilon(p_\epsilon)$ leaves $W^s_{\epsilon,loc}(\mathcal{A})$ at some large positive time and hence cannot approach p_ϵ. Most probably it ends up in the sink near the origin.

We observe that, if the conditions in (1) of Theorem 4.7 are met, the symmetry of our equations with respect to the coordinate transformation $(x, y) \mapsto (-x, -y)$ implies that not only one but two symmetric orbits homoclinic to p_ϵ exist. This will bear important consequences for the chaotic dynamics we are interested in.

We now present another criterion for the existence of the Šilnikov loops which is more convenient for calculations. It can be obtained as follows: Let n be an integer such that $\cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}-\pi < \gamma^{\infty}(k,\alpha,\Gamma)+2n\pi < \cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}+\pi$ and let $\bar{\gamma}^{\infty}(k,\alpha,\Gamma) = \gamma^{\infty}(k,\alpha,\Gamma)+2n\pi$. Assume also that the γ coordinate of q_0 is $\cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}+\pi$ (not $\cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}-\pi$). Then

Proposition 4.8. *If the value of the Hamiltonian (3.10) at $h = 0$ and $\gamma = \bar{\gamma}^{\infty}(k,\alpha,\Gamma)$ is less than its value at q_0, then p_0^{∞} will lie in $\mathcal{R}_0$. If the inequality is reverese, p_0^{∞} will lie in $\mathcal{S}_0$.*

PROOF: This statement follows immediately from the fact that the value of the Hamiltonian (3.10) for γ between $\cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}-\pi$ and $\cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}+\pi$ increases monotonically with the smallest distance of the orbits from p_0, where it attains its minimum.

In Appendix F we check if the conditions of Theorem 4.7 are satisfied for any value of k in the allowed interval between 0 and $\sqrt{2}$. We find numerically that the equation $M(I,\gamma_0;k,\alpha,\Gamma)=0$ cannot hold for any positive α and Γ with k from the above range. Therefore simple Šilnikov orbits cannot exist in our model. However, for the modified equations

$$\dot{x} = -k^2 y + \tfrac{1}{4}y^3 - \tfrac{3}{4}x^2 y + \epsilon\left[\Gamma\frac{y}{\sqrt{2I-x^2-y^2}}\sin\gamma - \alpha(1+\lambda)x\right] \qquad (4.8a)_\epsilon$$

$$\dot{y} = \left(k^2-2I\right)x + \tfrac{7}{4}x^3 + \tfrac{3}{4}xy^2 - \epsilon\left[\Gamma\frac{x}{\sqrt{2I-x^2-y^2}}\sin\gamma + \alpha(1+\lambda)y\right] \qquad (4.8b)_\epsilon$$

$$\dot{I} = -\epsilon\left[\Gamma\sqrt{2I-x^2-y^2}\cos\gamma + 2\alpha I\right] \qquad (4.8c)_\epsilon$$

$$\dot{\gamma} = 1 - I - x^2 + \epsilon\Gamma \frac{1}{\sqrt{2I - x^2 - y^2}} \sin\gamma \qquad (4.8d)_\epsilon$$

with any positive λ, the Melnikov function does possess zeros, at least when k is near $k = \sqrt{2}$. We can also verify that the condition in Proposition 4.8 for p_0^∞ to be contained in $\mathcal{R}_0$ is met. Therefore, a symmetric pair of Šilnikov loops exists in this case. We remark that the above system $(4.8)_\epsilon$ corresponds to the Sine-Gordon and Nonlinear Schroedinger equations with the addition of viscous damping as is explained at the end of Chapter 1.

At this point, we briefly discuss the dynamics associated with the existence of the pair of loops homoclinic to p_ϵ. We want to show that an invariant Cantor set of orbits exists near this pair, the dynamics of which can be described by the shift map on the set of biinfinite sequences of the two symbols, say + and −. The two symbols correspond to the passage of the trajectory near the two symmetric homoclinic orbits. In particular, we want to show that for every biinfinite sequence of the symbols + and −, an initial condition exists in the invariant Cantor set whose trajectory can be represented by this sequence. Therefore, the dynamics on the Cantor set can best be characterized as random jumping of the trajectories in it between positive and negative values of the coordinates x and y.

Since the techniques for establishing these facts are fairly standard, we relegate all the rigorous statements about the dynamics in the Cantor set to Appendix G. We remark, however, that the treatment presented there is largely simplified in order to bring out in the clearest possible way the most striking feature of this situation: the

random jumping of the trajectories. Much more precise descriptions of the ensuing dynamics exist (see Wiggins [1988], Chapter 3), but their consequences for what we can observe are the same as those obtained from our simplified treatment.

We also point out the fact that the same type of dynamics are present, not only at the precise parameter values when the pair of homoclinic orbits exist, but also in a little neighborhood of these values in the parameter space where the two orbits are broken. The reason for this will be discussed in Appendix G; however, we want to add that because of this phenomenon, the set of the parameter values for which our phase space exhibits chaotic dynamics has nonzero measure in the parameter space.

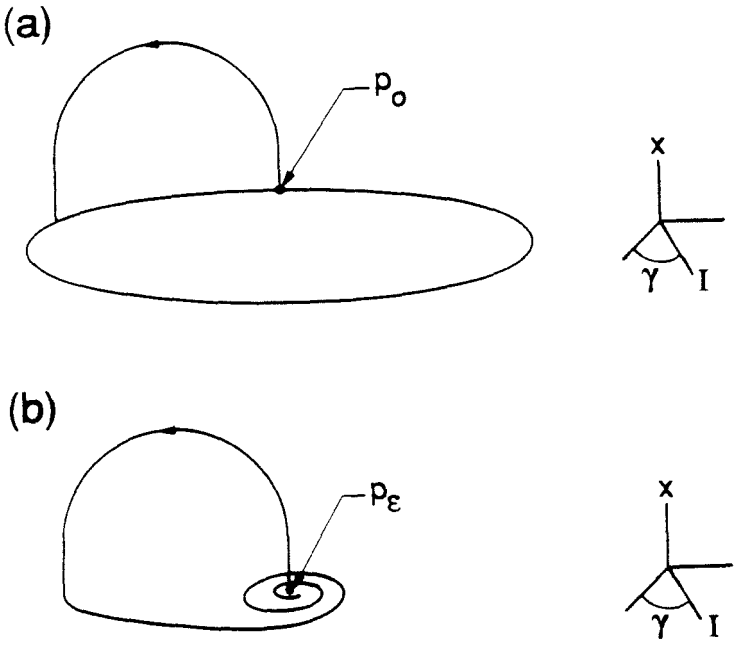

Figure 4.1. The unperturbed heteroclinic orbit emerging from p_0 and the saddle connection at p_ϵ.

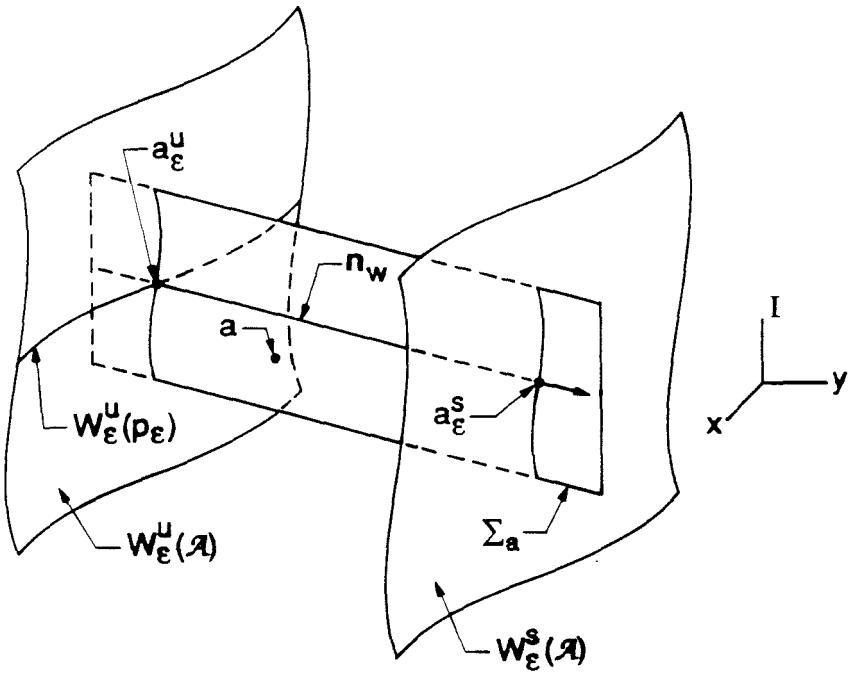

Figure 4.2. Σ_a intersects $W^u_\epsilon(p_\epsilon)$ transversely.

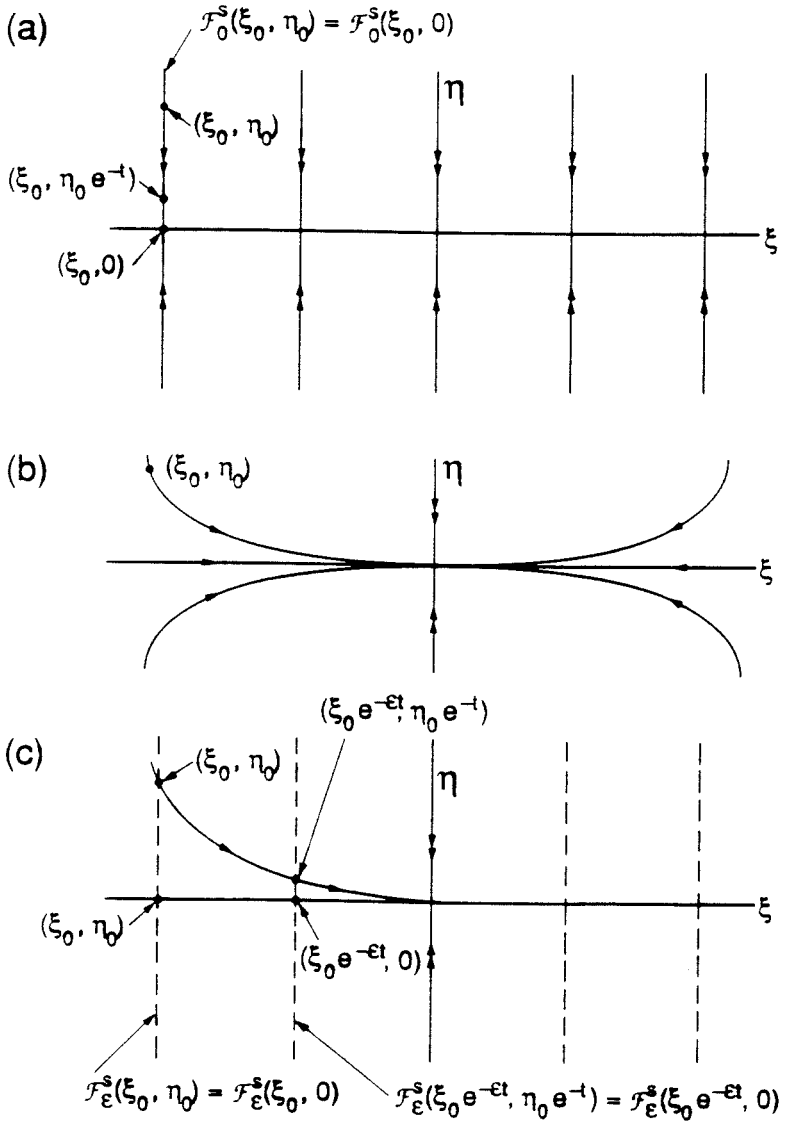

Figure 4.3. Unperturbed and perturbed stable fibers for system $(4.4)_\epsilon$:

(a) unperturbed orbits, (b) perturbed orbits,

(c) stable fibers in the perturbed system.

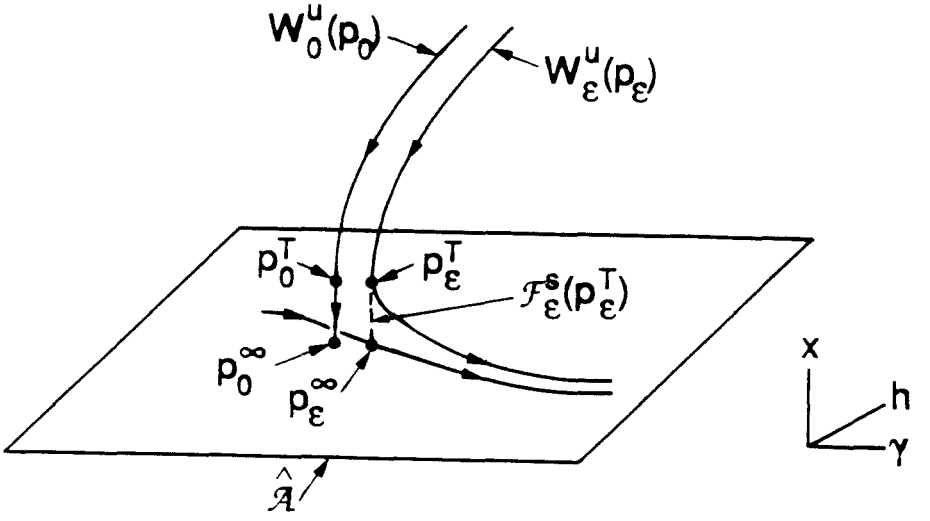

Figure 4.4. Points p_0^T, p_ϵ^T, p_0^∞ and p_ϵ^∞ with their stable fibers.

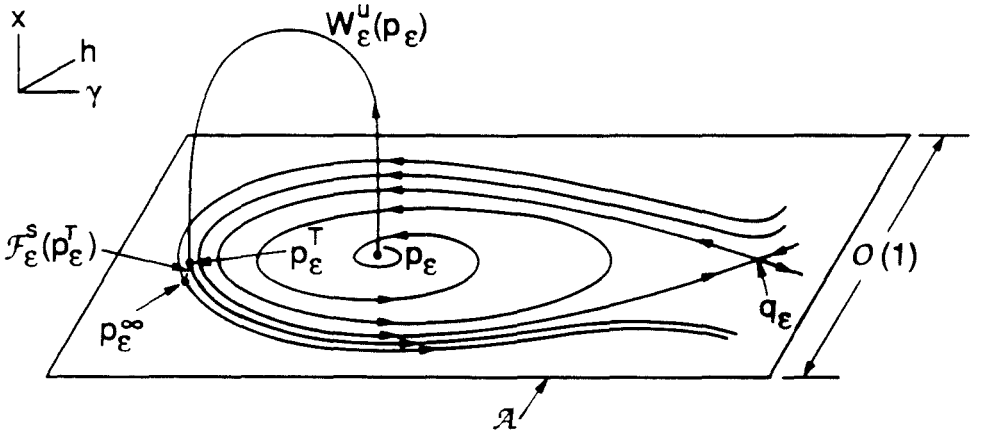

Figure 4.5. Sadle connection at p_ϵ.

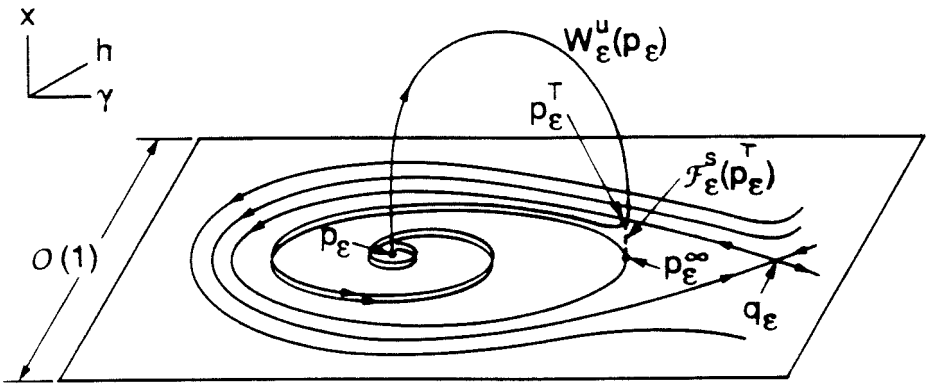

Figure 4.6. The case when $W^u_\epsilon(p_\epsilon)$ leaves $W^s_{\epsilon,loc}(\mathcal{A})$.

APPENDIX A

Perturbed Equations in Alternative Variables

The equations $(3.1)_\epsilon$ in the u, v, I and β coordinates are

$$\dot{u} = \left(k^2 - \tfrac{1}{2}I\right) v - \tfrac{3}{4}u^2 v + \tfrac{1}{4}v^3 - \epsilon\Gamma\cos\beta - \epsilon\alpha u \qquad (A.1a)_\epsilon$$

$$\dot{v} = -\left(k^2 + \tfrac{3}{2}I\right) u + \tfrac{7}{4}u^3 + \tfrac{3}{4}uv^2 + \epsilon\Gamma\sin\beta - \epsilon\alpha v \qquad (A.1b)_\epsilon$$

$$\dot{I} = \epsilon\Gamma(v\sin\beta - u\cos\beta) - 2\epsilon\alpha I \qquad (A.1c)_\epsilon$$

$$\dot{\beta} = 1 + k^2 - \tfrac{3}{2}I - \tfrac{3}{4}u^2 + \tfrac{1}{4}v^2\,. \qquad (A.1d)\epsilon$$

The Hamiltonian in the nondissipative case is given by

$$H = \tfrac{3}{4}I^2 + (1+k^2)I - \tfrac{7}{16}u^4 - \tfrac{3}{8}u^2v^2 + \tfrac{1}{16}v^4$$
$$+ \tfrac{1}{2}(k^2 + \tfrac{3}{2}I)u^2 + \tfrac{1}{2}(k^2 - \tfrac{1}{2}I)v^2 - \epsilon\Gamma(u\sin\beta + v\cos\beta)\,. \qquad (A.2)_\epsilon$$

APPENDIX B

Proof that Transformation (2.4) is Symplectic

Let $p = (p_1, \ldots, p_n)$, $q = (q_1, \ldots, q_n)$, $P = (P_1, \ldots, P_n)$ and $Q = (Q_1, \ldots, Q_n)$.

We say that the transformation $(p, q) \mapsto (P, Q)$ is *symplectic* if

$$\sum_{i=0}^{n} dp \wedge dq = \sum_{i=0}^{n} dP \wedge dQ$$

where $\wedge$ is the usual wedge product from tensor algebra.

The transformation (2.4) is given by

$$c = \sqrt{2I - x^2 - y^2}\, e^{i\gamma}$$

$$b = (x + iy) e^{i\gamma}\,.$$

We verify the fact that it is symplectic by the following direct calculation:

$$
\begin{aligned}
& dc_1 \wedge dc_2 + db_1 \wedge db_2 \\
=& \frac{1}{2i}\{dc^* \wedge dc + db^* \wedge db\} \\
=& \frac{1}{2i}\left\{ \left[(dx - idy)e^{-i\gamma} - i(x - iy)e^{-i\gamma}d\gamma\right] \wedge \left[(dx + idy)e^{i\gamma} + i(x + iy)e^{i\gamma}d\gamma\right] \right. \\
& + \left[\frac{1}{\sqrt{2I - x^2 - y^2}}(dI - xdx - ydy)e^{-i\gamma} - i\sqrt{2I - x^2 - y^2}e^{-i\gamma}\right] \\
& \left. \wedge \left[\frac{1}{\sqrt{2I - x^2 - y^2}}(dI - xdx - ydy)e^{i\gamma} + i\sqrt{2I - x^2 - y^2}e^{i\gamma}\right] \right\} \\
=& \frac{1}{2i}\left\{ 2i(dx \wedge dy) - i(x - iy)d\gamma \wedge (dx + idy) \right. \\
& \left. + i(x + iy)(dx - idy) \wedge d\gamma + 2i(dI - xdx - ydy) \wedge d\gamma \right\} \\
=& dx \wedge dy + dI \wedge d\gamma - \frac{1}{2}d\gamma \wedge (xdx + ydy - iydx + ixdy) \\
& + \frac{1}{2}(xdx + ydy + iydx - ixdy) \wedge d\gamma - (xdx + ydy) \wedge d\gamma \\
=& dx \wedge dy + dI \wedge d\gamma \,.
\end{aligned}
$$

APPENDIX C

Gronwall-Type Estimates

In this appendix we clarify what we mean by a simple Gronwall-type estimate. We first list the Gronwall Lemma. (See Coddington and Levinson, [1955], p. 37, or Guckenheimer and Holmes, [1983], p. 169.)

Lemma C1. *If the functions u,v and c are defined on* $[0,t]$, *c is nonnegative and differentiable and*

$$v(t) \leq c(t) + \int_0^t u(s)v(s)ds$$

then

$$v(t) \leq c(0)e^{\int_0^t u(s)ds} = \int_0^t e^{\int_s^t u(\sigma)d\sigma} ds\,.$$

Suppose now that we have a differential equation

$$\dot{x} = f(x) + \epsilon g(x,t,\epsilon) \qquad (C.1)_\epsilon$$

with $x \in \mathbf{R}^n$. If $x_\epsilon(t)$ is a solution of $(C.1)_\epsilon$ and $x_0(t)$ is a solution of $(C.1)_0$ such that $\|x_\epsilon(0) - x_0(0)\| = \mathcal{O}(\epsilon)$, then we want to show that $\|x_\epsilon(t) - x_0(t)\| = \mathcal{O}(\epsilon)$ for all finite t.

In order to do this, we first find the integral equation which holds for the difference $x_\epsilon(t)$ and $x_0(t)$. Integrating $(C.1)_\epsilon$ and $(C.1)_0$ and subtracting, we obtain

$$x_\epsilon(t) - x_0(t) = x_\epsilon(0) - x_0(0) + \int_0^t \left[f(x_\epsilon(s)) - f(x_0(s))\right] ds + \epsilon \int_0^t g(x_\epsilon(s),s,\epsilon)\, ds\,.$$

If

$$L = \max\{\, \|Df(x)\| \mid x \in \mathcal{D}\}$$

and

$$C = \max \{ \|g(x, s, \epsilon)\| \mid x \in \mathcal{D}, 0 \leq s \leq t \}$$

where $\mathcal{D}$ is some compact region containing $x_\epsilon(s)$ and $x_0(s)$ for all s with $0 \leq s \leq T$, then

$$\|x_\epsilon(t) - x_0(t)\| \leq \|x_\epsilon(0) - x_0(0)\| + L \int_0^t \|x_\epsilon(s) - x_0(s)\| ds + \epsilon C t .$$

We now apply the Gronwall Lemma with $c(t) = \|x_0(t) - x_0(0)\| + \epsilon C t$, $u(t) = L$ and $v(t) = \|x_\epsilon(t) - x_0(t)\|$ to obtain

$$\|x_\epsilon(t) - x_0(t)\|$$
$$\leq \|x_\epsilon(0) - x_0(0)\| e^{Lt} + \epsilon C \int_0^t e^{L(t-s)} ds \leq \left[\|x_\epsilon(0) - x_0(0)\| + \frac{\epsilon C}{L} \right] e^{Lt} .$$

By assumption, the expression in brackets is $\mathcal{O}(\epsilon)$; hence for all finite t we have $\|x_\epsilon(t) - x_0(t)\| = \mathcal{O}(\epsilon)$.

APPENDIX D

Proof of Proposition 4.1

In order to prove Proposition 4.2 we will need Lemma D.1, which we present here. This lemma is a modified version of the stable manifold theorem, which can be found, for instance, in Coddington and Levinson [1955], p. 330, as Theorem 4.1, and the modifications we had to make are taken from Theorems 4.4 and 4.5 on pages 337 and 338, respectively.

In order to present Lemma D.1, we have to set up some terminology. Let

$$\dot{\xi} = A(t,\lambda)\xi + f(\xi,\epsilon,\lambda)\,, \tag{D.1}$$

where $\xi \in \mathbb{C}^n$, $\epsilon \in \mathbb{C}$, $\lambda \in \mathbb{C}^p$. Let $\xi = (\xi_1, \xi_2)$, where $\xi_1 \in \mathbb{C}^m$, $\xi_2 \in \mathbb{C}^{n-m}$ and

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

with the real parts of the eigenvalues of A_1 less than $-\alpha$ and those of A_2 greater than $-\alpha$, where α is a positive number. Assume that $A(\epsilon,\lambda)$ is analytic in $|\epsilon| < \epsilon_0$ and $|\lambda - \bar{\lambda}| < \Lambda$ for some $\bar{\lambda} \in \mathbb{C}$ and positive ϵ_0 and Λ, and that f is analytic for $|\xi| < \xi_0$ and the same range of ϵ and λ as $A(\epsilon,\lambda)$. Also assume that $f(0,\epsilon,\lambda)$ and $D_\xi f(0,\epsilon,\lambda)$ are zero. This implies that for any positive d there exists a positive δ such that

$$|f(\xi) - f(\tilde{\xi})| \le d|\xi - \tilde{\xi}|\,, \tag{D.2}$$

when $|\xi| \le \delta$ and $|\tilde{\xi}| < \delta$.

Lemma D.1. *Let the above assumptions hold. There exists an invariant manifold W in the ξ space with the following properties:*

(1) W contains the origin.

(2) Any solution $\phi(t)$ with $\phi(0) \in W$ satisfies $|\phi(t)| < Ce^{-\alpha t}$ as $t \to \infty$ for some constant C.

(3) There exists a positive η such that a solution ϕ near the origin but not on W at $t = 0$, which satifies $|\phi(t)| \leq \eta$ for all $t > 0$, can not satisfy $|\phi(t)| \leq Be^{-\alpha t}$ for all t with any constant B.

(4) W can be represented as a graph $\xi_2 = \Psi(\xi_1, \epsilon, \lambda)$ with Ψ being analytic in ξ_1, t and λ, and $D_{\xi_1}\Psi(0, \epsilon, \lambda) = 0$.

PROOF: Put

$$U_1(t) = \begin{pmatrix} e^{tA_1} & 0 \\ 0 & 0 \end{pmatrix} \tag{D.3}$$

$$U_2(t) = \begin{pmatrix} 0 & 0 \\ 0 & e^{tA_2} \end{pmatrix}. \tag{D.4}$$

Then

$$e^{tA} = U_1(t) + U_2(t) \tag{D.5}$$

and

$$\dot{U}_j(t) = AU_j(t) \tag{D.6}$$

for $j = 1, 2$.

By assumption, there exist positive constants K and σ, with $\sigma < \alpha$, such that for $t \geq 0$

$$|U_1(t)| \leq Ke^{-(\alpha+\sigma)t}, \tag{D.7}$$

and for $t \leq 0$

$$|U_2(t)| \geq Ke^{-(\alpha-\sigma)t}. \tag{D.8}$$

Consider the integral equation

$$\theta(t,a) = U_1(t)a + \int_0^t U_1(t-s)f(\theta(s,a))ds - \int_t^\infty U_2(t-s)f(\theta(s,a))ds. \tag{D.9}$$

Choose d such that $\frac{2\epsilon k}{\sigma} < \frac{1}{2}$ and let $a = (a_1, 0)$ satisfy $2k|a| < \delta$. Solve (D.9) by successive approximations starting with $\theta_{(0)}(t,a) = 0$. It follows that $\theta_{(1)}(t,a) = U_1(t)a$ and that for positive t,

$$|\theta_{(1)}(t,a) - \theta_{(0)}(t,a)| \leq K|a|e^{-\alpha t}.$$

To make the induction step, assume that

$$|\theta_{(l)}(t,a) - \theta_{(l-1)}(t,a)| \leq \frac{K|a|e^{-\alpha t}}{2^{l-1}} \tag{D.10}$$

for $t \geq 0$. Then

$$
\begin{aligned}
&|\theta_{(l+1)}(t,a) - \theta_{(l)}(t,a)| \\
&\leq K \int_0^t e^{-(\alpha+\sigma)(t-s)} |f(\theta_{(l+1)}(s,a)) - f(\theta_{(l)}(s,a))| ds \\
&\quad + K \int_t^\infty e^{-(\alpha-\sigma)(t-s)} |f(\theta_{(l+1)}(s,a)) - f(\theta_{(l)}(s,a))| ds \\
&\leq \frac{K^2 d|a|}{2^{l-1}} \left[e^{-(\alpha+\sigma)t} \int_0^t e^{(\alpha+\sigma)s} e^{-\alpha s} ds + e^{-(\alpha-\sigma)t} \int_t^\infty e^{(\alpha-\sigma)s} e^{-\alpha s} ds \right] \qquad \text{(D.11)} \\
&= \frac{K^2 d|a|}{2^{l-1}\sigma} e^{-\alpha t} [1 - e^{-\alpha t} + 1] \\
&\leq \frac{2Kd}{\sigma} \frac{|a|K}{2^{l-1}} e^{-\alpha t} \\
&< \frac{|a|K}{2^l} e^{-\alpha t} .
\end{aligned}
$$

Summing the series

$$
\begin{aligned}
&[\theta_{(0)}(t,a) - \theta_{(1)}(t,a)] + [\theta_{(1)}(t,a) - \theta_{(2)}(t,a)] + \cdots \\
&\qquad + [\theta_{(l)}(t,a) - \theta_{(l-1)}(t,a)] + [\theta_{(l+1)}(t,a) - \theta_{(l)}(t,a)] + \cdots
\end{aligned}
$$

then yields the solution $\theta(t,a)$ with

$$
|\theta(t,a)| \leq 2K|a|e^{-\alpha t} . \qquad \text{(D.12)}
$$

The integral

$$
\int_t^\infty U_2(t-s) f(\theta(s,a)) ds
$$

converges uniformly since by (D.12), $|\theta(t,a)| < \delta$ and hence the integrand is

$$
\leq 2Kd|a|e^{-(\alpha-\sigma)(t-s)} e^{-\sigma s} = 2Kd|a|e^{-(\alpha-\sigma)t} e^{-\sigma s} .
$$

Therefore, we may differentiate (D.9) to get

$$
\dot{\theta}(t,a) = A\theta(t,a) + f(\theta(t,a)) .
$$

Note that, because of the uniform convergence of the iteration scheme, θ is an analytic function of a, ϵ, and λ for all $t \geq 0$.

From (D.9) it follows that the first m components of $\theta(t, a)$ satisfy

$$\theta_1(0, a) = a_1$$

and that the latter $n - m$ components are

$$\theta_2(0, a) = -\left[\int_0^\infty U_2(-s) f(\theta(s, a)) ds\right]_2 .$$

Let $\Psi(a_1) = \theta_2(0, a)$. Then the initial values $\xi = \theta(0, a)$ satisfy the equations $\xi_2 = \Psi(\xi_1)$ in the ξ-space, which define the manifold W. By the uniqueness of the solutions of differential equations, if p is any solution with $|p(0)|$ small and $p(0) \in W$, then $p(t) = \theta(t, a)$ for some a where θ is the solution of (D.9) satisfying $\theta(0, a) = p(0)$ and $p(t) \to 0$ as $t \to \infty$.

We will show next that no solution p of (D.1) with $|p(0)|$ small and $p(0)$ not on W can satisfy $|p(t)| \leq \delta$ for all t and $|p(t)| < Be^{-\alpha t}$ for all t and any B. Indeed, suppose that $|p(t)| \leq \delta$ and $|p(t)| < Be^{-\alpha t}$ for some B and for all t. Then it follows from (D.1) that

$$p(t) = e^{tA} p(0) + \int_0^t e^{(t-s)A} f(p(s)) ds .$$

Using (D.3) and (D.4) this can be written as

$$p(t) = U_1(t)p(0) + U_2(t)c + \int_0^t U_1(t-s) f(p(s)) ds - \int_t^\infty U_2(t-s) f(p(s)) ds , \quad \text{(D.13)}$$

where c is the constant vector

$$c = \int_0^\infty U_2(-s) f(p(s)) ds + p(0) .$$

Now $|f(p(s))| < d|p(s)| \leq dBe^{-\alpha s}$ since $|p(s)| < \delta$, and $|U_2(-s)| < Ke^{(\alpha-\sigma)s}$, and therefore, the integral in c converges.

The first term in (D.13) is bounded by $K|p(0)|e^{-\alpha t}$. The third one has the bound

$$KdB \int_0^t e^{-(\alpha+\sigma)(t-s)} e^{-\alpha s} ds = \frac{KdB}{\sigma} e^{-\alpha t}$$

and the fourth one is bounded by

$$KdB \int_t^\infty e^{-(\alpha-\sigma)(t-s)} e^{-\alpha s} ds = \frac{KdB}{\sigma} e^{-\alpha t},$$

which is the same bound as we found for the third term.

The second term, though, is a sum of polynomials, multiplied by exponentials $e^{-\mu t}$ where $\mu < \alpha - \sigma$. This term clearly cannot be bounded by $Be^{-\alpha t}$ for any constant B, unless we have $c = 0$ so that p satisfies (D.9).

To prove the uniqueness of solutions of (D.9), let θ and $\tilde{\theta}$ be solutions for the same a and let $|\theta(t,a)| \leq \delta$ and $|\tilde{\theta}(t,a)| < \delta$. Then (D.9) yields, with (D.2)

$$\begin{aligned}
|\tilde{\theta}(t,a) - \theta(t,a)|e^{\alpha t} &\leq Kde^{\alpha t} \int_0^t e^{-(\alpha+\sigma)(t-s)} |\tilde{\theta}(s,a) - \theta(s,a)| ds \\
&\quad + Kde^{\alpha t} \int_t^\infty e^{-(\alpha-\sigma)(t-s)} |\tilde{\theta}(s,a) - \theta(s,a)| ds \\
&= Kde^{-\sigma t} \int_0^t e^{\sigma s} e^{\alpha s} |\tilde{\theta}(s,a) - \theta(s,a)| ds \\
&\quad + Kde^{\sigma t} \int_t^\infty e^{(\alpha-\sigma)s} |\tilde{\theta}(s,a) - \theta(s,a)| ds .
\end{aligned}$$

If $M = \left\{ e^{\alpha t} |\tilde{\theta}(t,a) - \theta(t,a)| \,\middle|\, t \geq 0 \right\}$, then

$$\begin{aligned}
M &\leq MKde^{-\sigma t} [\frac{e^{\sigma t} - 1}{\sigma}] + MKde^{\sigma t} \frac{e^{-\sigma t}}{\sigma} \\
&\leq \frac{2MKd}{\sigma} .
\end{aligned}$$

Since we have chosen $d < \frac{\sigma}{2K}$, it follows that $M = 0$ and hence $p(0) \in W$.

Apart from taking $C = \delta$ in part (2) of the lemma, we only need to prove that $D_\xi \Psi(0) = 0$. This follows from the estimate

$$\begin{aligned}
|\theta_2(0,a)| &= \left| -\left[\int_0^\infty U_2(t-s) f(\theta(s,a)) ds \right]_2 \right| \\
&\leq dK \int_0^\infty e^{(\alpha-\sigma)s} |\theta(s,a)| ds \\
&\leq 2dK^2 |a| \int_0^\infty e^{-\sigma s} ds \\
&= \frac{2dK^2|a|}{\sigma} .
\end{aligned}$$

This shows that $\Psi(0) = \theta_2(0,0) = 0$ and that for any small enough d, there exists a δ such that if $|a| < \delta$, then

$$\frac{|\Psi(a_1) - \Psi(0)|}{|a_1|} < \frac{2dK^2}{\sigma} ,$$

which implies the assertion.

We now state and prove Proposition D.2, a slightly modified version of Proposition 4.2, of which Proposition 4.2 is a special case. The reason for the modification is that in order to apply Lemma D.1, we must allow ϵ, k, α and Γ to assume complex values. We then restrict ourselves to the normal situation with ϵ, k, α and Γ being positive to recover Proposition 4.2. We remark that $W^u_{\epsilon,loc}(p_\epsilon)$ is the local unstable manifold of p_ϵ only if $\epsilon\alpha \geq 0$. If $\epsilon\alpha < 0$, the unstable manifold of p_ϵ is three-dimensional, but $W^u_{\epsilon,loc}(p_\epsilon)$, as described in Proposition D.2, is still distinguished by the property of being the curve segment along which points approach p_ϵ in backward time with the fastest exponential rate.

Proposition D.2. *For small enough ϵ, there exists a one-dimensional manifold $W^u_{\epsilon,loc}(p_\epsilon)$, the strong local unstable manifold of the fixed point p_ϵ. It depends analytically on ϵ and the parameters k, α, and Γ, and is uniformly $\mathcal{O}(\epsilon)$ close to $W^u_{0,loc}(p_0)$.*

PROOF: In order to use Lemma D.1, we must transform the equations $(3.3)_\epsilon$ into a form to which Lemma D.1 applies. In particular, in the new coordinates, the point p_ϵ must always be at the origin and its unstable eigenspace must coincide with one of the coordinate axes. Moreover, all the transformations involved must be analytic.

We first show that I_ϵ and γ_ϵ, the I and γ coordinates of the point p_ϵ differ from $I = 1$ and $\gamma = \pi - \cos^{-1}\sqrt{2}\frac{\alpha}{\Gamma}$ by an $\mathcal{O}(\epsilon)$ quantity and that they are analytic functions of ϵ. To this end, we write $I_\epsilon = 1 + \epsilon w_\epsilon$, insert this in equation $(3.8)_\epsilon$ with zero left hand side and cancel a factor of ϵ. We end up with a pair of equations

$$\Gamma\sqrt{2(1+\epsilon w_\epsilon)}\cos\gamma_\epsilon + 2\alpha(1+\epsilon w_\epsilon) = 0 \qquad (D.14a)_\epsilon$$

$$-w_\epsilon + \frac{\Gamma}{\sqrt{2(1+\epsilon w_\epsilon)}}\sin\gamma_\epsilon = 0 \qquad (D.14b)_\epsilon$$

which at $\epsilon = 0$ become

$$\Gamma\sqrt{2}\cos\gamma_{\epsilon=0} + 2\alpha = 0 \qquad (D.14a)_0$$

$$-w_{\epsilon=0} + \frac{\Gamma}{\sqrt{2}}\sin\gamma_{\epsilon=0} = 0\,. \qquad (D.14b)_0$$

The solutions of equations $(D.14)_0$ are

$$w_{\epsilon=0} = \pm\frac{\Gamma}{\sqrt{2}}\sqrt{1-2(\tfrac{\alpha}{\Gamma})^2} \qquad (D.15a)$$

$$\gamma_{\epsilon=0} = \pi - \cos^{-1}\sqrt{2}\tfrac{\alpha}{\Gamma} \qquad (D.15b)$$

The Jacobian determinant of $(D.14)_\epsilon$ at $\epsilon = 0$ is

$$\det\begin{pmatrix} 0 & -\Gamma\sqrt{2}\sin\gamma_{\epsilon=0} \\ -1 & \frac{\Gamma}{\sqrt{2}}\cos\gamma_{\epsilon=0} \end{pmatrix} = -\Gamma\sqrt{2}\sin\gamma_{\epsilon=0} = \Gamma\sqrt{2\left(1-2\left(\tfrac{\alpha}{\Gamma}\right)^2\right)},$$

which is nonzero for $\alpha < \frac{\Gamma}{\sqrt{2}}$. The conclusion now follows from the implicit function theorem for analytic functions. It also follows that the coordinates I_ϵ and γ_ϵ of p_ϵ are analytic functions of α and Γ away from $\Gamma = 0$ and $\alpha = \frac{\Gamma}{\sqrt{2}}$.

In order to align the stable subspace of p_ϵ with one of the coordinate axes we first calculate the stability matrix of the system $(3.3)_\epsilon$ at p_ϵ. This matrix is given by

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

with

$$A_1 = \begin{pmatrix} -\epsilon\alpha & -k^2 - \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon \\ -2I_\epsilon - k^2 + \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon & -\epsilon\alpha \end{pmatrix}$$

and

$$A_2 = \begin{pmatrix} -\epsilon\left[\frac{\Gamma}{\sqrt{2I_\epsilon}}\cos\gamma_\epsilon + \alpha\right] & +\epsilon\Gamma\sqrt{2I_\epsilon}\sin\gamma_\epsilon \\ -\frac{\epsilon\Gamma}{(2I_\epsilon)^{\frac{3}{2}}}\sin\gamma_\epsilon & +\frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\cos\gamma_\epsilon \end{pmatrix}$$

and the two zero 2×2 matrices off the main diagonal. At $\epsilon = 0$, the matrix A_1 has positive and negative eigenvalues whose eigenspaces are perpendicular to Π_c. The matrix A_2, in turn, has zero eigenvalues at $\epsilon = 0$ and acts on vectors parallel to Π_c. We only need to align the unstable eigenspace of A with a coordinate axis. Therefore, because of the block diagonal form of A, we only need to rotate the x

and y coordinates associated with A_1. We first find its eigenvalues

$$\lambda = -\epsilon\alpha \pm \sqrt{\left(k^2 - \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon\right)\left(2I_\epsilon - k^2 + \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon\right)}$$

and the corresponding eigenvectors

$$\mathbf{e}_1 = \left(\sqrt{k^2 - \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}, -\sqrt{2I_\epsilon - k^2 + \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}\right)$$

$$\mathbf{e}_2 = \left(\sqrt{k^2 - \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}, \sqrt{2I_\epsilon - k^2 + \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}\right),$$

which are obviously analytic in ϵ, k, α, and Γ.

We now define the new coordinates ξ and η by the requirement that

$$(x, y) = \xi\mathbf{e}_1 + \eta\mathbf{e}_2 .$$

This requirement immediately gives the analytic transformation

$$x = (\xi + \eta)\sqrt{k^2 - \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}$$

$$y = (\eta - \xi)\sqrt{2I_\epsilon - k^2 + \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon},$$

and its inverse

$$\xi = \frac{x}{2\sqrt{k^2 - \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}} - \frac{y}{2\sqrt{2I_\epsilon - k^2 + \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}}$$

$$\eta = \frac{x}{2\sqrt{k^2 - \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}} + \frac{y}{2\sqrt{2I_\epsilon - k^2 + \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}} .$$

We can now apply Lemma D.1 to obtain an analytic strong unstable manifold of the origin in the $(\xi, \eta, I_p, \gamma_p)$ coordinates with $I_p = I - I_\epsilon$ and $\gamma_p = \gamma - \gamma_\epsilon$. This manifold is a graph of the form

$$\eta = \eta(\xi; k, \alpha, \Gamma; \epsilon)$$

$$I_p = I_p(\xi; k, \alpha, \Gamma; \epsilon)$$

$$\gamma_p = \gamma_p(\xi; k, \alpha, \Gamma; \epsilon),$$

with

$$\eta(0; k, \alpha, \Gamma; \epsilon) = 0$$

$$I_p(0; k, \alpha, \Gamma; \epsilon) = 0$$

$$\gamma_p(0; k, \alpha, \Gamma; \epsilon) = 0,$$

and

$$\frac{\partial \eta}{\partial \xi}(0; k, \alpha, \Gamma; \epsilon) = 0$$
$$\frac{\partial I_p}{\partial \xi}(0; k, \alpha, \Gamma; \epsilon) = 0$$
$$\frac{\partial \gamma_p}{\partial \xi}(0; k, \alpha, \Gamma; \epsilon) = 0\,.$$

In order to transform it back to the original coordinates we first express, say, y in terms of x from the equation

$$\frac{x}{2\sqrt{k^2 - \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}} + \frac{y}{2\sqrt{2I_\epsilon - k^2 + \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}}$$
$$-\eta\left(\frac{x}{2\sqrt{k^2 - \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}} - \frac{y}{2\sqrt{2I_\epsilon - k^2 + \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}}, k, \alpha, \Gamma; \epsilon\right) = 0\,.$$

We can do this since the derivative of its left hand side with respect to y is

$$\frac{1}{2\sqrt{2I_\epsilon - k^2 + \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}}\sin\gamma_\epsilon}}\,,$$

which is nonzero.

We thus get $y = y(x, I_\epsilon, \gamma_\epsilon; k, \alpha, \Gamma; \epsilon)$. We then write

$$
\begin{aligned}
I &= I_\epsilon + I_p \left(\frac{x}{2\sqrt{k^2 - \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}} \sin\gamma_\epsilon}} - \frac{y(x, I_\epsilon, \gamma_\epsilon; k, \alpha, \Gamma; \epsilon)}{2\sqrt{2I_\epsilon k^2 - \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}} \sin\gamma_\epsilon}}; k, \alpha, \Gamma; \epsilon \right) \\
&= I(x, I_\epsilon, \gamma_\epsilon; k, \alpha, \Gamma; \epsilon) \\
\gamma &= \gamma_\epsilon + \gamma_p \left(\frac{x}{2\sqrt{k^2 - \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}} \sin\gamma_\epsilon}} - \frac{y(x, I_\epsilon, \gamma_\epsilon; k, \alpha, \Gamma; \epsilon)}{2\sqrt{2I_\epsilon - k^2 + \frac{\epsilon\Gamma}{\sqrt{2I_\epsilon}} \sin\gamma_\epsilon}}; k, \alpha, \Gamma; \epsilon \right) \\
&= \gamma(x, I_\epsilon, \gamma_\epsilon; k, \alpha, \Gamma; \epsilon)\,,
\end{aligned}
$$

to obtain the equations determining $W^u_\epsilon(p_\epsilon)$. Straight forward estimates involving the mean value theorem, similar to those we use in Lemma 4.5, will now show the uniform closeness of $W^u_\epsilon(p_\epsilon)$ and $W^u_0(p_\epsilon)$.

APPENDIX E

Derivation of the Melnikov Function

In this appendix we derive the Melnikov function. We will work with the equations $(3.3)_\epsilon$ recast in more compact notation

$$\dot{\xi} = JD_\xi H(\xi, I; k) + \epsilon g^\xi(\xi, I, \gamma; \alpha, \Gamma) \qquad (E.1a)_\epsilon$$

$$\dot{I} = \epsilon g^I(\xi, I, \gamma; \alpha, \Gamma) \qquad (E.1b)_\epsilon$$

$$\dot{\gamma} = -D_I H(\xi, I; k) + \epsilon g^\gamma(\xi, I, \gamma; \alpha, \Gamma) \qquad (E.1c)_\epsilon$$

with $\xi = (x, y)$, $D_\xi = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$, $D_I = \frac{\partial}{\partial I}$, and

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

In the remainder of this section we will omit the explicit dependence on the parameters k, α, and Γ in order to make the notation less cumbersome.

We want to show that the Melnikov function is given by the formula

$$M(I, \gamma_0; k, \alpha, \Gamma) = \int_{-\infty}^{\infty} \langle \mathbf{n}_W(\xi(s, I), I), g(\xi(s, I), I, \gamma(s, I, \gamma_0)) \rangle ds$$

where $\mathbf{n}_W(\xi, I) = (D_\xi H(\xi, I), D_I H(\xi, I) - D_I H(0, I), 0)$ and $\xi(t, I)$ and

$$\gamma(t, I, \gamma_0) = -\int_0^t D_I H(\xi(s, I), I) ds + \gamma_0$$

are calculated from formulas (2.23) and (2.24), respectively, depending on the value of I.

Before performing the derivation itself, we recall the geometry associated with the splitting of the manifolds. We are interested in the perturbed system in a neighborhood of the unperturbed homoclinic manifold $W(\mathcal{A})$. This manifold is parametrized by

$$W(\mathcal{A}) = \{(\xi(-t_0, I), I, \gamma(-t_0, I, \gamma_0)) \,|\, t_0 \in R, I_1 < I < I_2, 0 \leq \gamma_0 \leq 2\pi\}.$$

At each point $a \in W(\mathcal{A})$ we attach the normal, $\mathbf{n}_W$, to $W(\mathcal{A})$ at a and uniquely determine a_ϵ^s and a_ϵ^u, the points of intersection of the line through $\mathbf{n}_W$ with $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$, between which we measure the distance. We define the signed distance between a_ϵ^s and a_ϵ^u by

$$d(t_0, I, \gamma_0; k, \alpha, \Gamma; \epsilon) = \frac{\langle \mathbf{n}_W(\xi(-t_0, I), I), a_\epsilon^u - a_\epsilon^s \rangle}{\|\mathbf{n}_W(\xi(-t_0, I), I)\|}$$

and Taylor expand it about $\epsilon = 0$ to obtain

$$d(t_0, I, \gamma_0; k, \alpha, \Gamma; \epsilon) = \epsilon \frac{\langle \mathbf{n}_W(\xi(-t_0, I), I), \frac{\partial a_\epsilon^u}{\partial \epsilon} - \frac{\partial a_\epsilon^s}{\partial \epsilon} \rangle}{\|\mathbf{n}_W(\xi(-t_0, I), I)\|} + \mathcal{O}(\epsilon^2)$$

where $d(t_0, I, \gamma_0; k, \alpha, \Gamma; 0) = 0$ since $a_0^s = a_0^u$. The Melnikov function is then defined to be

$$M(I, \gamma_0; k, \alpha, \Gamma) = \left\langle \mathbf{n}_W(\xi(-t_0, I), I), \frac{\partial a_\epsilon^u}{\partial \epsilon}|_{\epsilon=0} - \frac{\partial a_\epsilon^s}{\partial \epsilon}|_{\epsilon=0} \right\rangle .$$

We calculate $M(I, \gamma_0; k, \alpha, \Gamma)$ using the time dependent Melnikov function

$$\tilde{M}(t) = \left\langle \mathbf{n}_W(\xi(t - t_0, I), I), \frac{\partial a_\epsilon^u(t)}{\partial \epsilon}|_{\epsilon=0} - \frac{\partial a_\epsilon^s(t)}{\partial \epsilon}|_{\epsilon=0} \right\rangle ,$$

where $a^s_\epsilon(t) = (\xi^s_\epsilon(t), I^s_\epsilon(t), \gamma^s_\epsilon(t))$ and $a^u_\epsilon(t) = (\xi^u_\epsilon(t), I^u_\epsilon(t), \gamma^u_\epsilon(t))$ are the solutions of $(E.1)_\epsilon$ with the initial conditions $a^u_\epsilon(0) = a^u_\epsilon$ and $a^s_\epsilon(0) = a^s_\epsilon$. Our calculation will therefore involve the quantities

$$\xi_1^{u,s}(t) = \frac{\partial \xi_\epsilon^{u,s}(t)}{\partial \epsilon}|_{\epsilon=0}$$
$$I_1^{u,s}(t) = \frac{\partial I_\epsilon^{u,s}(t)}{\partial \epsilon}|_{\epsilon=0}$$
$$\theta_1^{u,s}(t) = \frac{\partial \theta_\epsilon^{u,s}(t)}{\partial \epsilon}|_{\epsilon=0}\,.$$

which satisfy the first variation equations

$$\dot{\xi_1^{s,u}} = JD^2_\xi H(\xi(t-t_0,I),I)\xi_1^{s,u} + JD_I D_\xi H(\xi(t-t_0,I),I)I_1^{s,u}$$
$$+ g^x(\xi(t-t_0,I),I,\gamma(t-t_0,I,\gamma_0)) \qquad (E.2a)$$

$$\dot{I_1^{s,u}} = g^I(\xi(t-t_0,I),I,\gamma(t-t_0,I,\gamma_0))I_1^{s,u} \qquad (E.2b)$$

$$\dot{\gamma_1^{s,u}} = -D_\xi D_I H(\xi(t-t_0,I),I)\xi_1^{s,u} - D^2_I H(\xi(t-t_0,I),I)I_1^{s,u}$$
$$+ g^\gamma(\xi(t-t_0,I),I,\gamma(t-t_0,I,\gamma_0))\,. \qquad (E.2c)$$

They exist on the time intervals $[t_0,\infty)$ and $(-\infty,t_0]$, respectively.

We split the time-dependent Melnikov function, $\tilde{M}(t)$, into two parts

$$\tilde{M}(t) = \Delta^u(t) - \Delta^s(t) \qquad (E.3)$$

with $\Delta^{s,u}(t) = \langle \mathbf{n}_W(\xi(t-t_0,I),I), a_1^{s,u}(t)\rangle$, and $a_1^{s,u}(t) = (\xi_1^{s,u}(t), I_1^{s,u}(t), \gamma_1^{s,u}(t))$. We then proceed to find a differential equation for $\Delta^{s,u}(t)$. We have

$$\dot{\Delta}^{s,u}(t) = \langle \dot{\mathbf{n}}_W(\xi(t-t_0,I),I), a_1^{s,u}(t)\rangle + \langle \mathbf{n}_W(\xi(t-t_0,I),I), \dot{a}_1^{s,u}(t)\rangle\,, \qquad (E.4)$$

where

$$\dot{\mathbf{n}}_W(\xi(t-t_0,I),I) = \left(D_\xi^2 H\dot{\xi}, D_\xi D_I H\dot{\xi}, 0\right) = (D_\xi^2 H\ JD_\xi H,\ D_\xi D_I H\ JD_\xi H, 0)\ ,$$

and $\dot{a}_1^{s,u}(t)$ are given by the equations $(E.2)$. Note that we suppress the arguments in H for convenience.

Substituting in $(E.4)$ we find

$$\dot{\Delta}^{s,u}(t) = \langle D_\xi^2 H\ JD_\xi H, \xi_1^{u,s}\rangle + D_\xi D_I H\ JD_\xi H I_1$$

$$+ \langle D_\xi^2 H, \xi_1^{u,s}\rangle + \langle D_\xi H, JD_I D_\xi H\ I_1^{u,s}\rangle + \langle \mathbf{n}_W, g\rangle.$$

Now $\langle D_\xi^2 H\ JD_\xi H, \mathbf{v}\rangle = -\langle D_\xi H, JD_\xi^2 H\mathbf{v}\rangle$ for any vector $\mathbf{v}$. Also

$$D_\xi D_I H\ JD_\xi H = \langle D_\xi D_I H, JD_\xi H\rangle = -\langle D_\xi H, JD_\xi D_I H\rangle\,.$$

Therefore

$$\dot{\Delta}^{s,u}(t) = \langle \mathbf{n}_W, g\rangle(\xi(t-t_0,I),I,\gamma(t-t_0,I,\gamma_0))\,.$$

From $(E.3)$ and the definition of the Melnikov function, $M(t)$, we now have

$$\begin{aligned} &M(I,\gamma_0;k,\alpha,\Gamma) \\ &= \tilde{M}(0) \\ &= \Delta^u(0) - \Delta^s(0) \\ &= \int_{-T^u}^{T^s} \langle \mathbf{n}_W, g\rangle(\xi(t-t_0),I,\gamma(t-t_0,I,\gamma_0))dt + \Delta^u(-T_u) - \Delta^s(T_s)\,. \end{aligned}$$

Now $\Delta^s(T_s) = \langle \mathbf{n}_W(\xi(T_s - t_0, I), I), a_1^s(T_s)\rangle$, and for large T_s, the normal $\mathbf{n}_W(T_s - t_0, I), I)$ shrinks to zero exponentially. But $a_1^s(T_s)$ can at best grow linearly as can be seen from the asymptotic form of the first variation equations and the fact that no

exponentially growing solutions are allowed for $\xi_1^s(t)$ since $\xi_\epsilon^s(t)$ must lie on $W_\epsilon^s(\mathcal{A})$. Therefore $\Delta^s(T_s \to \infty) = 0$, and similarly $\Delta^u(-T_u \to -\infty) = 0$. Changing the integration variable from $t - t_0$ to t, we get

$$M(I, \gamma_0; k, \alpha, \Gamma) = \int_{-\infty}^{\infty} \langle \mathbf{n}_W, g \rangle(\xi(t, I), I, \gamma(t, I, \gamma_0))dt\,.$$

We now comment on the fact that the Melnikov function does not contain t_0. This arises due to our choice of the parametrization of $W(\mathcal{A})$ by orbits. We recall that in this parametrization, the ξ, I and γ coordinates of any point a on $W(\mathcal{A})$ are specified by the equations

$$\xi = \xi(t - t_0)$$

$$I = \text{constant}$$

$$\gamma = \gamma(-t_0, I, \gamma_0) = -\int_0^{-t_0} D_I H(x(s), I)ds + \gamma_0\,.$$

with some t_0, I, and γ_0.

We see that γ_0 is not the γ coordinate of the point a. However, the choice of I and γ_0 uniquely determines the orbit on which a lies; *i.e.*, all the other points on this orbit are also described by the same I and γ_0.

This is very convenient for our purpose in the perturbed problem. There, our interest lies in answering the question whether a certain orbit, in particular, the unstable manifold $W_\epsilon^u(p_\epsilon)$ of p_ϵ in Chapter 4, is contained in the intersection of $W_\epsilon^s(\mathcal{A})$ and $W_\epsilon^u(\mathcal{A})$. Outside $W_{\epsilon,loc}^s(\mathcal{A})$ and $W_{\epsilon,loc}^u(\mathcal{A})$ such an intersection orbit can be parametrized via an unperturbed orbit because of the uniqueness of choice

of the point $a^s_\epsilon = a^u_\epsilon$ on the perturbed orbit given a point a on the unperturbed orbit. Therefore, the intersection orbit is also uniquely specified by choosing values of I and γ_0. As a consequence, zeros of the Melnikov function $M(I, \gamma_0; k, \alpha, \Gamma)$ stand in a unique correspondence with the intersection orbits.

Wiggins [1988] chooses a different parametrization of $W(\mathcal{A})$ where he only uses the unperturbed solutions to parametrize the ξ-coordinates:

$$W(\mathcal{A}) = \{(\xi(-t_0, I), I, \bar{\gamma}) \,|\, t_0 \in R, I_1 < I < I_2, 0 \leq \bar{\gamma} \leq 2\pi\}\,.$$

This parametrization is more geometric since for any point on $W(\mathcal{A})$, the parameter $\bar{\gamma}$ is its γ coordinate. However, this parametrization has the disadvantage that points on the same orbit now only share the same I variable. Instead of by the unique second parameter γ_0, the orbit is determined by the equation

$$\bar{\gamma} + \int_0^{-t_0} D_I H(x(s), I) ds = \text{constant}\,.$$

The dependence of the Melnikov function on t_0, I and $\bar{\gamma}$ therefore changes as well.

The solutions which we must now use for calculating the Melnikov function are $\xi(t - t_0)$, I and

$$\gamma(t - t_0, t_0, I, \bar{\gamma}) = -\int_{-t_0}^{t-t_0} D_I H(x(s), I) ds + \bar{\gamma}\,,$$

and the result we obtain is

$$M\left(I, -\int_{-t_0}^{0} D_I H(x(s), I) ds + \bar{\gamma}; k, \alpha, \Gamma\right)$$

where M is the Melnikov function in the t_0-I-γ_0 coordinates. This form of the Melnikov function is particularly suitable for dealing with Poincaré maps such as those for the nonresonant periodic orbits in Section 3.5 since fixing $\bar{\gamma}$ corresponds to fixing the Poincaré section Σ^P, and t_0 then parametrizes $W(\mathcal{A}) \cap \Sigma^P$ However, for determining intersection orbits, our parametrization yields a more economic form of the Melnikov function.

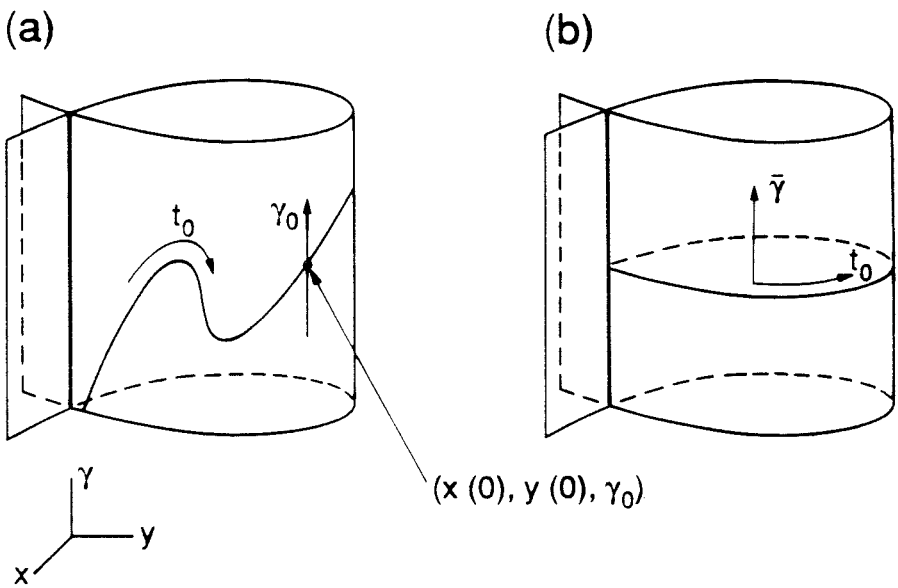

Figure E.1 The two ways of parametrizing $W(\mathcal{A})$.

APPENDIX F

Calculation of the Melnikov Function

In this appendix we calculate the Melnikov function to check for the intersections of $W^s_\epsilon(\mathcal{A})$ and $W^u_\epsilon(\mathcal{A})$. We perform the calculation using the equations $(3.3)_\epsilon$ and the formula

$$M(I,\gamma_0;k,\alpha,\Gamma) = \int_{-\infty}^{\infty} \langle \mathbf{n}_W, g\rangle \, dt$$

derived in the previous appendix.

From $(3.3)_\epsilon$ we compute the components of $\mathbf{n}_W$ to be

$$n^x_W = x\left[(2I - k^2) - \tfrac{7}{4}x^2 - \tfrac{3}{4}y^2\right]$$

$$n^y_W = y\left[-k^2 - \tfrac{3}{4}x^2 + \tfrac{1}{4}y^2\right]$$

$$n^I_W = x^2$$

$$n^\gamma_W = 0\,.$$

We also need the components g^x, g^y, g^I and g^γ of the perturbation g, which are

$$g^x = \Gamma\frac{y}{\sqrt{2I - x^2 - y^2}}\sin\gamma - \alpha x$$

$$g^y = -\Gamma\frac{x}{\sqrt{2I - x^2 - y^2}}\sin\gamma - \alpha y$$

$$g^I = -\Gamma\sqrt{2I - x^2 - y^2}\cos\gamma - 2\alpha I$$

$$g^\gamma = \Gamma\frac{1}{\sqrt{2I - x^2 - y^2}}\sin\gamma\,.$$

After some algebra we find the integrand of the Melnikov function to be equal to

$$\Gamma x\sqrt{2I - x^2 - y^2}(-x\cos\gamma + y\sin\gamma) - 4Ix^2 + \alpha(x^2 + y^2)\left(k^2 + \tfrac{7}{4}x^2 - \tfrac{1}{4}y^2\right).$$

Using formulas (2.12) and (2.19), we transform this expression into

$$-2\sqrt{2}\Gamma B\sqrt{I-B}\cos\theta\cos\beta - 4\alpha I x^2 + \alpha(x^2+y^2)\left(k^2+\tfrac{7}{4}x^2-\tfrac{1}{4}y^2\right)$$

where B, θ and β are given by formulas (2.23) and (2.24), respectively, depending on the value of k, and x and y can be computed from B and θ using (2.12). From (2.23c) and (2.24c) we see that we can write β in the form $\beta(t,I,\beta_0) = \hat{\beta}(t,I) + \beta_0$ where $\hat{\beta}(t,I)$ is an odd function of t. Now, from (2.19), (2.23b), (2.24b) and Figure 2.1 we find that $\beta_0 = \gamma_0 + \frac{\pi}{2}$ for $I > 4k^2$ and that $\beta_0 = \gamma_0$ for $I < 4k^2$. Also, from Figure 2.1 we find that for $I > 4k^2$ the function $x(t,I)$ is odd and the function $y(t,I)$ is even, and that for $I < 4k^2$ the function $y(t,I)$ is odd and the function $x(t,I)$ is even. This implies that the Melnikov function is

$$M(I,\gamma_0;k,\alpha,\Gamma) = \Gamma F(I,k)\cos\gamma_0 + \alpha G(I,k) \qquad (F.1)$$

with

$$F(I,k) = 2\sqrt{2}\int_{-\infty}^{\infty} B\sqrt{I-B}\cos\theta\cos\hat{\beta}\,dt$$

and

$$G(I,k) = \int_{-\infty}^{\infty}\left[-4I\alpha x^2 + \alpha(x^2+y^2)\left(k^2+\tfrac{7}{4}x^2-\tfrac{1}{4}y^2\right)\right]dt\,,$$

for $I > 4k^2$ and

$$F(I,k) = -2\sqrt{2}\int_{-\infty}^{\infty} B\sqrt{I-B}\cos\theta\cos\hat{\beta}\,dt$$

and

$$G(I,k) = \int_{-\infty}^{\infty}\left[-4I\alpha x^2 + \alpha(x^2+y^2)\left(k^2+\tfrac{7}{4}x^2-\tfrac{1}{4}y^2\right)\right]dt$$

for $I < 4k^2$.

The most important case occurs at the resonance, *i.e.*, when $I = 1$. For $k < \frac{1}{2}$ we there have

$$B = \frac{4k^2 \operatorname{sech}^2 k\sqrt{2-k^2}t}{1 - 7\frac{k}{2-k^2}\tanh^2 k\sqrt{2-k^2}t}$$

$$\cot\theta = \frac{k}{\sqrt{2-k^2}}\tanh k\sqrt{2-k^2}t$$

$$\hat{\beta} = -\frac{1}{\sqrt{7}}\tanh^{-1}\left(\sqrt{7}\frac{k}{\sqrt{2-k^2}}\tanh k\sqrt{2-k^2}t\right)$$

$$1 - B = (1-4k^2)\frac{1 + \frac{k}{2-k^2}\tanh^2 k\sqrt{2-k^2}t}{1 - 7\frac{k}{2-k^2}\tanh^2 k\sqrt{2-k^2}t}$$

$$B\sqrt{1-B}\cos\theta = 4k^2\sqrt{1-4k^2}\frac{\frac{k}{\sqrt{2-k^2}}\tanh k\sqrt{2-k^2}t \operatorname{sech}^2 k\sqrt{2-k^2}t}{(1 - 7\frac{k}{2-k^2}\tanh^2 k\sqrt{2-k^2}t)^{\frac{3}{2}}}$$

$$x^2 = 2B\frac{\frac{k}{2-k^2}\tanh^2 k\sqrt{2-k^2}t}{1 + \frac{k}{2-k^2}\tanh^2 k\sqrt{2-k^2}t}$$

$$y^2 = 2B\frac{1}{1 + \frac{k}{2-k^2}\tanh^2 k\sqrt{2-k^2}t}\,.$$

For $k > \frac{1}{2}$ the same expressions become

$$B = \frac{4(2-k^2)\operatorname{sech}^2 k\sqrt{2-k^2}t}{7 - \frac{2-k^2}{k^2}\tanh^2 k\sqrt{2-k^2}t}$$

$$\tan\theta = \frac{\sqrt{2-k^2}}{k}\tanh k\sqrt{2-k^2}t$$

$$\hat{\beta} = -\frac{1}{\sqrt{7}}\tanh^{-1}\left(\frac{1}{\sqrt{7}}\frac{\sqrt{2-k^2}}{k}\tanh k\sqrt{2-k^2}t\right)$$

$$1 - B = (4k^2-1)\frac{1 + \frac{2-k^2}{k^2}\tanh^2 k\sqrt{2-k^2}t}{7 - \frac{2-k^2}{k^2}\tanh^2 k\sqrt{2-k^2}t}$$

$$B\sqrt{1-B}\cos\theta = 4(2-k^2)\sqrt{4k^2-1}\frac{\operatorname{sech}^2 k\sqrt{2-k^2}t}{(7 - \frac{2-k^2}{k^2}\tanh^2 k\sqrt{2-k^2}t)^{\frac{3}{2}}}$$

$$x^2 = 2B\frac{1}{1+\frac{2-k^2}{k^2}\tanh^2 k\sqrt{2-k^2}t}$$

$$y^2 = 2B\frac{\frac{2-k^2}{k^2}\tanh^2 k\sqrt{2-k^2}t}{1+\frac{2-k^2}{k^2}\tanh^2 k\sqrt{2-k^2}t}.$$

From these expressions we can either calculate the Melnikov function numerically or approximate it analyticaly for small values of k and $2-k^2$, respectively.

In order to verify the conditions for $W^u_\epsilon(p_\epsilon)$ to be contained in $W^s_\epsilon(\mathcal{A})$, we must choose

$$\gamma_0 = \bar{\gamma}_0(k,\alpha,\Gamma) = \pi - \cos^{-1}\sqrt{2}\tfrac{\alpha}{\Gamma} - \psi(k)$$

where

$$\psi(k) = \frac{1}{\sqrt{7}}\tanh^{-1}\sqrt{\frac{7k^2}{2-k^2}} - \tan^{-1}\frac{k}{\sqrt{2-k^2}}$$

for $k < \frac{1}{2}$ and

$$\psi(k) = \tan^{-1}\frac{\sqrt{2-k^2}}{k} + \frac{1}{\sqrt{7}}\tanh^{-1}\sqrt{\frac{2-k^2}{7k^2}}$$

for $\frac{1}{2} < k < \sqrt{2}$. The Melnikov function therefore becomes

$$\begin{aligned} &M(1,\bar{\gamma}_0(k,\alpha,\Gamma);k,\alpha,\Gamma) \\ &\quad = \alpha\left(-\sqrt{2}F(1,k)\cos\psi(k)G(1,k)\right) + \sqrt{\Gamma^2-2\alpha^2}F(1,k)\sin\psi(k). \end{aligned} \tag{F.2}$$

The condition that the Melnikov function be zero is therefore given by the equation

$$a(k)\sqrt{2}\tfrac{\alpha}{\Gamma} + b(k)\sqrt{1-2(\tfrac{\alpha}{\Gamma})^2} = 0 \tag{F.3}$$

where

$$a(k) = -F(1,k)\cos\psi(k) + \tfrac{G(1,k)}{\sqrt{2}}$$

and

$$b(k) = F(1,k)\sin\psi(k)\,.$$

The equation $(F.3)$ is a relation between the quantities $\sqrt{2}\frac{\alpha}{\Gamma}$ and k and hence we can calculate $\sqrt{2}\frac{\alpha}{\Gamma}$ in terms of k from it. It is easy to see that the only valid solutions are those for which $\sqrt{2}\frac{\alpha}{\Gamma}$ is nonnegative. (If we allowed negative values as well, the equation $(F.3)$ would change so that we would instead have to solve the original equation $(F.3)$ for $|\sqrt{2}\frac{\alpha}{\Gamma}|$.)

On Figures F.1 and F.2 we plot the values of $a(k)$ and $b(k)$ for $0 < k < \sqrt{2}$. We see that both $a(k)$ and $b(k)$ are always negative, therefore no positive solutions of (F.3) can exist. This implies that in the original system $(1.3)_\epsilon$ we can not find any simple Šilnikov homoclinic loops.

If we repeat the above procedure for equations (4.8), the integrand of the Melnikov function will become

$$\Gamma x\sqrt{2I - x^2 - y^2}(-x\cos\gamma + y\sin\gamma) - 4Ix^2 + \alpha(1+\lambda)(x^2+y^2)\left(k^2 + \tfrac{7}{4}x^2 - \tfrac{1}{4}y^2\right)\,.$$

The expressions for $F(1,k)$ stay the same, however, the expression for $G(1,k)$ will become

$$\tilde{G}(I,k,\lambda) = \int_{-\infty}^{\infty}\left[-4I\alpha x^2 + \alpha(1+\lambda)(x^2+y^2)\left(k^2 + \tfrac{7}{4}x^2 - \tfrac{1}{4}y^2\right)\right]\,dt$$

We define the coefficients $a(k,\lambda)$ and $b(k,\lambda)$ in the same way as we defined $a(k)$ and $b(k)$ and plot them on Figures F.3 and F.4 for $\lambda = 0.1$. We see that there exists an interval of k in which the values of $a(k,\lambda)$ and $b(k,\lambda)$ have opposite signs.

For those values of k we can calculate $\sqrt{2}\frac{\alpha}{\Gamma}$ as a function of k so that the modified Melnikov function $M(1, \bar{\gamma}_0(k, \alpha, \Gamma); k, \lambda, \alpha, \Gamma)$ will have a zero.

We then calculate the difference $\Delta\mathcal{H}$ of the values of the Hamiltonian (3.10) at q_0 and at p_0^∞ with its γ coordinate shifted by a suitable multiple of 2π as in Proposition 4.8. For all values of k for which the Melnikov function has a zero we calculate this difference by first calculating $\sqrt{2}\frac{\alpha}{\Gamma}$ from the modified version of equation $(F.3)$, inserting the corresponding values at q_0 and p_0^∞ into (3.10), and subtracting. The normalized quantity $d(k) = \frac{\Delta\mathcal{H}}{\sqrt{2}\Gamma}$ is plotted on Figure F.5. Note that $d(k)$ is positive throughout its whole interval of definition. This implies that at those values of k for which the Melnikov function $M(1, \bar{\gamma}_0(k, \alpha, \Gamma); k, \lambda, \alpha, \Gamma)$ has zeros, a Šilnikov loop exists at p_ϵ.

We can verify this result independently by a perturbation calculation near $k = \sqrt{2}$. We put $2 - k^2 = \eta$ and obtain

$$B = \tfrac{4}{7}\eta^2 \operatorname{sech}^2 k\eta t + \mathcal{O}(\eta^4)$$

$$\theta = \tfrac{\eta}{k} \tanh k\eta t + \mathcal{O}(\eta^3)$$

$$\hat{\beta} = -\tfrac{\eta}{7k} \tanh k\eta t + \mathcal{O}(\eta^3)$$

$$1 - B = 1 + \mathcal{O}(\eta^2)$$

$$B\sqrt{1-B}\cos\theta = \tfrac{4}{7}\eta^2 \operatorname{sech}^2 k\eta t + \mathcal{O}(\eta^4)$$

$$\cos\hat{\beta} = 1 + \mathcal{O}(\eta^2)$$

$$x^2 = \tfrac{8}{7}\eta^2 \operatorname{sech}^2 k\eta t + \mathcal{O}(\eta^4)$$

$$y^2 = \mathcal{O}(\eta^4).$$

From the above expressions we easily calculate

$$F(1,k) = -\frac{16\sqrt{\eta}}{7k}$$

and

$$\tilde{G}(1,k,\lambda) = \frac{32\eta}{7k}(-1+\lambda).$$

We also have

$$\psi(k) = \frac{8\eta}{7k} + \mathcal{O}(\eta^3) = \frac{8\eta}{7\sqrt{2}} + \mathcal{O}(\eta^3).$$

Therefore, the expression for the Melnikov function becomes

$$M(1,\bar{\gamma}_0(k,\alpha,\Gamma);k,\lambda,\alpha,\Gamma) = \frac{32\eta}{7k}\left(-\frac{4\eta}{7}\sqrt{\Gamma^2-2\alpha^2}+\lambda\alpha\right) + \mathcal{O}(\eta^3).$$

This implies that $\frac{\alpha}{\Gamma} = \frac{4\eta}{7\lambda} + \mathcal{O}(\eta^2)$. We also see that at that value the partial derivative $\frac{\partial M}{\partial \alpha}(1,\bar{\gamma}_0(k,\alpha,\Gamma);k,\lambda,\alpha,\Gamma)$ is equal

$$\frac{\partial M}{\partial \alpha}(1,\bar{\gamma}_0(k,\alpha,\Gamma);k,\lambda,\alpha,\Gamma) = \frac{32\eta\lambda}{7k}$$

which is nonzero. Therefore, the conditions of Proposition 4.3 for $W^u_\epsilon(p_\epsilon)$ to be locally contained in $W^s_\epsilon(\mathcal{A})$ are met.

We can also immediately compute that the value of the Hamiltonian (3.10) at q_0 is $\Gamma\sqrt{2}+\mathcal{O}(\eta)$ and that the value of the Hamiltonian (3.10) at p_0^∞ is $-\Gamma\sqrt{2}+\mathcal{O}(\eta)$. Therefore, their difference is $2\Gamma\sqrt{2} + \mathcal{O}(\eta)$, which matches the results obtained numerically and presented on Figure F.5. We have thus confirmed that the Šilnikov loop exists in this case.

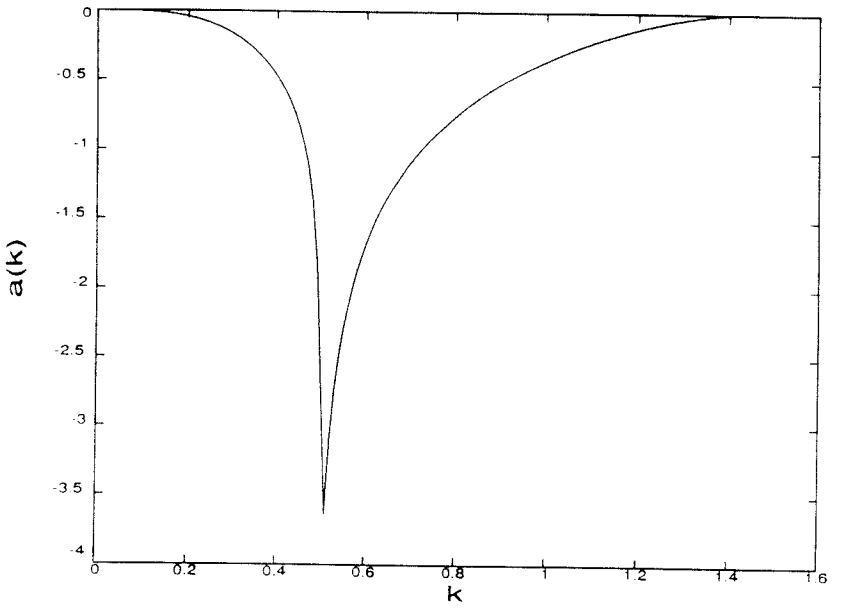

Figure F.1 The coefficient $a(k)$ at $\lambda = 0$.

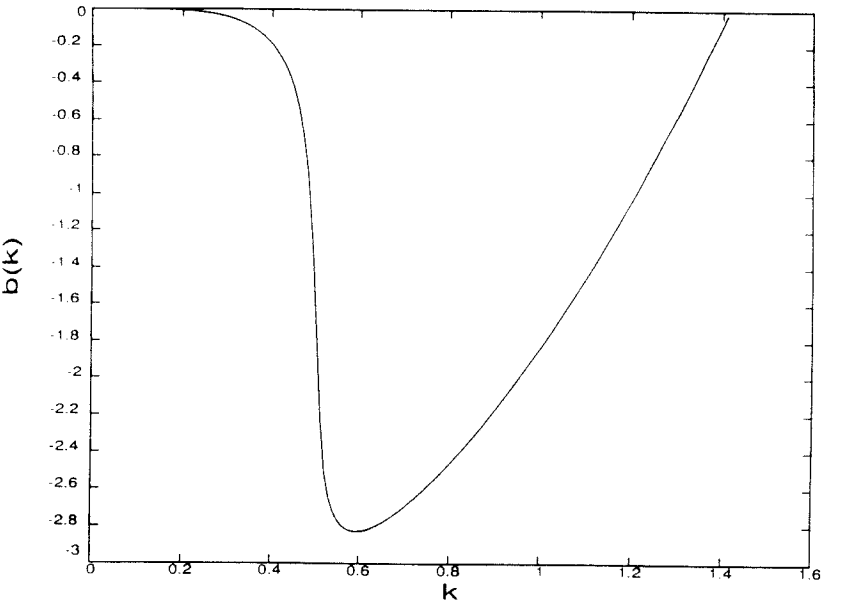

Figure F.2 The coefficient $b(k)$ at $\lambda = 0$.

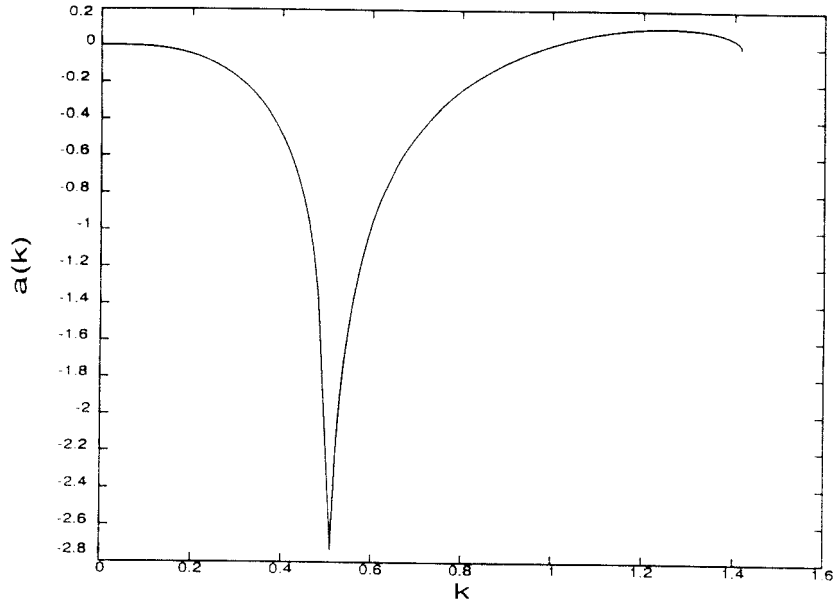

Figure F.3 The coefficient $a(k)$ at $\lambda = 0.1$.

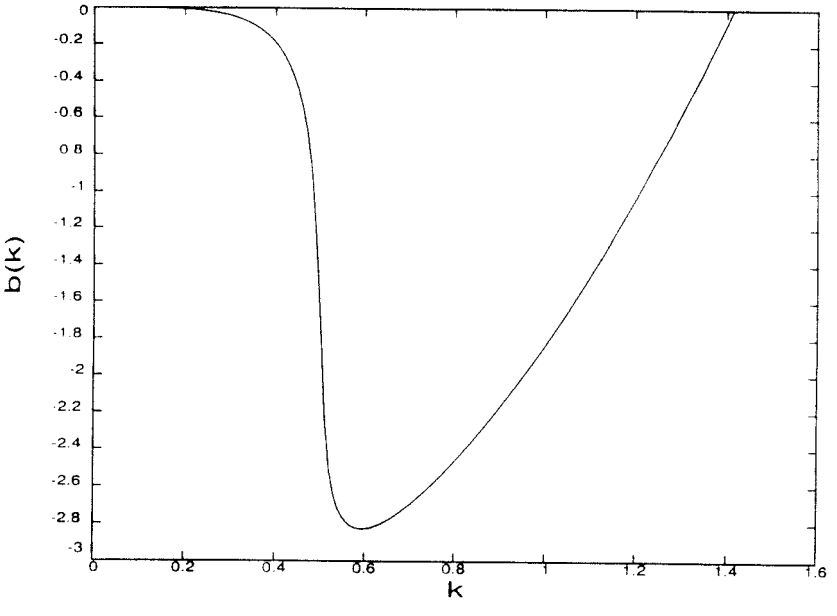

Figure F.4 The coefficient $b(k)$ at $\lambda = 0.1$.

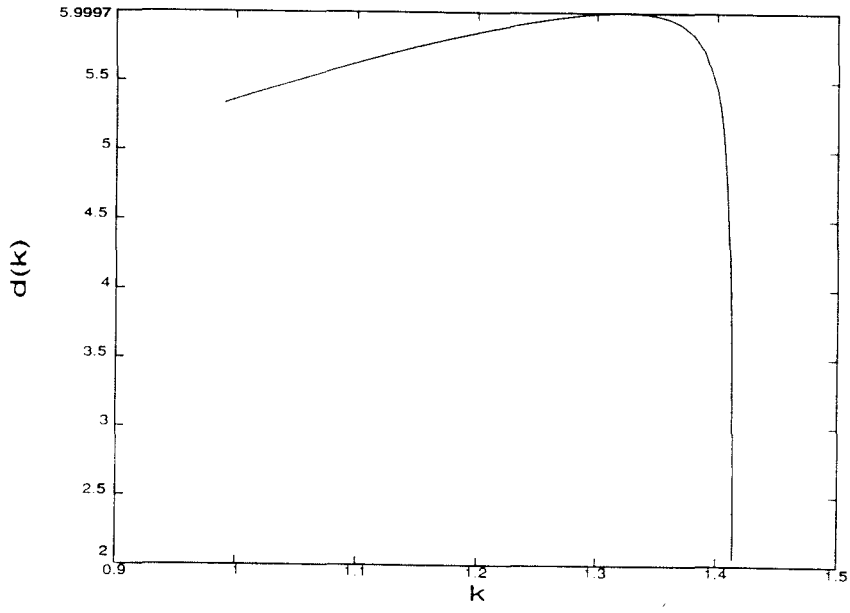

Figure F.5 The coefficient $d(k)$ at $\lambda = 0.1$.

APPENDIX G

Symmetric Šilnikov Example in Four Dimensions

In this appendix we present a discussion of the chaotic dynamics implied by a symmetric pair of Šilnikov loops homoclinic to a fixed point in the four-dimensional phase space. We will show that, under certain conditions imposed on the eigenvalues of the stability matrix at the fixed point, the existence of such a pair of loops will imply the presence of a chaotic invariant set in the phase space. We will consider a more general situation rather than our particular case, since there is no difference in the treatment.

This general situation is described by a system of the form

$$\dot{x} = -\rho x - \omega y + F^x(x, y, z, w) \tag{G.1a}$$

$$\dot{y} = \omega y - \rho x + F^y(x, y, z, w) \tag{G.1b}$$

$$\dot{z} = -\lambda z + F^z(x, y, z, w) \tag{G.1c}$$

$$\dot{w} = \nu w + F^w(x, y, z, w) \tag{G.1d}$$

where ρ, ω, λ and ν are positive and F^x, F^y, F^z and F^w are quadratic at the origin.

It is clear that the origin is a hyperbolic fixed point for (G.1) with the eigenvalues of the linearized vector field given by $-\rho \pm i\omega, -\lambda, \nu$. Hence it has a three-dimensional stable and a one-dimensional unstable manifold. We will assume that, locally, the unstable manifold of the origin coincides with the w axis and the stable manifold coincides with the x-y-z hyperplane. The vector field in our problem

satisfies an additional constraint on the eigenvalues, which is that $\lambda > \nu > \omega > \rho$. We will assume this to be true in the present discussion as well. We remark that, clearly, the choice of the origin as the fixed point we will be interested in, presents no loss of generality. It is less clear that we can even locally align its stable and unstable manifolds with the indicated linear subspaces. This fact needs a proof, which can be found in Wiggins [1988], page 184.

We furthermore assume that the vector field (G.1) is invariant under the transformation $(x, y, z, w) \mapsto (x, y, -z, -w)$, which is exactly what we have in our equations, and that it possesses a homoclinic orbit connecting the origin to itself. By the assumed symmetry there must in fact be two of them, call them Γ^+ and Γ^-, respectively, where the $+$ and the $-$ signs refer to whether the orbit leaves the origin in the positive or negative direction of w. The two orbits can be parametrized by $x = x^\Gamma(t)$, $y = y^\Gamma(t)$, $z = \pm z^\Gamma(t)$ and $w = \pm w^\Gamma(t)$ where $(x^\Gamma(t), y^\Gamma(t), z^\Gamma(t), w^\Gamma(t))$ is, say, a particular solution on Γ^+. By our assumption that $\lambda > \rho$, the pieces of the two orbits near the origin, which lie in the stable manifold, are tangent to the x-y plane.

In order to study the orbit structure of (G.1), we will follow the standard procedure of computing a local Poincaré map P defined in a neighborhood of Γ^+ and Γ^-. For this we first define four cross-sections to the symmetric vector field

(G.1):

$$\Pi_0^+ = \{(x,y,z,w) \,|\, \epsilon e^{-\frac{2\pi\rho}{\omega}} \leq x \leq \epsilon, y = 0, 0 < w \leq \epsilon, -\epsilon \leq z \leq \epsilon\}$$

$$\Pi_0^- = \{(x,y,z,w) \,|\, \epsilon e^{-\frac{2\pi\rho}{\omega}} \leq x \leq \epsilon, y = 0, -\epsilon \leq w < 0, -\epsilon \leq z \leq \epsilon\}$$

$$\Pi_1^+ = \{(x,y,z,w) \,|\, w = \epsilon\}$$

$$\Pi_1^- = \{(x,y,z,w) \,|\, w = -\epsilon\}.$$

For simpler notation, we also define $\Pi_0 = \Pi_0^+ \cup \Pi_0^-$. The Poincaré map P will take a point p on Π_0 and follow its trajectory as it passes by the origin, leaves a neighborhood of the origin, staying close to a solution on one of two homoclinic orbits until it returns to a point $P(p)$ on Π_0. Of course, we do not consider any points in Π_0 which the time flow of (G.1) will not bring back to Π_0.

We want to show that there exits an invariant Cantor set for P on Π_0, and that the action of P on the points in that Cantor set is topologically conjugate to the action of the shift map on the set of all biinfinite sequences of the symbols $+$ and $-$.

The exact calculation of the map P is obviously impossible, since we would have to solve (G.1) in order to find the trajectories of points in Π_0. However, we can approximate P by a simpler map P^L, which can be computed, and argue that the existence of a chaotic invariant Cantor set on Π_0 for P^L implies the existence of such a set for P. We will only compute P^L and show it possesses a chaotic Cantor set. For the proof that these results carry over to the exact map P, see Proposition 3.2.10 in Wiggins [1988], p. 198.

In order to approximate P we first represent it as a composite of four maps P_0^+, P_0^-, P_1^+ and P_1^-. Here P_0^+ is defined as the map which takes a point on Π_0^+ and brings it to Π_1^+. The map P_1^+ then brings points on Π_1^+ back to Π_0^+. The maps P_0^- and P_0^- are defined analogously on the negative side. The map P therefore coincides with the composite $P^+ = P_1^+ \circ P_0^+$ on Π_0^+ and with the composite $P^- = P_1^- \circ P_0^-$ on Π_0^-.

In order to define P^L we linearize the maps P^+ and P^- around the origin to obtain the approximate maps P_0^{L+} and P_0^{L-}, mapping Π_0^+ and Π_0^- into Π_1^+ and Π_1^-, respectively. To obtain P_1^{L+} and P_1^{L-} , we linearize P_1^+ and P_1^- around the intersection points of Π_1^+ and Π_1^- with Γ^+ and Γ^-, respectively; again, P^L coincides with the composite $P^{L+} = P_1^{L+} \circ P_0^{L+}$ on Π_0^{L+} and with the composite $P^{L-} = P_1^{L-} \circ P_0^{L-}$ on Π_0^-.

We now construct P^{L+} using the linearized flow of (G.1), given by

$$x(t) = e^{-\rho t}(x_0 \cos\omega t - y_0 \sin\omega t) \qquad (G.2a)$$

$$y(t) = e^{-\rho t}(x_0 \sin\omega t + y_0 \cos\omega t) \qquad (G.2b)$$

$$z(t) = z_0 e^{-\lambda t} \qquad (G.2c)$$

$$w(t) = w_0 e^{\nu t}\,. \qquad (G.2d)$$

For every point $(x, y, z, w) = (x_0, 0, z_0, w_0)$ on Π_0^+ we calculate from (G.2d) the time T it takes to reach Π_1^+. We then insert T into the other equations in (G.2) to obtain the image of that point under P^{L+}. Note that we have carefully chosen the x width of Π_0^+ so that orbits starting on Π_0^+ do not reintersect Π_0^+ before leaving a

neighborhood of the origin.

For any point in Π_0^+, the time of flight T from Π_0^+ to Π_1^+ is found by solving

$$\epsilon = w_0 e^{\nu T}$$

from which we obtain

$$T = \frac{1}{\nu} \log \frac{\epsilon}{w_0} . \tag{G.3}$$

Using (G.2) and (G.3) and dropping the subscript 0 in the coordinates, we compute the map P_0^{L+} from Π_0^+ to Π_1^+ to be

$$\begin{pmatrix} x \\ 0 \\ z \\ w \end{pmatrix} \mapsto \begin{pmatrix} x \left(\frac{w}{\epsilon}\right)^{\frac{\rho}{\nu}} \cos\left(\frac{\omega}{\nu} \log \frac{\epsilon}{w}\right) \\ x \left(\frac{w}{\epsilon}\right)^{\frac{\rho}{\nu}} \sin\left(\frac{\omega}{\nu} \log \frac{\epsilon}{w}\right) \\ z \left(\frac{w}{\epsilon}\right)^{\frac{\lambda}{\nu}} \\ \epsilon \end{pmatrix} . \tag{G.4}$$

To get an idea of the geometry of $P_0^{L+}(\Pi_0^+)$, it will be useful to consider a foliation of Π_0^+ by the slabs

$$R_k^+ = \{(x, y, z, w) \,|\, \epsilon e^{-\frac{2\pi\rho}{\omega}} \leq x \leq \epsilon, -\epsilon < z \leq \epsilon, \epsilon e^{-\frac{2\pi(k+1)\nu}{\omega}} \leq w \leq e^{-\frac{2\pi k\nu}{\omega}} \} .$$

Then we have

$$\Pi_0^+ = \bigcup_{k=0}^{\infty} R_k^+ .$$

It will also be useful to coordinatize the x-y part of Π_1^+ by polar coordinates. Denoting the x, y, z coordinates on Π_1^+ by x', y', z' in order to avoid confusion with the coordinates on Π_0^+ we have

$$r = \sqrt{x'^2 + y'^2}, \qquad \tan\theta = \frac{y'}{x'}$$

and in these coordinates P_0^{L+} is written as

$$\begin{pmatrix} x \\ 0 \\ z \\ w \end{pmatrix} \mapsto \begin{pmatrix} r \\ \theta \\ z' \\ \epsilon \end{pmatrix} = \begin{pmatrix} x\left(\frac{w}{\epsilon}\right)^{\frac{\rho}{\omega}} \\ \frac{\omega}{\nu}\log\frac{\epsilon}{w} \\ z\left(\frac{w}{\epsilon}\right)^{\frac{\lambda}{\nu}} \\ \epsilon \end{pmatrix}. \tag{G.5}$$

Now consider an $R_k^+ \subset \Pi_0^+$ for fixed k. Using (G.5) we make the following observations concerning $P_0^{L+}(R_k^+)$:

i) The two-dimensional sheets w = constant contained in R_k^+ are mapped to θ = constant under P_0^{L+}.

ii) The two vertical boundaries of R_k^+ that are parallel to the w-z plane are mapped to two-dimensional logarithmic spirals.

iii) The two-dimensional sheet $z = 0$ contained in R_k^+ is mapped to $z' = 0$ in Π_1^+.

We see that $P_0^{L+}(R_k^+)$ has roughly the shape of a hollow cylinder.

Defining P_1^{L+}, note that we can take it to be linear, the approximation being valid in an open set $U^+ \subset \Pi_1^+$. There $P_1^{L+} : U^+ \subset \Pi_1^+ \to \Pi_0^+$ is given by

$$\begin{pmatrix} x \\ y \\ z \\ \epsilon \end{pmatrix} \mapsto \begin{pmatrix} a & b & c & 0 \\ 0 & 0 & 0 & 0 \\ d & e & f & 0 \\ g & h & i & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ \epsilon \end{pmatrix} + \begin{pmatrix} \bar{x} \\ 0 \\ \bar{z} \\ 0 \end{pmatrix} \tag{G.6}$$

where $(\bar{x} = \epsilon(1 + e^{-2\pi\rho})/2, 0, \bar{z}, 0) = \Gamma^+ \cap \Pi_0^+$ and $\bar{z} = \mathcal{O}(\epsilon^2)$.

The upper half of the Poincaré map $P^{L+} = P_1^{L+} \circ P_0^{L+} : V^+ \subset \Pi_0^+ \to \Pi_0^+$ is thus computed with the aid of (G.4) and (G.6) to be

$$\begin{pmatrix} x \\ z \\ w \end{pmatrix} \mapsto \begin{pmatrix} ax\left(\frac{w}{\epsilon}\right)^{\frac{\rho}{\nu}}\cos\left(\frac{\omega}{\nu}\log\frac{\epsilon}{w}\right) + bx\left(\frac{w}{\epsilon}\right)^{\frac{\rho}{\nu}}\sin\left(\frac{\omega}{\nu}\log\frac{\epsilon}{w}\right) + cz\left(\frac{w}{\epsilon}\right)^{\frac{\lambda}{\nu}} + \bar{x} \\ dx\left(\frac{w}{\epsilon}\right)^{\frac{\rho}{\nu}}\cos\left(\frac{\omega}{\nu}\log\frac{\epsilon}{w}\right) + ex\left(\frac{w}{\epsilon}\right)^{\frac{\rho}{\nu}}\sin\left(\frac{\omega}{\nu}\log\frac{\epsilon}{w}\right) + fz\left(\frac{w}{\epsilon}\right)^{\frac{\lambda}{\nu}} + \bar{z} \\ gx\left(\frac{w}{\epsilon}\right)^{\frac{\rho}{\nu}}\cos\left(\frac{\omega}{\nu}\log\frac{\epsilon}{w}\right) + hx\left(\frac{w}{\epsilon}\right)^{\frac{\rho}{\nu}}\sin\left(\frac{\omega}{\nu}\log\frac{\epsilon}{w}\right) + iz\left(\frac{w}{\epsilon}\right)^{\frac{\lambda}{\nu}} \end{pmatrix} \tag{G.16}$$

where $V^+ = (P_0^+)^{-1}(U^+)$. If we choose Π_0^+ sufficiently small then $P^{L+}(R_k^+)$ appears as in Figure G.7.

We will now proceed to examine the geometry of $P_0^{L+}(R_k^+)$ more carefully. The maximal and minimal values of w in R_k^+ and z' and r in $P_0^{L+}(R_k^+)$ are

$$w_{max} = \epsilon e^{-\frac{2\pi k\nu}{\omega}} \qquad w_{min} = \epsilon e^{-\frac{2\pi(k+1)\nu}{\omega}}$$

$$z'_{max} = \epsilon e^{-\frac{2\pi k\lambda}{\omega}} \qquad z'_{min} = \epsilon e^{-\frac{2\pi(k+1)\lambda}{\omega}}$$

$$r_{max} = \epsilon e^{-\frac{2\pi k\rho}{\omega}} \qquad r_{min} = \epsilon e^{-\frac{2\pi(k+1)\rho}{\omega}}$$

respectively. Since we assumed $\lambda > \nu > \rho$ we have

$$\frac{z'_{max}}{w_{min}} = e^{\frac{2\pi\nu}{\omega}} e^{\frac{2\pi k(\nu-\lambda)}{\omega}} \to 0$$

$$\frac{r_{min}}{w_{max}} = e^{-\frac{2\pi\rho}{\omega}} e^{\frac{2\pi k(\nu-\rho)}{\omega}} \to \infty$$

$$\frac{z'_{max}}{r_{min}} = e^{\frac{2\pi\rho}{\omega}} e^{\frac{2\pi k(\rho-\lambda)}{\omega}} \to 0$$

as $k \to \infty$. By a well known theorem in linear algebra, every linear transformation is the product of a rotation, an inversion, and a stretching. In our case the map P_1^{L+} is orientation preserving, so there is no inversion; moreover it is nonsingular. Hence $P^{L+}(R_k^+)$ is just a deformed and rotated version of $P_0^{L+}(R_k^+)$, *i.e.*, it has roughly the shape of a cracked cylinder. Generically its "symmetry axis" is neither parallel nor orthogonal to the x-z plane. This fact will enable us to find the proper horizontal and vertical slabs, with the aid of which we will construct a chaotic invariant set.

We repeat the same procedure on the negative to find the map P^{L-} and thus complete the construction of P^L.

Having constructed P^L, we want to show that it possesses a chaotic invariant

set. Our discussion here will be rather heuristic, and we refer to Wiggins [1988] for the rigorous details.

In order to construct the chaotic invariant set, we must find two horizontal slabs H^+ and H^- in Π_0, which will be mapped onto two vertical slabs V^+ and V^- in Π_0 in such a manner that both V^+ and V^- intersect both H^+ and H^-. The terms "horizontal slab" and "vertical slab" have precise definitions as described in Wiggins [1988], Chapter 2. We, however, will proceed by defining H^+, H^-, V^+ and V^- and refer to Wiggins [1988] in order to check that they fit into the precise definitions and that the way they intersect is proper for the general framework described there.

We now define H^+, H^-, V^+ and V^-. We first define V^+ and V^-. In order to do this we first note that the asymptotic estimates of the size of the various coordinates given above ensure that the inner boundary of $P^{L+}(R_k^+)$ intersects the upper horizontal boundary of R_k^+ and the lower horizontal boundary of R_k^- each in two separate curves (see Figure G.8). This means that $P^{L+}(R_k^+)$ intersects R_k^+ and R_k^- each in two separate pieces. Choose one of the two pieces of $P^{L+}(R_k^+) \cap R_k^+$, the piece of $P^{L+}(R_k^+) \cap R_k^-$ directly underneath it and the connecting piece to form the vertical slab V^+. For the slab V^- choose the mirror image of V^+ obtained by interchanging $+$ and $-$ in the construction of V^+. The horizontal slabs H^+ and H^- will be the preimages of the vertical slabs V^+ and V^-, respectively.

We now have two horizontal slabs, H^+ and H^-, being mapped onto two vertical

slabs V^+ and V^-. The orbit carrying any point in H^+ to V^+ is close to the homoclinic orbit Γ^+ and the orbit carrying any point in H^- to V^- is close to the homoclinic orbit Γ^-. We construct the invariant set Λ for our mapping P^L as the intersection of all the forward images of V^+ and V^- and all the backward images of H^+ and H^-. The intersection of all the forward images of V^+ and V^- turns out to be a Cantor set of lines, and the intersection of all the backward images of H^+ and H^- a Cantor set of surfaces. The set Λ, their intersection, is therefore a Cantor set of points. The details of this construction are the same as those for the standard Smale Horseshoe map, which are well known and presented in many standard texts, such as Guckenheimer and Holmes [1985], Wiggins [1988], and Wiggins [1989]. The only detail we need from this construction is the fact that we assign to each point p in Λ an infinite sequence $\{\cdots, s_{-2}, s_{-1}, s_0, s_1, s_2, \cdots\}$ where $s_i \in \{+, -\}$ for all i and where $p \in H_{s_0}$, $P^L(p) \in H_{s_1}$, $(P^L)^2(p) \in H_{s_2}$, *etc.*, and $p \in V_{s_{-1}}$, $(P^L)^{-1}(p) \in V_{s_{-1}}$, $(P^L)^{-2}(p) \in V_{s_{-2}}$, *etc.* But, $(P^L)^{-1}(V_i) = H_i$; therefore we can interpret the sequence of the $+$ and $-$ to mean that $(P^L)^k(p) \in H_{s_k}$ for every integer k, which means that during the k-th iteration of the map P^L, our point will follow close to the orbit Γ^{s_k}. Since all biinfinite sequences of $+$ and $-$ are possible, we see that we have thus encoded the random jumping of the points in Λ near the upper and lower homoclinic orbits.

At the end we remark that neither the Poincaré map P nor the the resulting chaotic Cantor set are sensitive to small perturbations unlike the Šilnikov orbits

which break immediately. Therefore, the chaotic Cantor set exists in a neighborhood of every value of the parameters at which the two Šilnikov orbits exist. The details of this statement can be found in Wiggins [1988], p. 240 and the reference cited there. We note, however, that this discussion implies the existence of a chaotic Cantor set for a subset of the parameter space with nonzero measure.

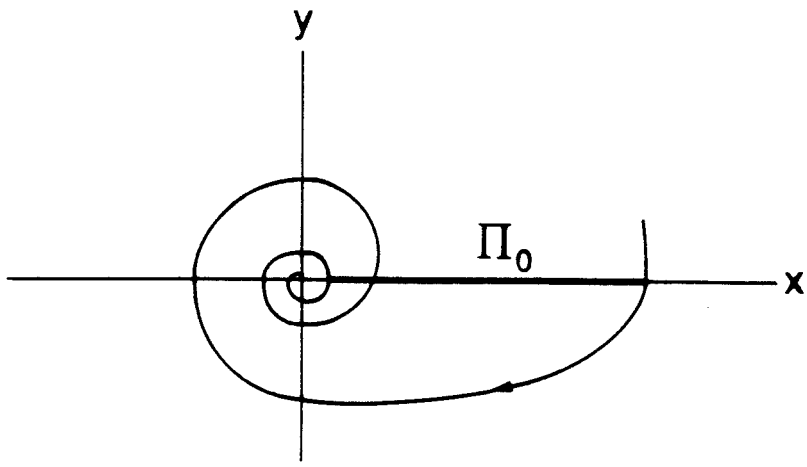

Figure G.1. Intersection of Π_0 with the x-y plane.

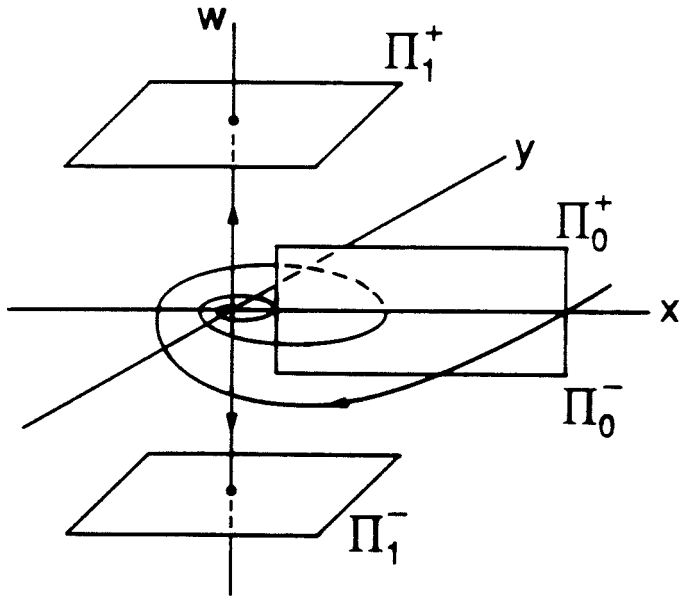

Figure G.2. The cross sections Π_0^+, Π_0^-, Π_1^+ and Π_1^-.

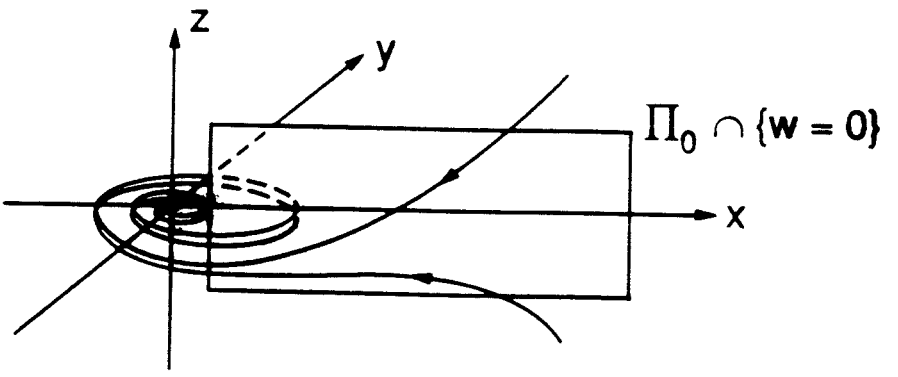

Figure G.3. Intersection of Π_0 with the x-y-z hyperplane.

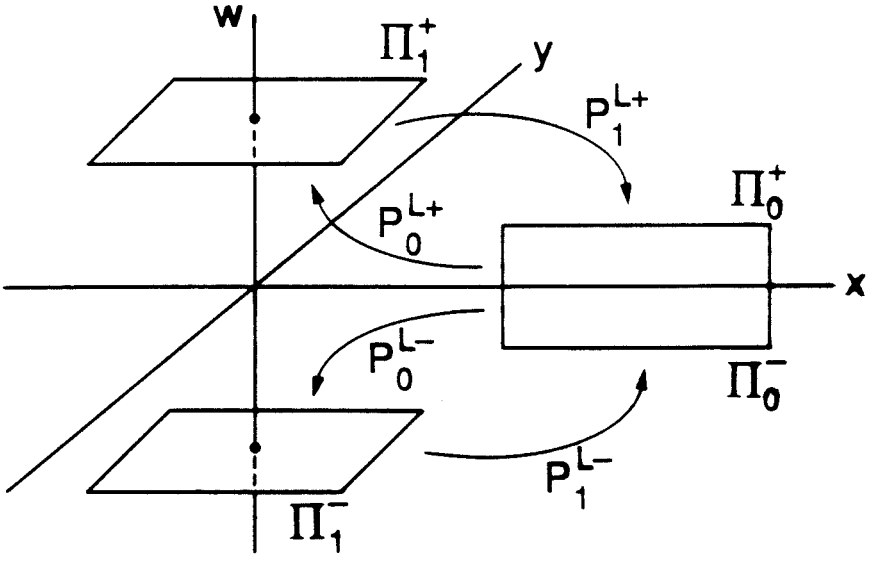

Figure G.4. Maps P_0^{L+}, P_0^{L-}, P_1^{L+} and P_1^{L-}.

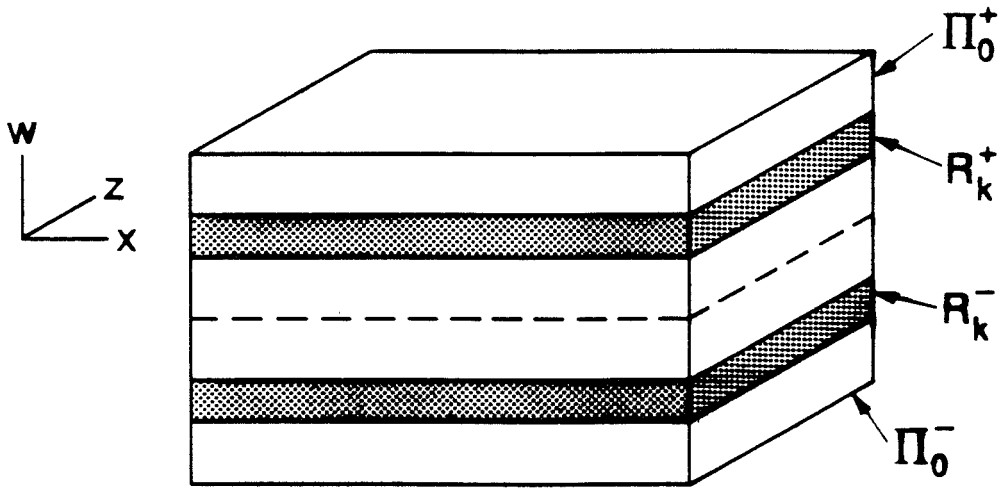

Figure G.5. Horizontal slabs R_k^+ and R_k^-.

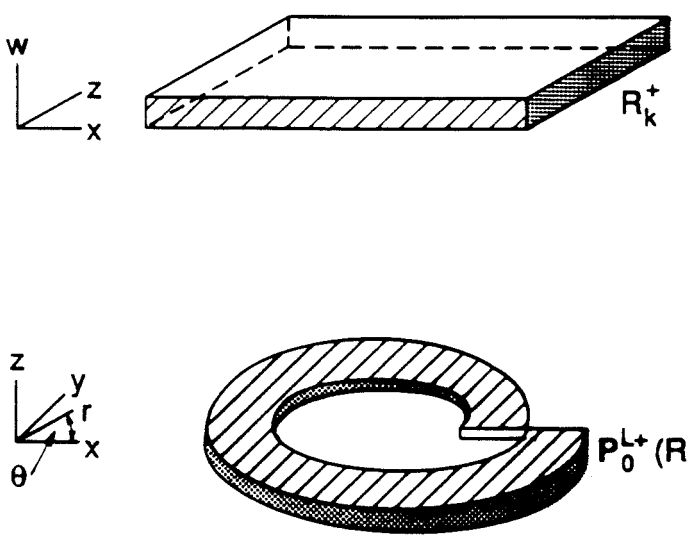

Figure G.6. The slab R_k^+ and its image $P_0^{L+}(R_k^+)$.

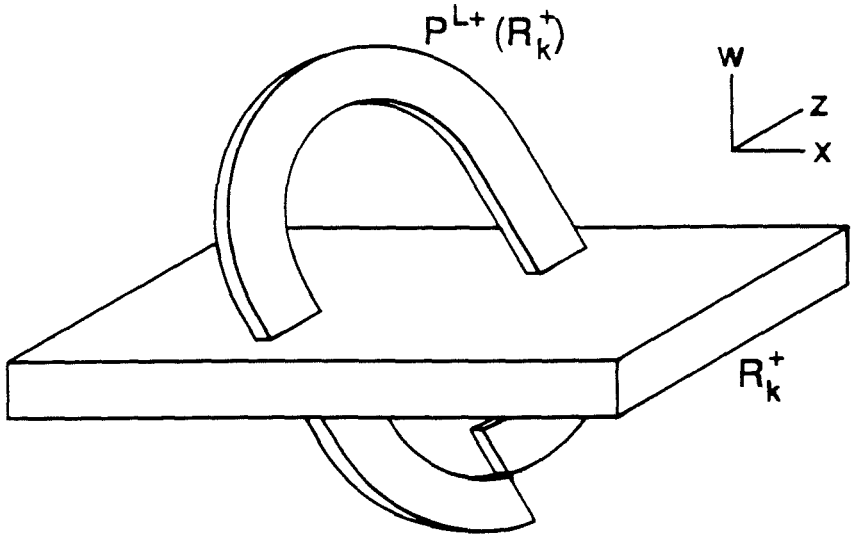

Figure G.7. The slab R_k^+ and its image $P^{L+}(R_k^+)$.

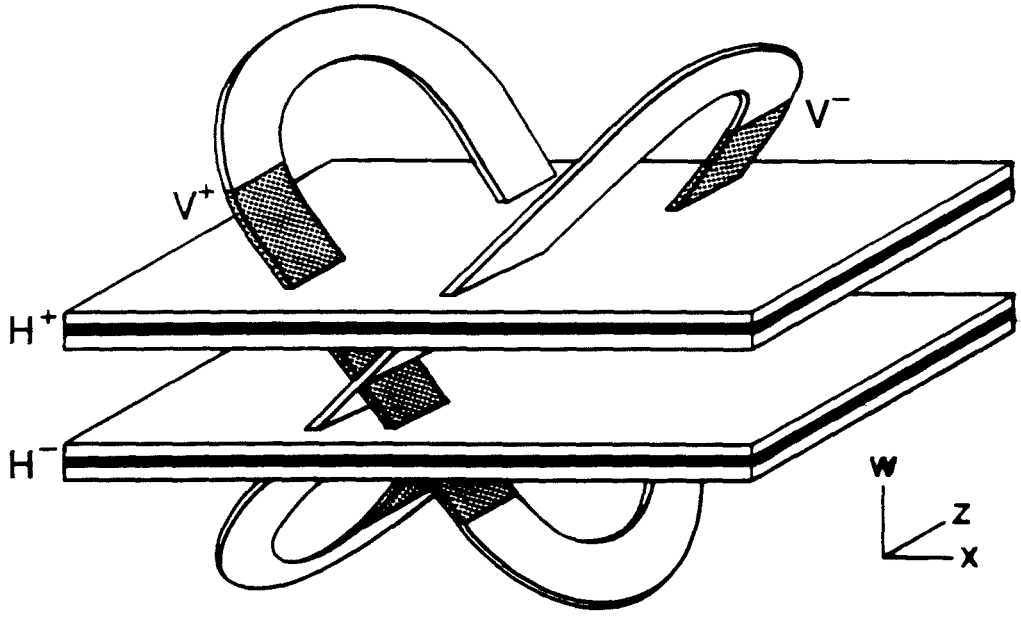

Figure G.8. The horizontal slabs H^+ and H^- with the vertical slabs V^+ and V^-.

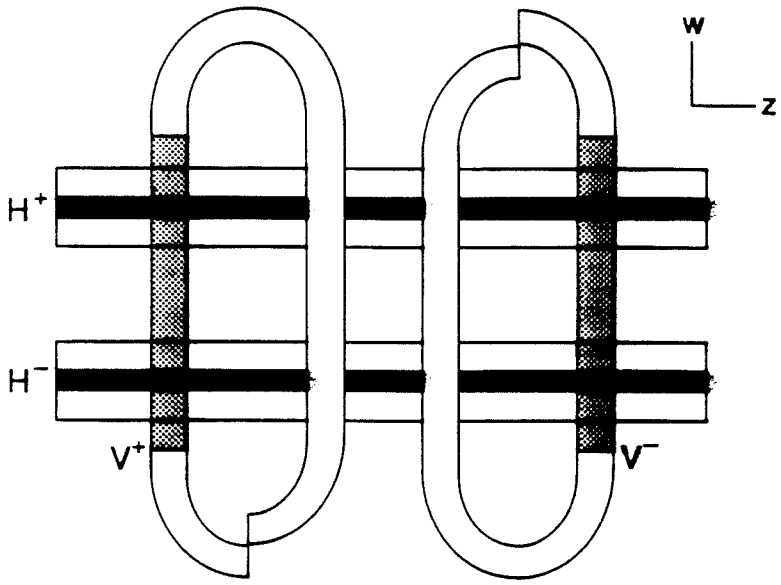

Figure G.9. The w-z components of the horizontal slabs H^+ and H^- and the vertical slabs V^+ and V^-.

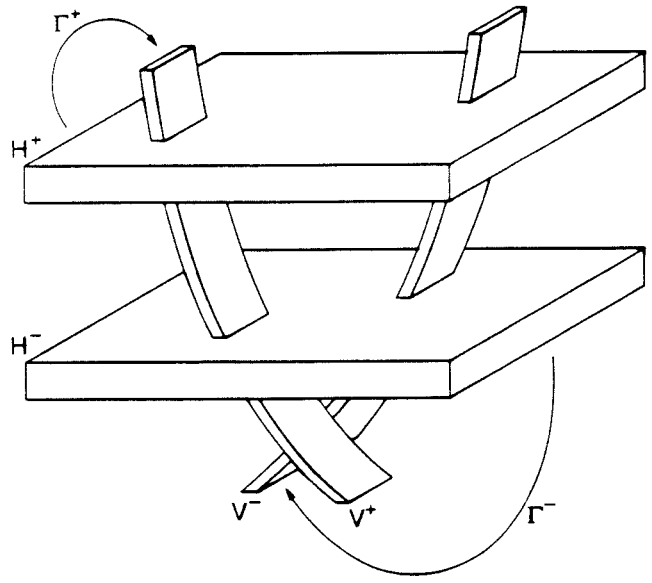

Figure G.10. The horizontal slabs H^+ and H^- are mapped onto the vertical slabs V^+ and V^-.

REFERENCES

Arnold, V.I. [1988] ed. *Dynamical Systems III.* Springer-Verlag: New York, Heidelberg, Berlin.

Bishop, A.R., Forest, M.G., McLaughlin, D.W. and Overman II, E.A. [1986]. A quasi-periodic route to chaos in a near-integrable PDE. *Physica 23D*, **198**, 293-328.

Bishop, A.R., Flesch, R., Forest, M.G., McLaughlin, D.W. and Overman II, E.A. [1989]. Correlations between chaos in a perturbed Sine-Gordon equation and a truncated model system. *Theoretical Division and Center for Nonlinear Studies, Los Alamos National Lab*, preprint.

Coddington, E.A. and Levinson, N. [1955]. *Theory of Ordinary Differential Equations.* McGraw-Hill: New York.

Ercolani, N. and McLaughlin, D.W. [1989]. Notes on Melnikov integrals for models of the driven pendulum chain. *Unpublished Notes.*

Ercolani, N., Forest, M.G. and McLaughlin, D.W. [1989]. Geometry of the modulational instability Part III: Homoclinic orbits for the periodic Sine-Gordon equations. *Physica D*, to appear.

Fenichel, N. [1971]. Persistence and smoothness of invariant manifolds for flows. *Ind. Univ. Math J.*, **21**, 193-225.

Fenichel, N. [1974]. Asymptotic stability with rate conditions. *Ind. Univ. Math J.*, **23**, 1109-1137.

Fenichel, N. [1977]. Asymptotic stability with rate conditions, II. *Ind. Univ. Math J.*, **26**, 81-93.

Fenichel, N. [1979]. Geometric singular perturbation theory for ordinary differential equations. *J. Diff. Eqns.*, **31**, 53-98.

Guckenheimer, J. and Holmes, P.J. [1983]. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* Springer-Verlag: New York, Heidelberg, Berlin.

Holmes, P.J. and Marsden, J.E. [1982a]. Horseshoes in perturbations of Hamiltonian systems with two degrees of freedom. *Comm. Math. Phys.*, **82**, 523-544.

Olver, P.J. [1986]. *Applications of Lie Groups to Differential Equations.* Springer-Verlag: New York, Heidelberg, Berlin.

Šilnikov, L.P. [1965]. A case of the existence of a denumerable set of periodic motions. *Sov. Math. Dokl.*, **6**, 163-166.

Wiggins, S. [1988]. *Global Bifurcations and Chaos.* Springer-Verlag: New York, Heidelberg, Berlin.

Wiggins, S. [1989]. *Nonlinear Dynamical Systems and Chaos.* Springer-Verlag: New York, Heidelberg, Berlin, to appear.

Zufiria J.A. [1988]. Oscillatory Spatially Periodic Weakly Nonlinear Gravity Waves on Deep Water. *J. Fluid Mech.*, **191**, 341-372.