NISTIR 5104

A Mechanism for Capture into Resonance

T. J. Burns
C. K. R. T. Jones

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards and Technology
Gaithersburg, MD 20899

NIST

NISTIR 5104

A Mechanism for Capture into Resonance

T. J. Burns
C. K. R. T. Jones

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Gaithersburg, MD 20899

January 1993

A MECHANISM FOR CAPTURE INTO RESONANCE

T. J. BURNS* AND C. K. R. T. JONES†

Abstract. We present a mechanism for capture into resonance in perturbed two-frequency Hamiltonian systems. When an isolated attractor of the averaged system passes through a resonance on a time scale which is asymptotically slower than that on which the damping works, it transfers its domain of attraction to the resonance.

1. Introduction. The method of averaging is one of the main perturbation techniques used to obtain information about solutions of systems of the form

$$
\begin{aligned}
\dot{\phi} &= \omega(I) + \varepsilon g(I, \phi, \varepsilon), \\
\dot{I} &= \varepsilon f(I, \phi, \varepsilon),
\end{aligned}
\tag{1.1}
$$

where ε is a small positive perturbation parameter, $\phi = (\phi_1, \ldots, \phi_m) \in T^m$ are fast phase variables, $I = (I_1, \ldots, I_n) \in R^n$ are slow variables which are first integrals of the unperturbed problem, $\omega(I) = (\omega_1(I), \ldots, \omega_m(I))$ are the frequencies of the unperturbed motion, and the functions f and g are 2π-periodic in ϕ. Define the *averaged system* for the slow variables to be the system obtained by eliminating the oscillatory part of f to leading order in ε,

$$
\dot{J} = \varepsilon F(J), \qquad F(J) = (2\pi)^{-2} \oint_{T^m} f(J, \phi, 0)\, d\phi. \tag{1.2}
$$

The idea behind the *averaging principle* is that the averaged system should provide a good approximation for the evolution of the slow variables I over a time interval of order $1/\varepsilon$ (see e.g. [3], [4]).

When $m = 1$, so that there is just one fast variable, if the frequency is bounded away from zero, i.e., $\omega(I) > c_1^{-1} > 0$, then the averaging principle holds true in great generality (see e.g. [3], [4]). (In this paper, c_i denote positive constants.) However, the resonant interaction of two or more fast variables can affect the validity of averaging. Indeed, Arnold [2] has given an example in the case $m = 2$ which shows that, without further assumptions, trajectories of the original system (1.1) can be captured in a resonance, with the result that $|I(t) - J(t)|$ can be $O(1)$ after a time of order $1/\varepsilon$. Thus, traditionally capture in resonance has been viewed as an obstacle to the applicability of the averaging principle in systems with multiple frequencies. As a partial remedy to this situation for the case where $m = 2$, Neishtadt has proved theorems giving estimates on the measure of the set of initial conditions that lead to solutions captured in resonance, which must therefore be excluded when applying the method of averaging; see below for more details. However, capture in resonance in perturbed two-frequency systems also

* Computing and Applied Mathematics Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899.

† Division of Applied Mathematics, Brown University, Providence, RI 02912; supported by ONR under grant N00014-92-J-1401 and NSF under grant DMS-910085.

offers an explanation of some of the interesting phenomena of resonant motion that occur in nature. Particularly striking examples occur in celestial mechanics, such as the 3 : 2 ratio of the spin to orbital periods of the planet Mercury in its motion around the Sun (see e.g. [10], [16]). From this viewpoint, capture in resonance is desirable, and results are sought that guarantee such capture to be an event of high probability. In this paper, we will describe a mechanism that can make capture in resonance very likely. The mechanism involves an interaction between the asymptotic structures of the averaged system and a resonance in the case where $m = 2$.

We shall consider a perturbed two-frequency Hamiltonian system of the form (1.1), where $0 < \varepsilon \ll 1$, and the right-hand sides have period 2π in $\boldsymbol{\phi}$ and are analytic on the set $\mathcal{K} = \{(\boldsymbol{I}, \boldsymbol{\phi}) : \boldsymbol{I} \in G \subset \boldsymbol{C}^2,\ |Im(\phi_i)| \le \rho,\ i = 1,2\}$, where $\rho > 0$ and G is a complex, compact domain. We assume that the system (1.1) is real for real values of the arguments of the right-hand sides, and consider only real solutions. A *resonance* is said to occur in the unperturbed system (1.1) whenever $(\boldsymbol{k}, \boldsymbol{\omega}) = 0$ for an irreducible integer coefficient vector $\boldsymbol{k} = (k_1, k_2) \neq 0$. Notice that there is no information about resonances in the averaged system (1.2). The original perturbed system (1.1) with $m = 2$ is said to satisfy Condition A if $\omega_2(\boldsymbol{I}) \neq 0$ and the rate of change of the frequency ratio ω_1/ω_2 along trajectories of the system is bounded away from zero. Arnold [2] has shown that when Condition A is satisfied, the orbits of the system cross the resonance zones transversely, and the averaging principle is valid for (1.1) with $m = 2$.

Since the work of Arnold, there have been a number of studies which give conditions for the validity of the averaging principle, for most initial conditions, in perturbed two-frequency systems of the form (1.1) (for references, see [3], [4], [18]). The strongest result of this type is due to Neishtadt [20], and it can be stated as follows. Assume $\omega_2 \neq 0$. System (1.1) is said to satisfy condition N if the rate of change of the frequency ratio ω_1/ω_2 along trajectories of the *averaged* system (1.2) is bounded away from zero, i.e.,

$$|L(\boldsymbol{J})| > c_2^{-1} > 0, \tag{1.3}$$

where

$$L(\boldsymbol{I}) = \left(\omega_1 \frac{\partial \omega_2}{\partial \boldsymbol{I}} - \omega_2 \frac{\partial \omega_1}{\partial \boldsymbol{I}}\right) \boldsymbol{F} \tag{1.4}$$

(scalar product in $\boldsymbol{R}^n$). Neishtadt has proved that if condition N holds, along with a certain nondegeneracy condition B (which is typically the case), then for most initial points $(\boldsymbol{I}_0, \boldsymbol{\phi}_0)$, if $\boldsymbol{I}(0) = \boldsymbol{J}(0)$, then

$$|\boldsymbol{I}(t) - \boldsymbol{J}(t)| < c_3\sqrt{\varepsilon}\,|\ln \varepsilon|, \qquad 0 \le t \le 1/\varepsilon. \tag{1.5}$$

The exceptional set of initial points of trajectories for which (1.5) does not hold has measure $\le c_4\sqrt{\varepsilon}$; it includes those trajectories which are captured by a resonance. Hence, capture in resonance can be viewed as an event with low probability, and passage through resonance is the typical behavior for a time $\sim 1/\varepsilon$ in perturbed two-frequency Hamiltonian systems which satisfy the conditions of Neishtadt's theorem.

Standing in contrast to Neishtadt's theorem are the many stable resonances which have been observed in the Solar System. In particular, the $3:2$ spin/orbit resonance of the planet Mercury, which is known to an accuracy of 10^{-4} [16], appears to be a stable resonance in a dissipatively perturbed two-frequency integrable system that is deemed unlikely by Neishtadt's result. However, it turns out that, in some models of this spin-orbit resonance [10], Neishtadt's condition N is violated repeatedly [17], [6]. In this paper, we will formulate a simple model problem which gives a clear focus on this phenomenon. An application to the motion of Mercury will be given in a forthcoming paper. We shall consider a perturbed two-frequency Hamiltonian system with effectively a single resonance, which has the property that, in a compact neighborhood of this resonance, Neishtadt's transversality condition holds for a time $\sim 1/\varepsilon^2$, then is violated, then holds again for a time $\sim 1/\varepsilon^2$, and so on, as a certain slow variable (related to very slow change in the eccentricity of Mercury's orbit) varies between zero and one (mod 1). We will show this provides an interesting *transfer of capture* mechanism for the transport of phase space (see Wiggins [23]). As far as we know, this transfer of capture idea was first proposed by Kyner [17] in a study of the spin-orbit resonance of Mercury.

Our example is as follows,

$$\begin{aligned} \dot{\phi}_1 &= I_1, \\ \dot{\phi}_2 &= 1, \\ \dot{I}_1 &= \varepsilon\left[\sin(\phi_2-\phi_1)+\tfrac{1}{2}\cos 2\pi I_2-\tfrac{1}{4}I_1\right], \\ \dot{I}_2 &= \varepsilon^2, \end{aligned} \tag{1.6}$$

for $(\boldsymbol{I},\boldsymbol{\phi})\in D\times \boldsymbol{T}^2$, where $D=\boldsymbol{R}\times\boldsymbol{S}^1$ and $\boldsymbol{S}^1=\boldsymbol{R}(\text{mod}\,1)$. There is effectively only one resonant phase, $\theta=\phi_1-\phi_2$, corresponding to the resonance curve $\eta=I_1-1=0$. Thus the system is equivalent to

$$\begin{aligned} \dot{\theta} &= \eta, \\ \dot{\psi} &= 1, \\ \dot{\eta} &= \varepsilon\left[B(\tau)-\sin\theta-\tfrac{1}{4}\eta\right], \\ \dot{\tau} &= \varepsilon^2, \end{aligned} \tag{1.7}$$

where

$$B(\tau)=\tfrac{1}{2}\cos 2\pi\tau-\tfrac{1}{4}, \tag{1.8}$$

$\psi=\phi_2$, $\tau=I_2$, and $(\eta,\tau,\theta,\psi)\in\boldsymbol{R}\times\boldsymbol{S}^1\times\boldsymbol{T}^2$. System (1.7) is a trivial example of a *partially averaged system*, in which the motion near a particular resonance is studied by retaining in the perturbed system only the harmonic terms specific to the resonance of interest (see [3], [4]). The averaged system for the slow variables in system (1.7) is given by

$$\begin{aligned} \dot{x} &= \tfrac{1}{4}\varepsilon\left[x_*(p)-x\right], \\ \dot{p} &= 0, \end{aligned} \tag{1.9}$$

where x_*, the *zero of the averaged torque*, is defined by

$$x_*(p) = 4B(p). \tag{1.10}$$

Condition N is violated at the resonance if $B(p) = 0$, which occurs at $p = \frac{1}{6}$ and $p = \frac{5}{6}$. Thus, $B(p) = 0$ is equivalent to the zero of the averaged torque coinciding with the resonance $\{\eta = 0\}$.

Henceforth, we shall ignore the equation for ψ in (1.7), and consider the resulting $1\frac{1}{2}$ degree of freedom system for $(\eta, \tau, \theta) \in \boldsymbol{R} \times \boldsymbol{S}^1 \times \boldsymbol{T}^1$,

$$\begin{aligned} \dot{\theta} &= \eta, \\ \dot{\eta} &= \varepsilon \left[B(\tau) - \sin\theta - \tfrac{1}{4}\eta \right], \\ \dot{\tau} &= \varepsilon^2. \end{aligned} \tag{1.11}$$

If we rescale the system by introducing the slower time $s = \sqrt{\varepsilon}t$ and the normalized distance to the resonance $z = \eta/\sqrt{\varepsilon}$, we obtain

$$\begin{aligned} \theta' &= z, \\ z' &= B(\tau) - \sin\theta - \tfrac{1}{4}\sqrt{\varepsilon}\, z, \\ \tau' &= \varepsilon^{3/2}, \end{aligned} \tag{1.12}$$

where $' = d/ds$. This system is in the "pendulum" normal form obtained by partial averaging in the neighborhood of a given resonance in more general systems of the form (1.1), which can have infinitely many resonances (see e.g. [4], [18]). Thus, the model problem we shall consider is equivalent to the equation of a pendulum with friction and very slowly varying periodic external torque, which changes sign twice during each period. Neishtadt's condition B for system (1.6) requires that the saddle equilibrium point of system (1.12) be nondegenerate. Since $|B(\tau)| \leq \frac{3}{4}$, condition B is always satisfied by (1.6).

We shall show that for the above model problem, permanent capture into resonance is quite likely. We formulate this phenomenon in two Theorems. The first concerns the existence of an attractor near the resonance curve, while the second addresses its domain of attraction and hence the likelihood of capture into the resonance through tending asymptotically to the resonant attractor.

THEOREM 1.1. *For sufficiently small ε, system (1.12) has a unique attracting periodic orbit $\mathcal{V}^\varepsilon$ which is located within an $O(\sqrt{\varepsilon})$-neighborhood of the resonance curve $z = 0$.*

In order to study the domain of attraction of $\mathcal{V}^\varepsilon$, we shall use certain sets of initial conditions that can be viewed as sample test sets. However, they are formulated to cover a neighborhood of the zero of the averaged torque, which we now define. We assume that Neishtadt's condition N holds initially in a set containing the resonance, so that the zero of the averaged torque satisfies $|x_*(\tau(0))| > 2d_0 > 0$, for some constant d_0. Without loss of generality, we may assume that $x_*(\tau(0)) > 0$. Then Neishtadt's

Theorem implies there exist positive numbers $\varepsilon_2(d_0)$ and $T_1(d_0)$ such that, for all $\varepsilon \in (0, \varepsilon_2]$, most solutions of system (1.11) with initial value $\eta(0)$ in some bounded interval containing $x_*(\tau(0))$ satisfy

$$\eta(T_1/\varepsilon) \in \{\eta \in \boldsymbol{R} : 0 < d_0 < x_*(\tau(0)) - d_0 \leq \eta \leq x_*(\tau(0)) + d_0\},$$

so that the corresponding solution of the scaled system (1.12) satisfies

$$[x_*(\tau(0)) - d_0]/\sqrt{\varepsilon} \leq z(T_1/\sqrt{\varepsilon}) \leq [x_*(\tau(0)) + d_0]/\sqrt{\varepsilon}. \tag{1.13}$$

The goal of the remainder of this paper is to study the fate of a set of initial points under the flow of system (1.12) which satisfy condition (1.13).

In order to simplify the notation, we assume that the initial values lie within a subset of $\{\tau = 0\}$ at scaled time $s = 0$, namely the "strip"

$$S_\varepsilon = \{(\theta, z, \tau) : \theta \in \boldsymbol{S}^1, \ \ [x_*(0) - d_0]/\sqrt{\varepsilon} \leq z \leq [x_*(0) + d_0]/\sqrt{\varepsilon}, \ \ \tau = 0\}. \tag{1.14}$$

We formulate the Capture Theorem in terms of this "test" set. We use the notation $\omega(\cdot)$ for the ω—limit set of a point.

THEOREM 1.2. *For system (1.12), there is a subset $\hat{S}_\varepsilon \subset S_\varepsilon$ for which $x \in \hat{S}_\varepsilon$ implies that $\omega(x) \subset \mathcal{V}^e$ and meas $(S_\varepsilon - \hat{S}_\varepsilon) = O(\exp(-c/\varepsilon))$, where c is a positive constant which is independent of ε.*

If the zero of the averaged torque did not pass through the resonance, then condition N would be satisfied in a set containing the resonance, and Neishtadt's Theorem would entail that most initial data enter a neighborhood of the zero of the averaged torque, which is an attractor in the averaged system (1.9). We offer the interpretation to Theorem 1.2 that if the zero of the averaged torque passes through the resonance *sufficiently slowly*, then it deposits its domain of attraction into the resonance. A key point is that the passage should be slow enough, which is encoded through the assumption that $\dot{\tau}$ is $O(\varepsilon^2)$. It is an open question how far this assumption can be relaxed. In the next section, we will present a more detailed discussion of the contents of this paper.

2. Outline of Results. In this section, we shall first review some results which are known for the unperturbed pendulum system obtained by setting $\varepsilon = 0$ in the scaled system (1.12), and then we shall review some results which are known for the "frozen" pendulum system obtained by setting $\tau' = 0$ in system (1.12). This will provide a framework which we will then use to outline the contents of the remaining sections of this paper.

When $\varepsilon = 0$, system (1.12) is Hamiltonian, τ is a fixed parameter, and there is a closed invariant curve of saddle-point equilibrium solutions,

$$\mathcal{U} = \{(\theta, z, \tau) \in \mathcal{S} : \theta = \pi - \Theta(\tau), \ z = 0\}, \tag{2.1}$$

and a closed invariant curve of center equilibrium solutions,

$$\mathcal{V} = \{(\theta, z, \tau) \in \mathcal{S} : \theta = \Theta(\tau),\ z = 0\}, \tag{2.2}$$

in the solid torus

$$\mathcal{S} = \mathcal{C} \times \boldsymbol{S}^1, \tag{2.3}$$

where $\mathcal{C}$ is the cylinder $\boldsymbol{S}^1 \times \boldsymbol{R}$, and

$$\Theta(\tau) = \arcsin B(\tau), \qquad |\Theta(\tau)| < \pi/2. \tag{2.4}$$

For $\varepsilon = 0$, the qualitative phase portraits of orbits of (1.12) on the cylinder $\mathcal{C}_\tau = \mathcal{S} \cap \{\tau = \text{const}\}$ are given in Fig. 2.1 for the two cases (a) $0 < B < 1$, and (b) $B = 0$. In each

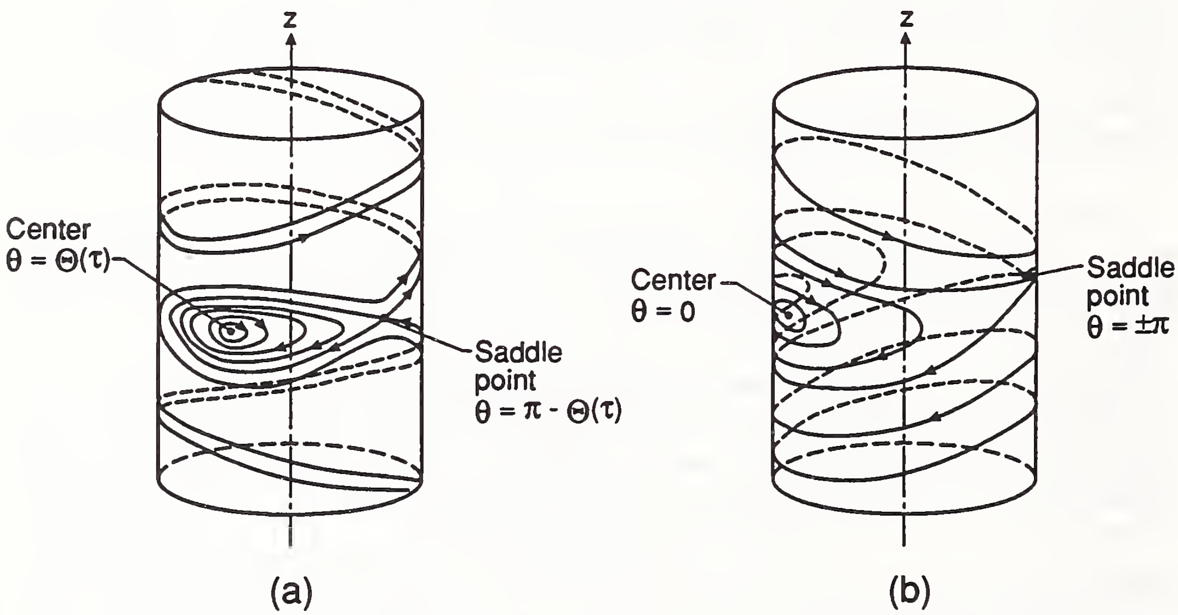

FIG. 2.1. *The undamped pendulum with constant torque.*

case, there are two equilibrium points on the cylinder, a saddle point and a center. In the first case, there is a homoclinic orbit which contains the center equilibrium point in its interior and approaches the saddle point as $t \to \pm\infty$. The solutions inside the region enclosed by the saddle point and the homoclinic orbit correspond to periodic oscillations of the pendulum, and this oscillatory region grows in area as B decreases toward zero. When $B = 0$, instead of a single homoclinic orbit containing the center equilibrium point, there is a pair of homoclinic orbits which encircle the cylinder, one in $\mathcal{C}_\tau \cap \{z > 0\}$ and the other in $\mathcal{C}_\tau \cap \{z < 0\}$, connecting the stable and unstable manifolds of the saddle point, and together these enclose the region of oscillation. Outside this region, the orbits encircle the cylinder, corresponding to complete rotations of the pendulum, with a counterclockwise direction of rotation in the upper region, a clockwise direction of rotation in the lower region, and the angular speed z is a 2π-periodic function of θ in both cases. The two types of periodic motion of the pendulum in this case, corresponding to oscillations and rotations, are called periodic solutions of the first and second kind, respectively. For $B < 0$, the transformation $\theta \to -\theta$, $z \to -z$ reduces the problem to the case with $B > 0$.

When $0 < \varepsilon \ll 1$, the system (1.12) is dissipative, but τ changes only by an amount $\sim \varepsilon$ over an interval in $s \sim 1/\sqrt{\varepsilon}$. Thus, assume for the moment that $0 < \varepsilon \ll 1$, but that τ is still a fixed parameter. Then the "frozen" dissipative system

$$\begin{aligned} \theta' &= z, \\ z' &= B(\tau) - \sin\theta - \tfrac{1}{4}\sqrt{\varepsilon}\, z, \\ \tau' &= 0, \end{aligned} \tag{2.5}$$

is equivalent to the equation of a pendulum with viscous damping and constant torque, and it is known (see e.g. [1], [22]) that, if the ratio of the torque to the damping coefficient, which is the scaled zero of the averaged torque $x_*(\tau)/\sqrt{\varepsilon} = B(\tau)/(\frac{1}{4}\sqrt{\varepsilon})$, is greater than a certain critical value ($\approx 4/\pi$; see [22]), then its qualitative phase portrait on the cylinder $\mathcal{C}_\tau = \mathcal{S} \cap \{\tau = \text{const}\}$ is as indicated in Fig. 2.2(a). There are two

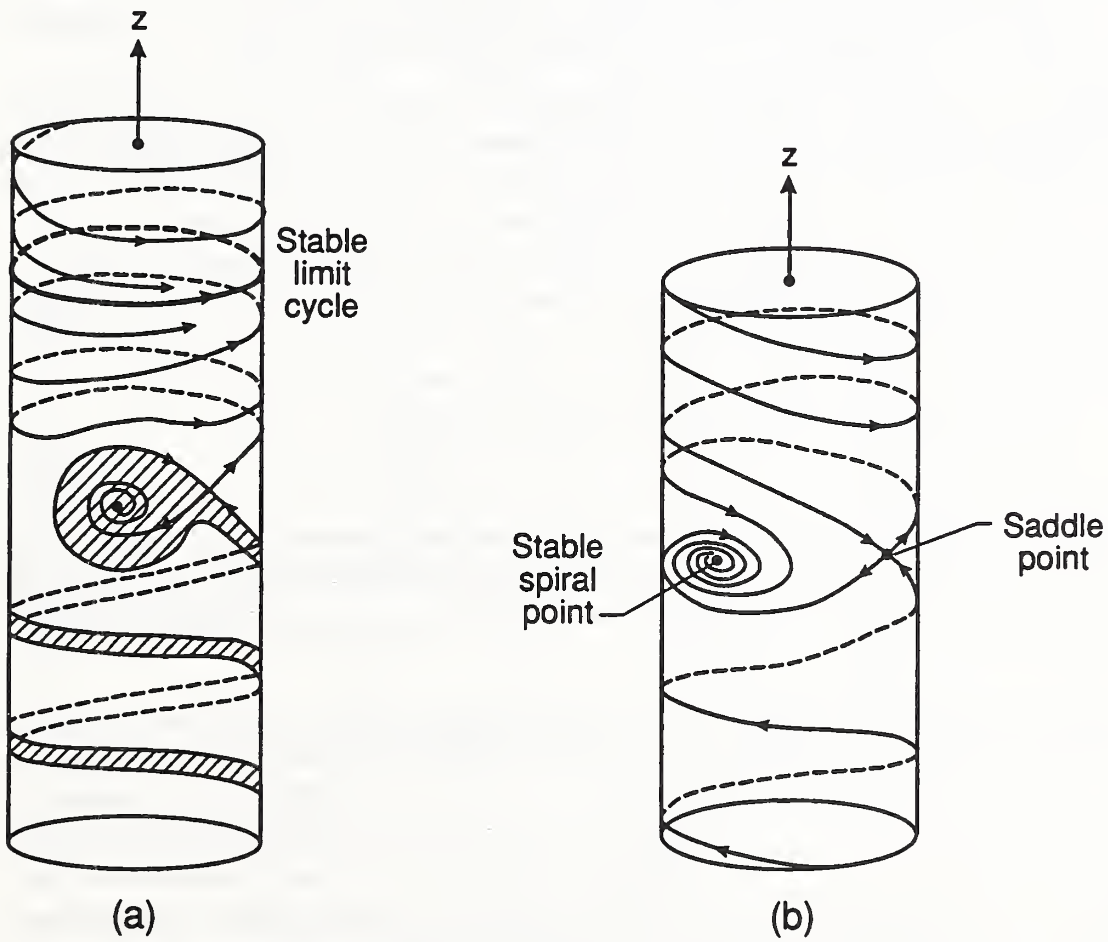

FIG. 2.2. *The damped pendulum with constant torque.*

attractors, a stable spiral point at the resonance $\{z = 0\}$, and a rotary solution which can be shown to be near the zero of the averaged torque. Since the orbit of the attracting rotary solution is a closed curve encircling the cylinder, it is called a limit cycle of the second kind, to distinguish it from a limit cycle of the first kind, which encircles an equilibrium point on the phase cylinder without encircling the cylinder itself. As the parameter B decreases, the limit cycle moves smoothly down the corresponding cylinder, until the critical value of the parameter ratio is reached at $\tau = \tau_A$, when the orbit merges with the stable and unstable manifolds of the saddle point to become a homoclinic saddle connection. For smaller nonnegative values of B, the saddle connection breaks, and every orbit except the saddle point and its stable manifold approaches the spiral point asymptotically with increasing time, as indicated in Fig. 2.2(b). Once again, for $B \leq 0$, the transformation $\theta \to -\theta$, $z \to -z$ reduces the problem to the case with $B \geq 0$. Thus, as B decreases and becomes negative, the spiral point remains the only attractor in the system, until the negative of the critical parameter value is reached when $\tau = \tau_B$, and a saddle connection appears below the $z = 0$ circle on the cylinder, then it breaks, and the limit cycle reappears and moves down the cylinder as B decreases further. After B reaches its minimum value at $\tau = \frac{1}{2}$, the limit cycle moves back up the cylinder as τ increases, disappears when the lower saddle connection forms when $\tau = \tau_C$, reappears when the upper saddle connection forms when $\tau = \tau_D$, and so on periodically, as the parameter τ varies between zero and one (mod 1).

Thus, in the "frozen" dissipative system, the invariant closed curve (2.2) of center equilibrium points $\mathcal{V}$ becomes a normally hyperbolic invariant closed curve of stable spiral equilibrium points. Also, there exist numbers $\tau_A, \tau_B, \tau_C, \tau_D \in \boldsymbol{S}^1$, with $0 < \tau_A < \tau_B < \tau_C < \tau_D < 1$, such that, for all τ in the arc $(\tau_D, \tau_A) \subset \boldsymbol{S}^1$ containing 0, there is an attracting invariant surface in $\mathcal{S} \cap \{z > 0\}$, and for all τ in the arc $(\tau_B, \tau_C) \subset \boldsymbol{S}^1$ containing $\frac{1}{2}$, there is an attracting invariant surface in $\mathcal{S} \cap \{z < 0\}$. These surfaces are composed of asymptotically stable rotary solutions which correspond to zeros of the averaged torque. At each of the endpoints of these arcs, the surface develops a cusp at the saddle equilibrium point, and for τ in the closed set $[\tau_A, \tau_B] \cup [\tau_C, \tau_D]$, no such surface exists, so the only attractor in $\mathcal{S} \cap ([\tau_A, \tau_B] \cup [\tau_C, \tau_D])$ is the curve of stable spiral points. This leads to the conjecture that, in the "unfrozen" system (1.12) with $\tau' = \varepsilon^{3/2} > 0$, there is an attractor $\mathcal{V}^\varepsilon$ associated with the resonance, and there is also an attractor associated with the zero of the averaged torque for as long as it is located away from the resonance. It also suggests that the following *transfer of capture* scenario is possible in the "unfrozen" system.

Suppose that initially the zero of the averaged torque is located outside the resonance region near $\eta = 0$ in system (1.12). Without loss of generality, we may assume that it is located above the resonance in $\mathcal{S} \cap \{z > 0\}$. If a solution of the system is initially located above the resonance in $\mathcal{S} \cap \{z > 0\}$, then it will be attracted to the zero of the averaged torque. If a solution of the system is initially located below the resonance in $\mathcal{S} \cap \{z < 0\}$, then by Neishtadt's theory, as time increases this solution is likely to pass through the resonance, and thus also will be attracted to the zero of the averaged torque. However, as time continues to increase, the zero of the averaged torque will move downward and cross the region of the resonance, and while this happens, the domain of attraction of the zero of the averaged torque could be transferred to the domain of attraction of the resonance attractor. Of course, this heuristic argument does not take into account the interaction between the $O(\sqrt{\varepsilon})$ damping and the effect of the $O(\varepsilon^{3/2})$ rate of change of the parameter τ with respect to the scaled time s. Nevertheless, in this paper, we will prove that this transfer of capture is indeed what happens. In particular, the results we will present in the following sections will establish that very nearly all solutions of system (1.12) are captured by $\mathcal{V}^\varepsilon$ as $s \to \infty$.

In Section 3, we will prove there exist numbers $\tau_\#$, $\tau^\#$, with $\tau_\# < \tau^\#$, and an overflowing invariant set in $\mathcal{S} \cap \{z > 0, \ \ \tau_\# \leq \tau \leq \tau^\#\}$ which contains the surface $\{(\theta, z, \tau) \in \mathcal{S} : x = x_*(\tau), \ \tau_\# \leq \tau \leq \tau^\#\}$. Furthermore, we will show that, under the flow of system (1.12), any solution with initial value in the strip S_ε enters this invariant set before $\tau(s) = \tau^\#$. In Section 4, we will prove there also exist homoclinic orbits in the time-dependent system. In Section 5, we will prove that the curve of equilibrium points $\mathcal{V}$ perturbs to a periodic attractor $\mathcal{V}^\varepsilon$ which is $O(\sqrt{\varepsilon})$-close to $\mathcal{V}$. Finally, in Section 6, we will prove our main result about almost certain capture in resonance, Theorem 1.2.

3. The Finite-Time Rotary Attractor.

We shall henceforth unfold the phase space $\mathcal{S}$ to form the phase space $\boldsymbol{R}^2 \times \boldsymbol{S}^1$ wherein $(\theta, z) \in \boldsymbol{R}^2$ and $\tau \in \boldsymbol{S}^1$. In this section, we will define the component of the finite-time rotary attractor which lies in $\{z > 0\}$ and establish some of its properties, on an appropriate interval in τ, which will be used in the proof of Theorem 1.2 in Section 6. It will turn out that the intersection of the attractor with each strip $\{\tau = \text{const}\}$ is the limit cycle of the second kind of the "frozen" system (2.5) for the given value of the parameter τ. A completely analogous discussion applies to the component which lies in $\{z < 0\}$.

Continuing the outline in the preceding section of known results for the scaled "frozen" system (2.5) (see [22]), the associated equations for the trajectories of system (2.5) in $\boldsymbol{R}^2 \times \boldsymbol{S}^1$ are given by

$$\frac{dz}{d\theta} = \frac{B(\tau) - \sin\theta}{z} - \tfrac{1}{4}\sqrt{\varepsilon}, \tag{3.1}$$

$$\frac{d\tau}{d\theta} = 0, \tag{3.2}$$

for $z \neq 0$. Let A_τ denote the unique limit cycle of the second kind of the "frozen" system (2.5),

$$A_\tau = \{(\theta, z) : \ \theta = \psi(s, \tau), \ z = w(s, \tau), \ \ s \in \boldsymbol{R}\}, \tag{3.3}$$

for each $\tau \in (\tau_D, \tau_A)$. As a function of s for fixed τ, A_τ has minimum period Λ_τ in s given by

$$\Lambda_\tau = \int_0^{2\pi} \frac{d\theta}{z_p(\theta, \tau)}, \tag{3.4}$$

where $z_p(\theta, \tau)$ is the unique 2π-periodic solution of Eq. (3.1), so that

$$w(s, \tau) = z_p(\psi(s, \tau), \tau) > 0. \tag{3.5}$$

It can also be shown that

$$\psi(s,\tau) = \chi(s,\tau) + 2\pi s/\Lambda_\tau\,, \tag{3.6}$$

where $\chi(s,\tau)$ is a function of period Λ_τ in s, with $\chi(0,\tau) = 0$, for each fixed $\tau \in (\tau_D,\, \tau_A)$. Thus, for fixed $\tau \in (\tau_D,\, \tau_A)$, the limit cycle of the second kind is an attracting 2π-periodic function of θ, $A_\tau = \{(\theta, z):\ z = z_p(\theta,\tau),\ \theta \in \mathbf{R}\}$. Also, for each $\tau \in (\tau_D,\, \tau_A)$, it can be shown that the mean of z_p with respect to θ satisfies

$$M_\theta[z_p] = \frac{1}{2\pi}\int_0^{2\pi} z_p(\theta,\tau)\, d\theta = B(\tau)/(\tfrac{1}{4}\sqrt{\varepsilon}) = x_*(\tau)/\sqrt{\varepsilon}\,. \tag{3.7}$$

Thus, for a given fixed τ, the mean $M_\theta[z_p]$ is equal to the zero of the averaged torque (1.10) scaled by $1/\sqrt{\varepsilon}$. If $\tau_a,\ \tau_b \in [0,\, \tau_A)$ satisfy $\tau_b > \tau_a$, then by monotonicity of B on this interval, $B(0) \geq B(\tau_a) > B(\tau_b) > 0$, and it can be shown in this case that $z_p(\theta,\tau_a) > z_p(\theta,\tau_b) > 0$ for all $\theta \in [0, 2\pi]$. Furthermore, $\min\{z_p(\theta,\tau) : \theta \in [0,2\pi]\} \downarrow 0$ as $\tau \uparrow \tau_A$. Similarly, it can be shown that if $\tau_a,\ \tau_b \in (\tau_D,\, 0]$ satisfy $\tau_b > \tau_a$, then by monotonicity of B on this interval, $B(0) \geq B(\tau_b) > B(\tau_a) > 0$, and it can be shown that $z_p(\theta,\tau_b) > z_p(\theta,\tau_a) > 0$ for all $\theta \in [0,2\pi]$, and also $\min\{z_p(\theta,\tau) : \theta \in [0,2\pi]\} \downarrow 0$ as $\tau \downarrow \tau_D$. Thus, for any small positive number $d_1 < d_0$, independent of ε, there exist positive numbers $\tau_\dagger < \tau_\# < \tau^\# < \tau^\dagger \in (\tau_D,\, \tau_A)$, such that $z_p(\theta,\tau) \geq d_1$, for all $(\theta,\tau) \in [0,2\pi] \times [\tau_\#,\, \tau^\#]$, and $z_p(\theta,\tau) \geq \frac{1}{2}d_1$, for all $(\theta,\tau) \in [0,2\pi] \times [\tau_\dagger,\, \tau^\dagger]$.

Definition. For a given value of d_1, the *rotary attractor* is the smooth surface

$$\mathcal{A} = \cup_{\tau\in[\tau_\#,\tau^\#]}\, A_\tau. \tag{3.8}$$

Note that $\mathcal{A}$ is not invariant in the full system (1.12). We nevertheless refer to this closed set as an attractor because, as we shall now show, it possesses an attracting neighborhood $\mathcal{P}$, relative to an appropriate τ-interval.

Let γ be the number such that $(1-\gamma)B(\tau_\#) = B(\tau_\dagger)$ and $(1-\gamma)B(\tau^\#) = B(\tau^\dagger)$, so that $0 < \gamma < 1$. For each $\tau \in [\tau_\#, \tau^\#]$, let $B^0_\tau = B(\tau)$, $B^+_\tau = (1+\gamma)B^0_\tau$, and $B^-_\tau = (1-\gamma)B^0_\tau$. Also, let $z^0_\tau(\theta)$, $z^+_\tau(\theta)$, and $z^-_\tau(\theta)$ denote the trajectories of the rotary limit cycles A_τ, A^+_τ, and A^-_τ of the "frozen" system (2.5) with $B(\tau)$ replaced by B^0_τ, B^+_τ, and B^-_τ, respectively. The next Lemma establishes that, on the appropriate τ-interval, the limit cycles remain bounded away from the resonance, and A_τ lies below A^+_τ and above A^-_τ. Recall that $B(0) = \frac{1}{4}$.

Lemma 3.1. *For each $\tau \in [\tau_\#, \tau^\#]$ and for all $\theta \in \mathbf{R}$,*

$$0 < \tfrac{1}{2}d_1 \leq z^-_\tau(\theta) < z^0_\tau(\theta) < z^+_\tau(\theta) \leq 6/\sqrt{\varepsilon}\,. \tag{3.9}$$

Proof. Let

$$G^0_\tau(\theta,z) = (B^0_\tau - \sin\theta - \tfrac{1}{4}\sqrt{\varepsilon}z)/z, \qquad G^+_\tau(\theta,z) = (B^+_\tau - \sin\theta - \tfrac{1}{4}\sqrt{\varepsilon}z)/z.$$

Then for all $\theta \in \boldsymbol{R}$ and $z > 0$, $G^0_\tau(\theta, z) < G^+_\tau(\theta, z)$. Furthermore, since the means (3.7) of the two limit cycle trajectories z^0_τ and z^+_τ satisfy $M_\theta[z^0_\tau] < M_\theta[z^+_\tau]$, it follows there must exist a number $\bar{\theta} \in [0, 2\pi]$ such that $z^0_\tau(\bar{\theta}) < z^+_\tau(\bar{\theta})$. By a standard comparison theorem (see e.g. [5, Corollary 2, p. 26]) and 2π-periodicity of both functions in θ, it follows that $z^0_\tau(\theta) < z^+_\tau(\theta)$ for all $\theta \in \boldsymbol{R}$. Similarly, it can be shown that $z^-_\tau(\theta) < z^0_\tau(\theta)$ for all $\theta \in \boldsymbol{R}$. The lower bound follows from the discussion preceding the definition of $\mathcal{A}$. The upper bound is established as follows. The curve z^+_τ is at its highest point when $\tau = 0$ and $dz^+_\tau/d\theta = 0$. By (3.1), this implies that $z^+_\tau(\theta) \leq [(1+\gamma)B(0)+1]/(\frac{1}{4}\sqrt{\varepsilon}) \leq 6/\sqrt{\varepsilon}$. □

Let

$$\mathcal{P} = \bigcup_{\tau \in [\tau_\#, \tau^\#]} \{(\theta, z, \tau) : \ \theta \in \boldsymbol{R}, \ \ z^-_\tau(\theta) < z < z^+_\tau(\theta)\}. \tag{3.10}$$

LEMMA 3.2. *For Eq. (1.12), $\mathcal{P}$ is positively invariant relative to the set $\boldsymbol{R}^2 \times [\tau_\#, \tau^\#]$.*

Proof. For each $\tau \in [\tau_\#, \tau^\#)$, the vector field of (1.12) is always directed into the set $\mathcal{P}$. To see this, we compare the first two components of the vector field of system (1.12) with the corresponding vector fields which define the bounding curves A^+_τ and A^-_τ. Along the upper bounding curve, we have for $\tau \in [\tau_\#, \tau^\#]$,

$$(z, B^+_\tau - \sin\theta - \tfrac{1}{4}\sqrt{\varepsilon}z) \wedge (z, B^0_\tau - \sin\theta - \tfrac{1}{4}\sqrt{\varepsilon}z) = \left(B^0_\tau - B^+_\tau\right) z = -\gamma B^0_\tau < 0,$$

since $B^0_\tau > 0$ on $[\tau_\#, \tau^\#]$. Similarly, along the lower bounding curve, we have for $\tau \in [\tau_\#, \tau^\#]$,

$$(z, B^-_\tau - \sin\theta - \tfrac{1}{4}\sqrt{\varepsilon}z) \wedge (z, B^0_\tau - \sin\theta - \tfrac{1}{4}\sqrt{\varepsilon}z) = \left(B^0_\tau - B^-_\tau\right) z = \gamma B^0_\tau > 0.$$

It is also clear that trajectories of system (1.12) must exit $\mathcal{P}$ along the face $\mathcal{P} \cap \{\tau = \tau^\#\}$. □

The next Lemma establishes that S_ε enters $\mathcal{P}$ under the flow of the full system (1.12). Let $s^\# = \tau^\#/\varepsilon^{3/2}$.

LEMMA 3.3. $S_\varepsilon \cdot s^\# \subset \mathcal{P} \cap \{\tau = \tau^\#\}$.

Proof. The points in $\sqrt{\varepsilon}S_\varepsilon = \{(\theta, \eta, \tau) : \eta = \sqrt{\varepsilon}z, \text{ where } (\theta, z, \tau) \in S_\varepsilon\}$ (see (1.14)) are bounded away from the resonance at $\eta = 0$ in system (1.11). Thus, we may use the method of averaging (see e.g. [3]) to assert there exists a near-identity change of variables $w = \eta + \varepsilon h(\eta, \tau, \theta)$ such that $(\eta, \tau, \theta) \mapsto (w, \tau, \theta)$ is a diffeomorphism on $\sqrt{\varepsilon}S_\varepsilon \times [0, \tau^\#] \times \boldsymbol{R}$ which transforms the equations for the slow variables in system (1.11) into the system

$$\dot{w} = -\tfrac{1}{4}\varepsilon[w - x_*(\tau)] + O(\varepsilon^2), \tag{3.11}$$

$$\dot{\tau} = \varepsilon^2. \tag{3.12}$$

Letting $w = v + x_*(\tau)$, we then have that

$$\dot{v} = -\tfrac{1}{4}\varepsilon v + O(\varepsilon^2). \tag{3.13}$$

Hence, for $0 \le t \le \tau^{\#}/\varepsilon^2$,

$$\eta(t) = x_*(\tau) + \alpha \exp\left(-\tfrac{1}{4}\varepsilon t\right) + O(\varepsilon),$$

where α is a constant which depends on the initial data. Thus, for all solutions of system (1.11) with initial value for η in $\sqrt{\varepsilon}S_\varepsilon$, there exists a positive constant T_2 such that, for sufficiently small ε, $\tau(T_2/[\varepsilon/|\ln\varepsilon|]) = O(\varepsilon|\ln\varepsilon|) < \tau^{\#}$ and $v(T_2/[\varepsilon/|\ln\varepsilon|]) = O(\varepsilon)$. Hence, $\eta(T_2/[\varepsilon/|\ln\varepsilon|]) - x_*(\tau(T_2/[\varepsilon/|\ln\varepsilon|])) = O(\varepsilon)$, which in turn implies that, in the scaled system (1.12), $z(T_2/[\sqrt{\varepsilon}/|\ln\varepsilon|]) - x_*(\tau(T_2/[\sqrt{\varepsilon}/|\ln\varepsilon|]))/\sqrt{\varepsilon} = O(\sqrt{\varepsilon})$. Using a comparison argument similar to that used in the proof of Lemma 3.1, it follows there is an s^* such that $\tau(s^*) < \tau^{\#}$ and every orbit of system (1.12) with initial value for $z \in S_\varepsilon$ is contained in $\mathcal{P} \cap \{\tau = \tau(s^*)\}$. The conclusion of the Lemma then follows from the positive invariance of $\mathcal{P}$. □

Let $\tau_* = \frac{1}{6}$, so that $B(\tau_*) = 0$, and let $\hat{\tau}$ be a number such that $B'(\tau) < 0$ on $(0, \hat{\tau}]$ and such that $\hat{\tau} > \tau_*$. Consider $A^+_{\tau^{\#}}$, the upper bounding curve of $\mathcal{P}_{\tau^{\#}} = \mathcal{P} \cap \{\tau = \tau^{\#}\}$. The set $A^+_{\tau^{\#}} \times [\tau^{\#}, \hat{\tau}]$ forms a barrier, preventing $S_\varepsilon \cdot s$ from drifting up, since for $\tau = \tau^{\#}$, the vector field points down on $A^+_{\tau^{\#}}$, and this effect is only strengthened as $B(\tau)$ decreases. Thus, we have the following Lemma, wherein $\hat{\mathcal{P}}_{\tau^{\#}}$ is the set of all points below $A^+_{\tau^{\#}} \times [\tau^{\#}, \hat{\tau}]$. Recall that, at the end of Section 1, we have assumed that $\tau = 0$ when $s = 0$.

LEMMA 3.4. *For all* $\tau \in [\tau^{\#}, \hat{\tau}]$, $S_\varepsilon \cdot s \subset \hat{\mathcal{P}}_{\tau^{\#}}$.

4. Heteroclinic Orbits. In the phase space $\boldsymbol{R}^2 \times \boldsymbol{S}^1$, wherein $(\theta, z) \in \boldsymbol{R}^2$ and $\tau \in \boldsymbol{S}^1$, there are infinitely many curves of saddle points when $\varepsilon = 0$. Indeed, setting $\varepsilon = 0$ in (1.12) renders

$$\begin{aligned} \theta' &= z\,, \\ z' &= B(\tau) - \sin\theta\,, \\ \tau' &= 0\,, \end{aligned} \tag{4.1}$$

which is the equation of a forced pendulum with the forcing B depending on the parameter τ. In each slice with τ fixed, (4.1) has infinitely many critical points at

$$s_k(\tau) = (\Theta(\tau) + 2k\pi,\ 0) \tag{4.2}$$

and

$$u_k(\tau) = (-[\Theta(\tau) + \pi] + 2k\pi,\ 0), \tag{4.3}$$

where $k \in \boldsymbol{Z}$ and $\Theta(\tau) = \arcsin B(\tau)$, as in (2.4). The $s_k(\tau)$ are the centers discussed in Section 2, while the points $u_k(\tau)$ are saddle points in each τ-slice.

Since B is periodic, each set

$$\mathcal{U}_k = \{u_k(\tau) : \tau \in [0,1]\}$$

is a simple closed curve in the phase space $\boldsymbol{R}^2 \times \boldsymbol{S}^1$. Moreover, it is normally hyperbolic relative to the flow induced by (4.1). By Fenichel's Theorems on invariant manifolds (see [8], [9], [14]), it must therefore perturb to an invariant manifold for (1.12) when $\varepsilon > 0$ but sufficiently small. Moreover, this invariant manifold is also a simple closed curve. Since $\tau' > 0$ when $\varepsilon > 0$, this manifold must be a periodic orbit, which we denote $\mathcal{U}_k^\varepsilon$. Fenichel's theorems also guarantee that the local stable and unstable manifolds for $\mathcal{U}_k$ perturb to the same for $\mathcal{U}_k^\varepsilon$.

We next address the question of whether the periodic orbits $\mathcal{U}_k^\varepsilon$ have any heteroclinic orbits between them. When $\varepsilon = 0$, the condition $B(\tau) = 0$ (which gives an unforced pendulum in (4.1)) exactly guarantees the existence of two heteroclinic orbits between $\mathcal{U}_k$ and $\mathcal{U}_{k+1}$, one in $z > 0$ ($\mathcal{U}_k \to \mathcal{U}_{k+1}$) and the other in $z < 0$ ($\mathcal{U}_{k+1} \to \mathcal{U}_k$). In particular, this structure occurs at $\tau = \tau_*$, where $\tau_* = \frac{1}{6}$, so that $B(\tau_*) = 0$ and $B'(\tau_*) < 0$. We shall see that both of these heteroclinic orbits perturb to a heteroclinic orbit when $\varepsilon > 0$, between the relevant periodic orbit. However, the structure holds on a small interval of τ's, and to uncover it, we must blow τ up near $\tau = \tau_*$. To this end, set

$$\rho = \frac{\tau - \tau_*}{\sqrt{\varepsilon}}, \tag{4.4}$$

where $\rho \in \mathcal{I}$, some closed, bounded interval containing 0.

Now,

$$B(\tau) = B(\tau_* + \sqrt{\varepsilon}\rho) = \sqrt{\varepsilon}\beta(\rho, \varepsilon).$$

We rewrite (1.12) as

$$\begin{aligned} \theta' &= z\,, \\ z' &= -\sin\theta + \sqrt{\varepsilon}\beta(\rho,\varepsilon) - \tfrac{1}{4}\sqrt{\varepsilon}z\,, \\ \rho' &= \varepsilon\,, \end{aligned} \tag{4.5}$$

which limits, when $\varepsilon = 0$, to

$$\begin{aligned} \theta' &= z\,, \\ z' &= -\sin\theta\,, \\ \rho' &= 0\,, \end{aligned} \tag{4.6}$$

an unforced pendulum with redundant parameter ρ. Equation (4.6) has a tube of heteroclinic orbits connecting each pair of saddle points, and our goal here is to study which (if any) survive the perturbation. This is achieved by a Melnikov argument.

Replace $\mu = \sqrt{\varepsilon}$ in (4.5), and append $\mu' = 0$,

$$\begin{aligned} \theta' &= z\,, \\ z' &= -\sin\theta + \mu\hat{\beta}(\rho,\mu) - \tfrac{1}{4}\mu z\,, \\ \rho' &= \mu^2\,, \\ \mu' &= 0\,. \end{aligned} \tag{4.7}$$

As for (4.5), when $\mu = 0$ we have a tube of heteroclinic loops connecting $\mathcal{U}_0$ to $\mathcal{U}_1$, see Fig. 4.1. We need only consider these connections as all others are reproduced by

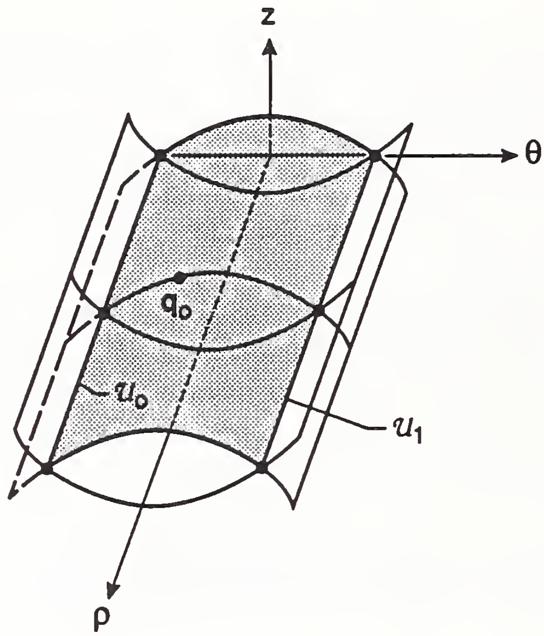

FIG. 4.1. *The heteroclinic loop.*

translation in θ by $2k\pi$. Let $q_0 = q_0(\rho) \in \boldsymbol{R}^2 \times \boldsymbol{S}^1$ be a point on the heteroclinic orbit in $\mu = 0$, $z > 0$, see Fig. 4.1. Moreover, choose $q^-(\rho, \mu)$ smoothly on the unstable manifold W^u of $\mathcal{U}_0$, so that $q^-(\rho, 0) = q^0(\rho)$. Similarly, choose $q^+(\rho, \mu) \in W^s$, the stable manifold of $\mathcal{U}_1$, smoothly so that $q^+(\rho, 0) = q^0(\rho)$. We set $d(\rho, \mu) = f(q_0(\rho)) \wedge (q^-(\rho, \mu) - q^+(\rho, \mu))$, where $q = (z, \theta)$ and $f(q) = (z, -\sin\theta)$, the first two components of the vector field of (4.7) with $\mu = 0$. This measures the distance between W^u and W^s as a function of μ.

Since $d(\rho, 0) = 0$ and d is smooth, we have that $d(\rho, \mu) = \mu D(\rho, \mu)$ and

$$\left.\frac{\partial d}{\partial \mu}\right|_{\mu=0} = D(\rho, 0) = M(\rho),$$

which is the so-called adiabatic Melinkov function (see [15]). Thus,

$$M(\rho) = f(q_0(\rho)) \wedge \left(\left.\frac{\partial q^+}{\partial \mu}\right|_{\mu=0} - \left.\frac{\partial q-}{\partial \mu}\right|_{\mu=0} \right).$$

The quantities $\partial q^- / \partial \mu$ and $\partial q^+ / \partial \mu$ can be viewed as the first two components of a solution of the equation of variations of (4.7), with ρ and μ components 0 and 1 respectively. We write the equation of variations using differential forms, when $\mu = 0$, as

$$\begin{aligned} d\theta' &= dz\,, \\ dz' &= -\cos\theta d\theta + (\hat{\beta} - \tfrac{1}{4}z)d\mu\,, \\ d\rho' &= 0\,, \\ d\mu' &= 0\,. \end{aligned} \tag{4.8}$$

Let $\boldsymbol{v}_1, \boldsymbol{v}_2$ be two solutions of (4.8), consider

$$\omega = d\theta \wedge dz(\boldsymbol{v}_1, \boldsymbol{v}_2),$$

and calculate

$$\omega' = d\theta' \wedge dz + d\theta \wedge dz'$$

$$= (\hat{\beta} - \tfrac{1}{4}z) d\theta \wedge d\mu,$$

from (4.8). We parametrize the underlying heteroclinic orbit so that it is at $\boldsymbol{q}_0$ at $s = 0$. Then set $\boldsymbol{v}_1$ to be the vector field of (4.7) and set $\boldsymbol{v}_2 = (\delta\theta, \delta z, 0, 1)$, with $(\delta\theta(0), \delta z(0)) = \partial \boldsymbol{q}^- / \partial\mu \Big|_{\mu=0}$. It follows that $d\theta \wedge d\mu(\boldsymbol{v}_1, \boldsymbol{v}_2) = z$ and so, denoting $\omega_- = d\theta \wedge d\mu(\boldsymbol{v}_1, \boldsymbol{v}_2)$,

$$\omega_-' = (\hat{\beta} - \tfrac{1}{4}z)z.$$

With the obvious notation, ω_+ is formed using $\partial \boldsymbol{q}^+ / \partial\mu \Big|_{\mu=0}$, and we also have

$$\omega_+' = (\hat{\beta} - \tfrac{1}{4}z)z.$$

We conclude that

$$\omega_-(0) - \omega_+(0) = \int_{-\infty}^{+\infty} (\hat{\beta} - \tfrac{1}{4}z)z \, ds.$$

But

$$\omega_-(0) - \omega_+(0) = \frac{\partial \boldsymbol{q}^-}{\partial\mu} \wedge \boldsymbol{f}(\boldsymbol{q}_0) - \frac{\partial \boldsymbol{q}^+}{\partial\mu} \wedge \boldsymbol{f}(\boldsymbol{q}^0),$$

which implies that

$$\boldsymbol{f}(\boldsymbol{q}_0) \wedge \left(\frac{\partial \boldsymbol{q}^+}{\partial\mu} - \frac{\partial \boldsymbol{q}^-}{\partial\mu} \right) = \int_{-\infty}^{+\infty} (\hat{\beta} - \tfrac{1}{4}z)z \, ds.$$

Hence, as in Guckenheimer and Holmes [11, p. 201],

$$M(\rho) = 2(\pi\hat{\beta} - 1).$$

We see that $M(\rho) = 0$ if $\hat{\beta} = 1/\pi$. Similarly there is a heteroclinic orbit below ($z < 0$) near a ρ-value at which $\hat{\beta} = -1/\pi$. Since $\hat{\beta}$ is decreasing, this gives $\rho_*^- < 0 < \rho_*^+$ and $\tau_*^- < \tau_* < \tau_*^+$ at which connections are made above and below the θ-axis, respectively, where

$$\tau_*^{\pm} = \tau_* + \sqrt{\varepsilon}\rho_*^{\pm}. \tag{4.9}$$

We summarize the results of this section in the following Lemma.

LEMMA 4.1. *There exist ρ_*^-, $\rho_*^+ \in \mathcal{I}$ satisfying $\rho_*^- < 0 < \rho_*^+$, at which heteroclinic connections are made between $\mathcal{U}_k$ and $\mathcal{U}_{k+1}$ above and below the θ-axis, respectively, for all $k \in \boldsymbol{Z}$.*

5. The Resonance Attractor. Recall that $\mathcal{V}$ (Eq. (2.2)) is a closed curve of center equilibrium points of system (1.12) in the solid torus $\mathcal{S}$ (Eq. (2.3)) when $\varepsilon = 0$, which corresponds to countably many curves

$$\mathcal{V}_k = \{s_k(\tau) : \tau \in [0,1]\}, \qquad k \in \boldsymbol{Z} \tag{5.1}$$

in the phase space $\boldsymbol{R}^2 \times \boldsymbol{S}^1$. In this section, we will prove that each of these curves perturbs to a periodic attractor, and we will also prove a Lemma about the domains of attraction of these solutions.

THEOREM 5.1. *For sufficiently small ε, system (1.12) has a unique attracting periodic orbit $\mathcal{V}_k^\varepsilon$ which is $O(\sqrt{\varepsilon})$-close to $\mathcal{V}_k$, for each $k \in \boldsymbol{Z}$.*

We shall prove this Theorem using two Lemmas. We first prove that each $\mathcal{V}_k$ perturbs to a *quasiperiodic* attractor.

LEMMA 5.2. *There exists a positive constant ε_3 such that, for $0 < \varepsilon \leq \varepsilon_3$, system (1.12) has a unique smooth quasiperiodic attractor $\mathcal{V}_k^\varepsilon$ which is $O(\sqrt{\varepsilon})$-close to $\mathcal{V}_k$, for each $k \in \boldsymbol{Z}$.*

Proof. We may restrict our attention to the case $k = 0$. Let $\Theta(\tau) = \arcsin B(\tau)$. Define a change of dependent variables in system (1.12) by

$$\theta = \Theta(\tau) + \mu u, \qquad z = \mu^3 d\Theta/d\tau + \mu v, \tag{5.2}$$

where $\mu = \sqrt{\varepsilon}$. Suppose $|u|$, $|v| \leq \nu$, for some positive constant ν. Then for sufficiently small μ, system (1.12) is equivalent to the system

$$u' = v, \tag{5.3}$$
$$v' = -\Omega^2(\tau)u + \tfrac{1}{2}\mu B(\tau)u^2 - \tfrac{1}{4}\mu v + \rho(u,v,\tau,\mu), \tag{5.4}$$
$$\tau' = \mu^3, \tag{5.5}$$

where

$$\Omega(\tau) = \sqrt{\cos\Theta(\tau)} = \left\{1 - [B(\tau)]^2\right\}^{1/4} \geq 7^{1/4}/2, \tag{5.6}$$

and

$$|\rho|_{C^0} = O(\mu^2). \tag{5.7}$$

Solve (5.5) to get $\tau = \mu^3 s$, and define another change of dependent variables by

$$u(s) = a(s)\cos\sigma + b(s)\sin\sigma, \qquad v(s) = -\Omega(\tau)a(s)\sin\sigma + \Omega(\tau)b(s)\cos\sigma, \tag{5.8}$$

where

$$\sigma(s,\mu) = \int_0^s \Omega(\mu^3 r)\, dr\,. \tag{5.9}$$

Then system (5.3)–(5.5) is transformed into the system

$$\begin{aligned} da/ds &= -R(a,b,s,\mu)\sin(\sigma)/\Omega(\tau)\,, && (5.10)\\ db/ds &= R(a,b,s,\mu)\cos(\sigma)/\Omega(\tau)\,, && (5.11)\end{aligned}$$

where

$$\begin{aligned} R(a,b,s,\mu) &= \tfrac{1}{2}\mu B(\tau)\,(a\cos\sigma + b\sin\sigma)^2 \\ &\quad +\mu\left(-\tfrac{1}{4}+\mu^2 d\Omega/d\tau\right)(-\Omega a\sin\sigma + \Omega b\cos\sigma) + \rho. \qquad (5.12)\end{aligned}$$

If we set $\boldsymbol{x} = (a, b)$, system (5.10)–(5.11) can be written in the form

$$(5.13)\qquad d\boldsymbol{x}/ds = \mu \boldsymbol{q}(s, \boldsymbol{x}, \mu) = \mu \boldsymbol{q}_1(s, \boldsymbol{x}) + O(\mu^2), \qquad \boldsymbol{x} \in D \subset \boldsymbol{R}^2,$$

where $\boldsymbol{q}(\cdot, \boldsymbol{x}, \mu) \in C^\infty(\boldsymbol{R}, \boldsymbol{R}^2)$ is quasiperiodic, and D is compact. This is the standard form of Bogoliubov for the method of averaging (see e.g. [4]). The limit

$$(5.14)\qquad \lim_{T\to\infty} \frac{1}{T}\int_0^T q_1(s, \boldsymbol{x})\, ds = \boldsymbol{Q}_1(\boldsymbol{x}) = -\tfrac{1}{8}\boldsymbol{x}$$

exists uniformly in $\boldsymbol{x}$; $\boldsymbol{Q}_1(0) = 0$; both eigenvalues of its Jacobian, evaluated at $\boldsymbol{x} = 0$, $\partial \boldsymbol{Q}_1(0)/\partial \boldsymbol{x}$, are real and negative; and the corresponding averaged system is given by the uncoupled linear system of equations in $\boldsymbol{R}^2$,

$$(5.15)\qquad d\boldsymbol{y}/ds = -\tfrac{1}{8}\mu \boldsymbol{y}.$$

It thus follows by results of Hale [12, Theorem V.3.1, p. 190] that there is a unique solution of (5.10)–(5.11), $\boldsymbol{x}_\dagger \in C(\boldsymbol{R}\times[0,\hat{\mu}], \boldsymbol{R}^2)$, which is a quasiperiodic attractor. The fact that this solution is smooth follows from [13, Lemma 4.1, p. 101]. By substituting this solution back into the original system (1.12), using the changes of variables (5.8) and (5.2), we have proved there is a unique quasiperiodic attractor $\mathcal{V}_k^\varepsilon$ which lies within an $O(\sqrt{\varepsilon})$-neighborhood of $\mathcal{V}_k$, for each $k \in \boldsymbol{Z}$, for all $0 < \varepsilon \le \varepsilon_3 = \hat{\mu}^2$. □

We next prove a result about the domains of attraction of these quasiperiodic orbits. Let $\boldsymbol{\xi} \cdot s = (\theta(s), z(s), \tau(s))$ be a solution of system (1.12). As in the preceding section, let $\rho = (\tau - \tau_*)/\mu$, where $B(\tau_*) = 0$ and $B'(\tau^*) < 0$, and $\mathcal{I}$ is a compact $O(1)$ interval in ρ containing 0 on which $B(\tau) = \mu\hat{\beta}(\rho, \mu)$. As long as $\rho \in \mathcal{I}$, system (1.12) can be written as the perturbed Hamiltonian system

$$(5.16)\qquad \begin{aligned} \theta' &= \partial H/\partial z + \mu f_1, \\ z' &= -\partial H/\partial \theta + \mu f_2, \\ \rho' &= \mu f_3, \end{aligned}$$

where H is the Hamiltonian

$$(5.17)\qquad H(z, \theta) = \tfrac{1}{2}z^2 - 2\cos^2 \tfrac{1}{2}\theta,$$

and $f_1(z,\theta,\rho,\mu) \equiv 0$, $f_2(z,\theta,\rho,\mu) = \hat{\beta}(\rho,\mu) - \frac{1}{4}z$, $f_3(z,\theta,\rho,\mu) = \mu$. When $\mu = 0$, energy is conserved along the orbits, i.e. $H = h$, where h is a constant. Let

$$\kappa^2 = \tfrac{1}{2}(h+2), \tag{5.18}$$

so that, when $\mu = 0$, $\kappa = 0$ on the stable center equilibrium orbits $(0 \bmod 2\pi, 0, \rho)$ of (5.16), and $\kappa = 1$ on the unstable saddle equilibrium orbits $(\pi \bmod 2\pi, 0, \rho)$ and along the homoclinic orbits which connect their stable and unstable manifolds. Let $h(s) = H(z(s),\theta(s))$ be the value of the Hamiltonian along the solution $\boldsymbol{\xi}\cdot s$. The conditions of the following Lemma specify that at scaled time $\bar{s}$, $\boldsymbol{\xi}\cdot\bar{s}$ is located inside the region of oscillation about the quasiperiodic orbit $\mathcal{V}_j^\varepsilon$.

LEMMA 5.3. *Suppose there exist an integer j and positive constants $\bar{s}$ and α such that $\rho(\bar{s}) \in Int\,(\mathcal{I})$, $(\theta(\bar{s}),0)$ lies between the saddle equilibrium points $\boldsymbol{u}_j(\tau(\bar{s}))$ and $\boldsymbol{u}_{j+1}(\tau(\bar{s}))$ (4.3) of system (1.12) with $\varepsilon = 0$, and $h(\bar{s}) < -\alpha\mu$. Then $\boldsymbol{\xi}\cdot s$ enters the domain of attraction of $\mathcal{V}_j^\varepsilon$ as $s \to \infty$.*

Proof. Neishtadt [19], [21] has shown that averaging can be used to approximate the time-varying energy h, even near a separatrix. Let

$$T(h) = \oint_{H=h} ds \tag{5.19}$$

be the period of the unperturbed motion (the notation means that the integral is calculated along the solution of the unperturbed system with $H(z,\theta) = h$), and let $f_i^0 = f_i(z,\theta,\rho,0)$. Consider the averaged system

$$\begin{aligned} \bar{h}\,' &= \frac{\mu}{T(\bar{h})}\oint_{H=\bar{h}} \left(\frac{\partial H}{\partial z} f_1^0 + \frac{\partial H}{\partial \theta} f_2^0 + \frac{\partial H}{\partial \rho} f_3^0\right) ds \\ &= \frac{\mu}{T(\bar{h})}\oint_{H=\bar{h}} z(s)\left[\hat{\beta}(\bar{\rho}(s),0) - \tfrac{1}{4}z(s)\right] ds, \end{aligned} \tag{5.20}$$

$$\bar{\rho}\,' = \frac{\mu}{T(\bar{h})}\oint_{H=\bar{h}} \frac{\partial H}{\partial \rho} f_3^0\, ds = 0, \tag{5.21}$$

and suppose that $\bar{h}(\bar{s}) = h(\bar{s})$, $\bar{\rho}(\bar{s}) = \rho(\bar{s})$. Then for $\mu\bar{s} \le \mu s \le S$, where S is a positive constant, Neishtadt has proved that $h(s) < -\alpha\mu$, and

$$\left|h(s) - \bar{h}(s)\right| + |\rho(s) - \bar{\rho}(s)| = O(\mu|\ln\mu|). \tag{5.22}$$

Equation (5.20) can be written in the form

$$\bar{h}\,' = -\frac{4\mu}{K(\kappa)}\left[E(\kappa) - (1-\kappa^2)K(\kappa)\right], \tag{5.23}$$

where K and E are the complete elliptic integrals of the first and second kind, respectively. It can be shown that the right-hand side of equation (5.23) is strictly negative for $\kappa \in (0,1)$, and it equals 0 when $\kappa = 0$, which corresponds to $\bar{h} = -2$, by (5.18). Thus, as s increases, the initially small but negative solution of the averaged equation (5.23) decreases monotonically towards $\bar{h} = -2$, which corresponds to the stable center equilibrium solution $(2j\pi, 0, \rho_0)$ of system (1.12) with $\mu = 0$. Furthermore, using (5.18), equation (5.23) can be written in the form

$$\kappa K(\kappa)\kappa' = -\mu\left[E(\kappa) - (1-\kappa^2)K(\kappa)\right].$$

By a well-known identity in elliptic integrals [7, Eq. (710.04)], this implies that

$$W' = -\mu W,$$

where

$$W(\kappa) = E(\kappa) - (1-\kappa^2)K(\kappa) = \tfrac{1}{8}\pi I(\kappa), \tag{5.24}$$

and I is the value of the action of the unperturbed nonlinear pendulum with Hamiltonian (5.17) and energy $h = 2(\kappa^2 - 1)$, so that W increases monotonically with κ. By hypothesis, initially $\kappa = \tilde{\kappa}(\mu) < \sqrt{1 - \frac{1}{2}\alpha\mu} < 1$, which implies that $W = \tilde{W}(\mu) < 1$ initially as well. Let $\sigma = \mu(s - \tilde{s})$. Then

$$\frac{dW}{d\sigma} = -W, \tag{5.25}$$

so that $W(\sigma) = \tilde{W}\exp(-\sigma) = \frac{1}{2}\tilde{W}$ when $\sigma = \mu(s - \tilde{s}) = \ln 2$. Hence, when $s = \check{s} = \tilde{s} + \ln 2/\mu$, $\kappa = \check{\kappa}(\mu)$, where $0 < \check{\kappa} < \tilde{\kappa}$. We claim there exists a constant $0 < c_6 < 1$, independent of μ, such that $\check{\kappa} < c_6$ for sufficiently small μ. If not, by continuous dependence of κ on the parameter μ, there would exist a sequence $\{\mu_n\}$ with $\lim_{n\to\infty}\mu_n = 0$ such that $\check{\kappa}(\mu_n) \uparrow 1$. But by (5.24),

$$\tfrac{1}{2} > \tfrac{1}{2}\tilde{W} = E(\check{\kappa}) - (1-\check{\kappa}^2)K(\check{\kappa}),$$

and by properties of the elliptic integrals,

$$\left[E(\check{\kappa}) - (1-\check{\kappa}^2)K(\check{\kappa})\right] \uparrow 1 \text{ as } \check{\kappa} \uparrow 1,$$

which is a contradiction. It follows that, for sufficiently small μ, there exists a positive constant c_7, independent of μ, such that $\check{h} < -c_7 < 0$, where $\check{h} = h(\check{s})$ and $\check{s} = \tilde{s} + \ln 2/\mu$. We also note that $\check{\kappa}$ is strictly $O(1)$. This follows from the fact that W is initially strictly $O(1)$, and there is insufficient time for the damping to decrease it to $o(1)$. Since $W \sim \kappa^2$ as $\kappa \downarrow 0$, $\check{\kappa}$ must also be $O(1)$. Thus, we must still show that $\xi \cdot s$ enters the domain of attraction of the periodic orbit $\mathcal{V}_j^\epsilon$ as s increases.

For $s \geq \check{s}$, we argue as follows. Define the energy

$$G(z,\theta,\tau) = \tfrac{1}{2}z^2 + V(\theta,\tau), \tag{5.26}$$

with

$$V(\theta,\tau) = \begin{cases} -2\cos^2\frac{1}{2}\theta - B(\tau)[\theta - \theta_{j+1}^u(\tau)] & \text{if } B(\tau) \geq 0, \\ -2\cos^2\frac{1}{2}\theta - B(\tau)[\theta - \theta_j^u(\tau)] & \text{if } B(\tau) \leq 0. \end{cases} \tag{5.27}$$

where $\theta_k^u(\tau) = -[\Theta(\tau) + \pi] + 2k\pi$, so that $(\theta_k^u(\tau), 0)$ is the k-th saddle equilibrium point (4.3) of system (4.1) at $\{\tau = \text{const}\}$. For fixed τ, the level curves $G = \gamma(\tau)$ represent orbits of solutions of the undamped pendulum equation with constant torque $B(\tau)$ (Eq. (4.1)),

$$l_{\gamma(\tau)} = \left\{(\theta, z) \in \boldsymbol{R}^2 : z = \pm\sqrt{2[\gamma(\tau) - V(\theta,\tau)]}\right\}. \tag{5.28}$$

Since $\partial V(\theta,\tau)/\partial\theta = \sin\theta - B(\tau)$ and $\partial^2 V(\theta,\tau)/\partial\theta^2 = \cos\theta$, in each slice $\{\tau = \text{const}\}$, the function V has a local minimum at $\theta_k^s(\tau) = \Theta(\tau) + 2k\pi$ and a local maximum at $\theta_k^u(\tau)$, for each $k \in \boldsymbol{Z}$ (see [22]). Hence, when $B(\tau) > 0$ and $\gamma(\tau) = V(\theta_{j+1}^u(\tau), \tau)$, the set (5.28) consists of the saddle point $\boldsymbol{u}_{j+1}(\tau)$ and the homoclinic separatrix loop which leaves and enters this saddle point and encloses the center $\boldsymbol{s}_j(\tau)$. Similarly, when $B(\tau) < 0$ and $\gamma(\tau) = V(\theta_j^u(\tau), \tau)$, the set (5.28) consists of the saddle point $\boldsymbol{u}_j(\tau)$ and the homoclinic separatrix loop which leaves and enters this saddle point and encloses the center $\boldsymbol{s}_j(\tau)$. When $B(\tau) = 0$, the set (5.28) consists of both saddle points and the two heteroclinic orbits connecting them. Also, when $\gamma(\tau) = V(\theta_j^s(\tau), \tau)$, the set (5.28) consists of the single equilibrium point $\boldsymbol{s}_j(\tau)$.

Let $g(s)$ be the value of G at the point $\boldsymbol{\xi} \cdot s$, and define $\check{g} = g(\check{s})$. Then $\check{g} = \check{h} + O(\mu)$ since $B(\check{\tau}) = O(\mu)$, where $\check{\tau} = \tau(\check{s})$, so that for sufficiently small μ, there is a positive constant c_8 such that $\check{g} < -c_8$. Consequently, the point $\check{\boldsymbol{\xi}} = \boldsymbol{\xi} \cdot \check{s}$ is located on the level curve $G(\theta, z, \check{\tau}) = \check{g}$, which is a simple closed curve containing the center equilibrium point $\boldsymbol{s}_j(\check{\tau})$. We also know there is a slightly larger constant $g^\dagger$ for which the same is true, and this curve contains the one corresponding to $\check{g}$ in its interior, see Fig. 5.1. We next find such a curve for each $\tau \in \boldsymbol{S}^1$. Define

$$a = \begin{cases} \left[g^\dagger - V(\theta_j^s(\check{\tau}), \check{\tau})\right] / \left[V(\theta_{j+1}^u(\check{\tau}), \check{\tau}) - V(\theta_j^s(\check{\tau}), \check{\tau})\right] & \text{if } B(\check{\tau}) \geq 0, \\ \left[g^\dagger - V(\theta_j^s(\check{\tau}), \check{\tau})\right] / \left[V(\theta_j^u(\check{\tau}), \check{\tau}) - V(\theta_j^s(\check{\tau}), \check{\tau})\right] & \text{if } B(\check{\tau}) \leq 0, \end{cases} \tag{5.29}$$

so that $0 < a < 1$, and also define

$$e(\tau) = \begin{cases} aV(\theta_{j+1}^u(\tau), \tau) + (1-a)V(\theta_j^s(\tau), \tau) & \text{if } B(\tau) \geq 0, \\ aV(\theta_j^u(\tau), \tau) + (1-a)V(\theta_j^s(\tau), \tau) & \text{if } B(\tau) \leq 0. \end{cases} \tag{5.30}$$

The energy $e(\tau)$ thus satisfies $e(\check{\tau}) = g^\dagger$, and $V(\theta_j^s(\tau), \tau) < e(\tau) < V(\theta_{j+1}^u(\tau), \tau)$ if $B(\tau) \geq 0$, and $V(\theta_j^s(\tau), \tau) < e(\tau) < V(\theta_j^u(\tau), \tau)$ if $B(\tau) \leq 0$. Thus, for each $\tau \in \boldsymbol{S}^1$, the level curve $l_{e(\tau)}$ corresponds to a periodic solution of the first kind of the undamped system (4.1) encircling the quasiperiodic attractor. Consequently, the set

$$\mathcal{D}_j = \cup_\tau \{(\theta, z, \tau) \in \mathcal{S} : G(z, \theta, \tau) \leq e(\tau)\} \tag{5.31}$$

is a closed neighborhood of $\mathcal{V}_j^\epsilon$ which contains $\boldsymbol{\xi} \cdot s$ in its interior for all $s \geq \check{s}$ (see Fig. 5.1), since $\check{\boldsymbol{\xi}}$ lies in the interior of $\mathcal{D}_j$, and the vector field of system (1.12) is either tangent to or directed into $\mathcal{D}_j$. This can be verified by comparing the first two components of the vector field of the undamped system (4.1) with those of the damped system (1.12) along the boundary of $\mathcal{D}_j$,

$$(z, B(\tau) - \sin\theta) \wedge (z, B(\tau) - \sin\theta - \tfrac{1}{4}\mu z) = -\tfrac{1}{4}\mu z^2.$$

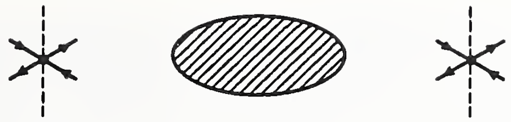

FIG. 5.1. *Section at $\tau = \check\tau$ of the neighborhood $\mathcal{D}_j$ of the quasiperiodic attractor.*

Hence, $g(s)$ is bounded for all $s \geq \check{s}$.

Now, by Lemma 5.2, there exists a positive constant c_9, independent of μ, such that, if

$$g(s) - V(\theta^e_j(\tau(s)), \tau(s)) < c_9\mu,$$

then $\xi \cdot s$ is in the domain of attraction of $\mathcal{V}^\varepsilon_j$. Suppose that $\xi \cdot s$ never enters this domain of attraction, so that

$$g(s) - V(\theta^e_j(\tau(s)), \tau(s)) \geq c_9\mu$$

for all $s \geq \check{s}$. Then

$$\begin{aligned}
\frac{dg}{ds} &= \frac{\partial G}{\partial z}\frac{dz}{ds} + \frac{\partial G}{\partial \theta}\frac{d\theta}{ds} + \frac{\partial G}{\partial \tau}\frac{d\tau}{ds} \\
&= z(s)\left[B(\tau(s)) - \sin\theta(s) - \tfrac{1}{4}\mu z(s)\right] + [\sin\theta(s) - B(\tau(s))]\, z(s) + O(\mu^3) \\
&= -\tfrac{1}{4}\mu\, [z(s)]^2 + O(\mu^3) \\
&= -\tfrac{1}{4}\mu\, \{2\,[g(s) - V(\theta(s), \tau(s))]\} + O(\mu^3) \\
&\leq -\tfrac{1}{4}\mu \left\{2\left[g(s) - V(\theta^e_j(s), \tau(s))\right]\right\} + O(\mu^3) \\
&\leq -\tfrac{1}{2}c_9\mu^2 + O(\mu^3) \\
&< 0
\end{aligned}$$

for sufficiently small μ, which implies that $g(s) - g(\check{s}) \to -\infty$ as $s \to \infty$. This contradicts the fact that g remains bounded for all $s \geq \check{s}$. Hence, $\xi \cdot s$ enters the domain of attraction of $\mathcal{V}^\varepsilon_j$ as $s \to \infty$. □

To complete the proof of Theorem 5.1, we must still prove that the attracting quasiperiodic orbits $\mathcal{V}^\varepsilon_k$ of Lemma 5.2 are actually periodic. We will do this by finding an invariant neighborhood of the quasiperiodic attractor in much the same way we did for the rotary attractor in Section 3 and then using a fixed-point argument. In the "frozen" system (2.5), there is a compact curve of stable spiral equilibrium points $\hat{\mathcal{V}}^\varepsilon_k = \cup_\tau\{(\theta, z, \tau) : \theta = \Theta(\tau) + 2k\pi,\ z = 0\}$, and the real parts of the eigenvalues of system (2.5) linearized about the equilibrium point for each fixed τ are both equal to $-\frac{1}{8}\sqrt{\varepsilon}$. From this it follows there is a tubular neighborhood $\hat{\mathcal{T}}^\varepsilon_k$ of the curve $\hat{\mathcal{V}}^\varepsilon_k$ into which the vector field of system (2.5) is directed, and the component of this vector field normal to $\hat{\mathcal{T}}^\varepsilon_k$ is $O(\sqrt{\varepsilon})$. Hence, $\hat{\mathcal{T}}^\varepsilon_k$ is an $O(1)$ positively invariant neighborhood of $\hat{\mathcal{V}}^\varepsilon_k$

in the "frozen" system. Since the "frozen" system differs from the "un-frozen" system (1.12) only in the τ-component of the vector field, and this component is only $O(\varepsilon^{3/2})$ in (1.12), it follows there is also an $O(1)$ positively invariant tubular neighborhood $\mathcal{T}_k^\varepsilon$ of the quasiperiodic orbit $\mathcal{V}_k^\varepsilon$ in system (1.12). By periodicity of the vector field (1.12) in τ, there is a Poincaré map $\boldsymbol{P} : \boldsymbol{R}^2 \to \boldsymbol{R}^2$ which takes the plane $\{\tau = 0\}$ into itself, and since $\mathcal{T}_k^\varepsilon$ is invariant under the flow of system (1.12), $\boldsymbol{P}$ maps the section $\mathcal{T}_k^\varepsilon \cap \{\tau = 0\}$ of this invariant tube into itself. Hence, by Brouwer's Fixed Point Theorem, there is a fixed point of $\boldsymbol{P}$ in $\mathcal{T}_k^\varepsilon \cap \{\tau = 0\}$, for each $k \in \boldsymbol{Z}$. By properties of the Poincaré map, these fixed points correspond to periodic solutions of period $1/\varepsilon^{3/2}$ of system (1.12). By Lemma 5.3, each periodic solution must lie in the domain of attraction of the corresponding quasiperiodic solution $\mathcal{V}_k^\varepsilon$, which by Lemma 5.2, is unique in a neighborhood of $\mathcal{V}_k$. It thus follows that $\mathcal{V}_k^\varepsilon$ is in fact a periodic solution of system (1.12) for each $k \in \boldsymbol{Z}$. □

6. Proof of Capture Theorem. In Section 4 we determined some values of τ, namely τ_*^- and τ_*^+, at which heteroclinic orbits existed. From those considerations, we see that $|\tau_*^+ - \tau_*^-| = O(\sqrt{\varepsilon})$ but in ρ-variables, $|\rho_*^+ - \rho_*^-| = O(1)$, where $\rho = (\tau - \tau_*)/\sqrt{\varepsilon}$ and $\tau_* = \frac{1}{6}$, so that $B(\tau_*) = 0$. We shall determine various values of τ to be used in the proof of the Theorem. These values will all be within $O(\sqrt{\varepsilon})$ of τ_* but be bounded apart in ρ. We shall use repeatedly in this section that $B'(\tau) < 0$ on $(0, \hat{\tau}]$ where $\hat{\tau} > \tau_*$.

The proof strategy consists in following a vertical section of the strip S_ε under the action of the flow. Let $L = L_{\hat{\theta}} \subset S_\varepsilon$ be a vertical segment

(6.1) $$L_{\hat{\theta}} = \{(\hat{\theta}, z, 0) : \hat{\theta} \text{ fixed and } [x_*(0) - d_0]/\sqrt{\varepsilon} \le z \le [x_*(0) + d_0]/\sqrt{\varepsilon}\},$$

see Fig. 6.1. We shall thus consider sets of the form $L \cdot s_i$ where s_i will take on various

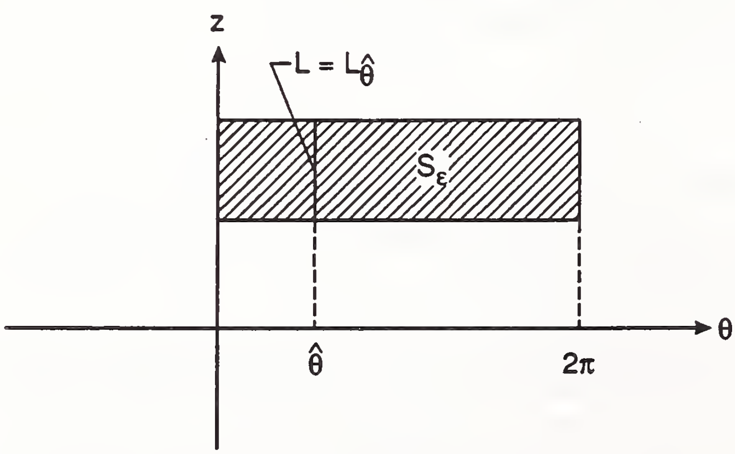

FIG. 6.1. *The vertical section $L = L_{\hat{\theta}}$.*

values. Indeed, we shall determine values of τ, denoted $\tau_0, \cdots \tau_5$, and the values of s will be related to these τ-values as the scaled times needed to attain the relevant τ-slice, i.e.,

$$s_i = \tau_i/\varepsilon^{3/2}, \qquad i = 0, \ldots, 5.$$

Suppose that the statement of the Theorem can be shown to hold for $L = L_{\hat{\theta}}$, and each $\hat{\theta} \in [0, 2\pi]$, wherein the estimate on the measure of the set not captured in resonance is uniform in $\hat{\theta}$. It then follows that the Theorem will hold in general by integrating the estimate in θ.

In the following we set $\mathcal{B}_k(\tau)$ to be a neighborhood of $\boldsymbol{u}_k(\tau)$, which is a square in the eigenvector coordinates. The square $\mathcal{B}_k(\tau)$ is to be small enough so that the system can be assumed to be effectively linear in $\mathcal{B}_k(\tau)$, see below. From the discussion in Section 4, $\boldsymbol{u}_k(\tau)$ perturbs to a saddle periodic orbit $\boldsymbol{u}_k(\tau, \varepsilon)$ for ε sufficiently small. If ε is small enough $\boldsymbol{u}_k(\tau, \varepsilon) \in \mathcal{B}_k(\tau)$. In the following, we have set $\boldsymbol{u}_k(\tau, \varepsilon) = (\theta_k(\tau, \varepsilon), z_k(\tau, \varepsilon), \tau)$.

We firstly set $\tau = \tau_0$ in the following Lemma.

LEMMA 6.1. *There is a $\tau_0 < \tau_*^-$ for which all points in*

$$L \cdot s \cap \{\theta = \theta_k(\tau, \varepsilon), \quad z > 0\}$$

must lie in $\mathcal{B}_k(\tau)$, for any $\tau \in [\tau_0, \hat{\tau}]$ and any section L of S_ε.

Proof. This is an easy consequence of Lemma 3.4, with $\tau_0 = \tau^\#$. □

Definition. A *left primary segment* $W_k^{s,L}(\tau)$ of the stable manifold $W_k^s(\tau)$ of $\boldsymbol{u}_k(\tau, \varepsilon)$ (relative to the Poincaré map on the τ-slice) is the connected component of $W_k^s(\tau) \cap \{\theta_{k-1}(\tau, \varepsilon) \le \theta \le \theta_k(\tau, \varepsilon)\}$ containing $\boldsymbol{u}_k(\tau, \varepsilon)$. Similarly, a *right primary segment* $W_k^{u,R}(\tau)$ of the unstable manifold $W_k^u(\tau)$ of $\boldsymbol{u}_k(\tau, \varepsilon)$ is the connected component of $W_k^u(\tau) \cap \{\theta_k(\tau, \varepsilon) \le \theta \le \theta_{k+1}(\tau, \varepsilon)\}$ containing $\boldsymbol{u}_k(\tau, \varepsilon)$; see Fig. 6.2.

The corresponding right and left primary segments of $W_k^s(\tau)$ and $W_k^u(\tau)$ are defined in the obvious way. With these definitions, we can determine the first values of τ. From

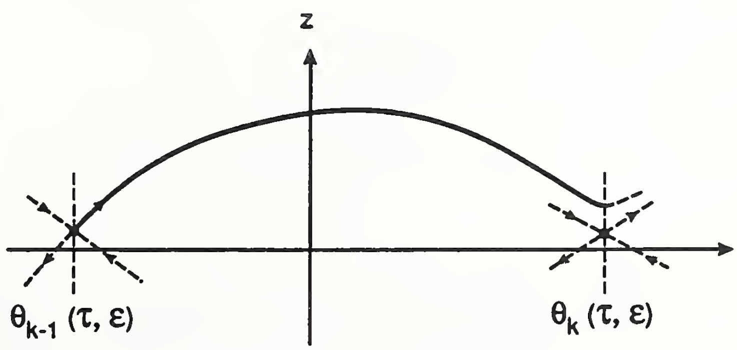

FIG. 6.2. *Primary segment of $W_{k-1}^u(\tau)$.*

the discussion in Section 4, we can set τ_1 and τ_5 so that

$$\tau_*^- < \tau_1 < \tau_5 < \tau_*^+$$

with the property that, in $\boldsymbol{R}^2 \backslash \{\mathcal{B}_{k-1}(\tau) \cup \mathcal{B}_k(\tau)\}$

$$W_{k-1}^{u,R}(\tau) \cap W_k^{s,L}(\tau) = \emptyset,$$

for all $\tau \in [\tau_1, \tau_5]$. Moreover, we can guarantee that $W_{k-1}^{u,R}(\tau)$ and $W_k^{s,L}(\tau)$ are separated by a distance bounded below by $c_{10}\sqrt{\varepsilon}$ in $\boldsymbol{R}^2 \backslash \{\mathcal{B}_{k-1}(\tau) \cup \mathcal{B}_k(\tau)\}$, where c_{10} is a constant depending only on τ_1 and τ_5. Similarly $W_k^{u,L}(\tau)$ and $W_{k-1}^{s,R}(\tau)$ are also uniformly separated in $\boldsymbol{R}^2 \backslash \{\mathcal{B}_{k-1}(\tau) \cup \mathcal{B}_k(\tau)\}$, for all $\tau \in [\tau_1, \tau_5]$.

In the next Lemma we shall need to make estimates on the behavior of trajectories as they pass near the saddle periodic orbit $\mathcal{U}_k$. This is facilitated by the use of Fenichel coordinates, see Fenichel [9], Jones and Kopell [14]. These coordinates straighten out the stable and unstable manifolds of $\mathcal{U}_k$, even for $\varepsilon \neq 0$. In Fenichel coordinates, near $\mathcal{U}_k$, system (1.12) can be rewritten as

$$
\begin{aligned}
a' &= \lambda_+ a + a g_1(a, b, \tau)\,, \\
b' &= \lambda_- b + b g_2(a, b, \tau)\,, \\
\tau' &= \varepsilon^{3/2}\,,
\end{aligned}
\tag{6.2}
$$

where $\lambda_+ > 0 > \lambda_-$ and $|g_i| < \delta$. These are valid in $\mathcal{B}_k = \cup_\tau \mathcal{B}_k(\tau)$, which we take to be a neighborhood of the periodic orbit $\mathcal{U}_k = \{(0,0,\tau) : \tau \in [0,1]\}$. Indeed, we set

$$\mathcal{B}_k(\tau) = \{(a, b, \tau) : |a| \leq \eta, |b| \leq \eta\}.$$

The quantity δ depends on η and can be made as small as desired by choosing η small. Note that, in Fenichel coordinates, the set $\{a = 0\}$ is the (local) stable manifold of $\mathcal{U}_k$ and $\{b = 0\}$ is its (local) unstable manifold. The first quadrant $\{a > 0,\ b > 0\} \cap \mathcal{B}_k(\tau)$ we shall denote $\mathcal{R}_k(\tau)$; this is the region between the (local) stable and unstable manifold of $\mathcal{U}_k(\tau)$ that points "upwards" in (θ, z, τ) coordinates. As above, we denote $\mathcal{R}_k = \cup_\tau \mathcal{R}_k(\tau)$.

We next choose a value of τ_2 such that, in ρ coordinates, ρ_1, ρ_2 and ρ_5 are $O(1)$ apart (using the obvious correspondence between τ and ρ), and we show this forces $L \cdot s$ to be close to $W_k^u \cup W_k^s$ inside $\mathcal{B}_k(\tau)$.

LEMMA 6.2. *For any $y \in L \cdot s_2 \cap \mathcal{B}_k(\tau_2)$,*

$$d(y, W_k^u(\tau) \cup W_k^s(\tau)) = O\left(\exp\left(-c_{11}/\varepsilon\right)\right).$$

Proof. We firstly observe that if $x \cdot s_2 \in \mathcal{R}_k(\tau_2)$, with $x \in L$, and $x \cdot s \in \mathcal{B}_j(\tau)$ for some $s_1 \leq s \leq s_2$, then we must have $x \cdot s \in \mathcal{R}_j(\tau)$. This follows easily from the topological configuration of the stable and unstable manifold of the saddle periodic orbit. If $x \cdot s$ enters a part of $\mathcal{B}_j(\tau)$ other than $\mathcal{R}_j(\tau)$ then this argument shows that it can never enter $\mathcal{R}_i(\tau)$ at another i or later τ.

The key point in the proof of this Lemma is to show that the trajectory $\{x \cdot s : s \in [s_1, s_2]\}$ can only pass through a bounded number of the neighborhoods $\mathcal{B}_k$. It takes scaled time $O(1)$ between leaving $\mathcal{B}_k$ and entering $\mathcal{B}_{k+1}$. In such a passage, clearly $z \geq c_{12} > 0$ for some constant c_{12}. The energy $E_0(z, \theta) = \frac{1}{2}z^2 - [\cos\theta + B(\tau_0)\theta]$ must decrease during such a passage by a constant K. If $x \cdot s$, for $s \in [s_1, s_2]$, has N such jumps then E_0 has decreased by, at least, NK. While the trajectory passes through $\mathcal{B}_k$, E_0 can increase, and we must estimate this increase.

Suppose that the trajectory $x \cdot s$, $s \in [s_1, s_2]$, passes through $\mathcal{B}_i, \mathcal{B}_{i+1}, \ldots, \mathcal{B}_{i+N}$; note that it cannot miss a $\mathcal{B}_j$ by Lemma 6.1. In each $\mathcal{B}_{i+j}, 0 \leq j < N$, the trajectory must be above W^s_{i+j+1} for it to be able to continue in the sequence. Suppose that $x \cdot s$ enters $\mathcal{B}_{i+j}$ at a point a_0^j. Then it follows that $a_0^j \geq c_{13}\sqrt{\varepsilon}$ for some constant $c_{13} > 0$, independent of j. The time S^j spent in $\mathcal{B}_{i+j}$ is estimated by

$$S^j \leq \frac{1}{\lambda_+ - \delta} \ln\left(\frac{\eta}{a_0^j}\right)$$

from integrating (6.2). Hence

$$S^j \leq c_{14}|\ln \varepsilon|, \tag{6.3}$$

where $c_{14} = c_{14}(\lambda_+, \eta)$. In $\mathcal{R}_{i+j}$, it is possible that $z < 0$ but it has a lower bound $z \geq -c_{15}\sqrt{\varepsilon}$ since the periodic orbit is within $O(\sqrt{\varepsilon})$ of $z = 0$ by Fenichel's Theorems. An easy calculation shows that

$$\frac{dE_0}{ds} \leq c_{16}\sqrt{\varepsilon} \tag{6.4}$$

in such an s-interval. But from (6.3) and (6.4), E_0 can increase by, at most, $c_{17}\sqrt{\varepsilon}|\ln \varepsilon|$. Since there are N such possible increases, we see that E_0 has a net decrease of at least

$$N(K - c_{17}\sqrt{\varepsilon}|\ln \varepsilon|).$$

As $\varepsilon \to 0$ and $N \to +\infty$, this increases without bound. But this contradicts the fact that $x \cdot s_2 \in \mathcal{B}_k$ for some k. It follows that $x \cdot s$ passes through a bounded number of $\mathcal{B}_k$'s.

The trajectory spends time $O(1/\varepsilon)$ between $\tau = \tau_1$ and $\tau = \tau_2$ (which we postulated to be $O(1)$ apart in ρ-coordinates). Since it spends $O(1)$ scaled time between $\mathcal{B}_k$'s and hits only finitely many, it must spend $O(1/\varepsilon)$ in some $\mathcal{B}_k$. In that $\mathcal{B}_k$, the trajectory passes through $\mathcal{R}_k$ and from (6.2) we must have a_0^k exponentially small. It follows that $d(x \cdot s, W_k^u(\tau) \cup W_k^s(\tau))$ is exponentially small while the trajectory lies in $\mathcal{R}_k$.

However, this passage through $\mathcal{R}_k$ must be the last as the next neighborhood hit would be in a part of $\mathcal{B}_k$ other than $\mathcal{R}_k$. This proves the Lemma. □

The preceding Lemmas have set up a configuration for $L \cdot s_2$, namely that it intersects $\theta = \theta_k(\tau_2, \varepsilon)$ $(z > z_k(\tau, \varepsilon))$ inside $\mathcal{B}_k(\tau_2)$ (Lemma 6.1) and it lies exponentially close to $W_k^u(\tau_2)$ at the exit point of $\mathcal{B}_k(\tau_2)$. We can thus divide $L \cdot s_2$ into segments $\mathcal{L}^k$ as follows. Let the original section $L = L_{\hat{\theta}}$ be parametrized by $z_0 \leq z \leq z_1$, and set $z^k = \min\{z \geq z_0 : (\theta, z, 0) \cdot s_2$ has θ component equal to $\theta_k(\tau_2, \varepsilon)\}$. If $L^k = L \cap \{z^k \leq z \leq z^{k+1}\}$, then set $\mathcal{L}^k = L^k \cdot s_2$, so that $\mathcal{L}^k$ is a segment of $L \cdot s_2$ which reaches from $\theta = \theta_k(\tau, \varepsilon)$ to $\theta = \theta_{k+1}(\tau, \varepsilon)$; see Fig. 6.3. We also define $\mathcal{L}^{j-1} = L^{j-1} \cdot s_2$, where $L^{j-1} = \{z \leq z^j\}$; here, z^j is the smallest z for which $L \cdot s_2$ intersects $\theta = \theta_j(\tau, \varepsilon)$. We make a similar definition for $\{z \geq z^j\}$. In the next Lemma, we shall study the interaction between $\mathcal{L}^k$ and the stable manifold of the saddle point.

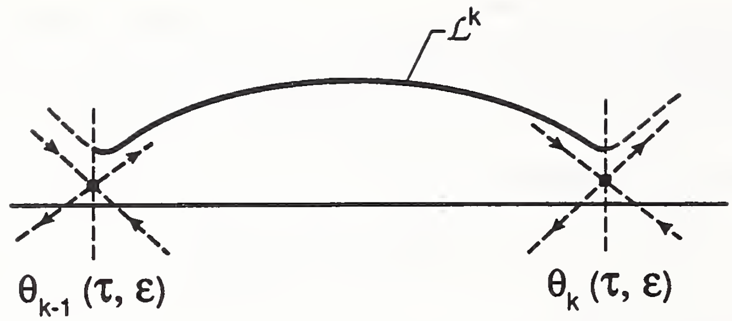

FIG. 6.3. *The segments $\mathcal{L}^k$.*

LEMMA 6.3. *Each segment $\mathcal{L}^k$ has exactly one intersection with $W^{s,L}_{k+1}(\tau_2)$.*

Proof. It follows from Lemma 6.2 that there is at least one intersection, as $L \cdot s_2$ lies below $W^{s,L}_{k+1}$ in $\mathcal{R}_k(\tau_2)$, since it is exponentially close to $W^{u,R}_k$ at that point and by the Melnikov calculation, $W^{u,R}_k$ and $W^{s,L}_{k+1}$ are $O(\sqrt{\varepsilon})$ apart.

It remains to show that there is at most one such intersection. We shall compare the tangent vector to $L \cdot s_2$ at such an intersection point with a tangent vector to $W^{s,L}_{k+1}(\tau_2)$. Indeed, let $v_1(s)$ be tangent to $L \cdot s$ so that, at $s = 0$, $v_1(0) = (0,1,0)$ in (θ, z, τ)-space. If $\hat{v}$ is tangent to $W^{s,L}_{k+1}(\tau_2)$, it suffices to show that the basis in (θ, z) given by $(\hat{v}, v_1(0))$ (forgetting the τ components) has positive orientation, for then $L \cdot s_2$ can cross $W^{s,L}_{k+1}(\tau_2)$ in only one direction, and multiple intersections are prevented.

First, let $v_1(s)$ and $v_2(s)$ be any tangent vectors carried along by the flow of (1.12). The area spanned by (v_1, v_2) in the (θ, s)-plane is given by $d\theta \wedge dz(v_1, v_2)$. As in Section 4 above, $d\theta$ and dz satisfy the equations of variation, which for (1.12) are given by

$$\begin{aligned} d\theta' &= dz\,, \\ dz' &= -\cos\theta d\theta + B'(\tau)d\tau - \tfrac{1}{4}\sqrt{\varepsilon}dz\,, \\ d\tau' &= 0\,. \end{aligned} \tag{6.5}$$

Setting $\Gamma = d\theta \wedge dz(v_1, v_2)$, we calculate

$$\Gamma' = B'(\tau)d\theta \wedge d\tau - \tfrac{1}{4}\sqrt{\varepsilon}\Gamma. \tag{6.6}$$

We specify that v_1 should be the vector field itself, $v_1 = (\theta', z', \tau')$, and v_2 a tangent vector, $v_2 = (\delta\theta, \delta z, 0)$, which at $s = 0$ takes the value $(0,1,0)$. In this case, (6.6) becomes

$$\Gamma' = -\varepsilon^{3/2}\delta\theta B'(\tau) - \tfrac{1}{4}\sqrt{\varepsilon}\Gamma. \tag{6.7}$$

Now, the trajectory must stay in $z > 0$, as otherwise it will move to the left as s increases; but this contradicts the C^1-closeness to the saddle orbit as guaranteed by Fenichel's Theorems. Then $\Gamma(0) = \theta' = z > 0$ and if $\Gamma = 0$, $\Gamma' > 0$ (as long as $z > 0$), we see that $\Gamma > 0$ at the intersection point in $\tau = \tau_2$. This says that the basis formed by the vector field and the tangent vector to $L \cdot s_2$ (in that order) has positive orientation, which is not quite what we need. However, it is also true that the basis formed by the tangent vector to $W_k^{s,L}(\tau_2)$, namely $\hat{v}$, and the vector field (in that order) also has positive orientation (by repeating the argument). The Lemma follows. □

The next step is to choose a $\tau = \tau_3$ (which is $O(1)$ larger than τ_2 in ρ). On account of Lemma 6.2, we can choose $\tau = \tau_3$ so that $L \cdot s_3$ either lies in the capture set or in some $\mathcal{B}_k(\tau_3)$. We shall now find how much of it lies in $\mathcal{B}_k(\tau_3)$.

Let $y \cdot s_3 \in L \cdot s_3 \cap W_k^{s,L}(\tau_3)$. Then $y \cdot s_3 \in \mathcal{B}_k(\tau_3)$. Set $\Sigma_k^\varepsilon(\tau)$ to be the strip of width ε^2 about $W_k^s(\tau)$ inside $\mathcal{B}_k(\tau)$, i.e., $\Sigma_k^\varepsilon(\tau) = \{(a, b, \tau) : |a| \leq \varepsilon^2\}$ in Fenichel coordinates. Further, recalling that η is the length of a side of $\mathcal{B}_k(\tau)$, set $\hat{\Sigma}_k^\varepsilon(\tau) = \{(a, b, \tau) \in \Sigma_k^\varepsilon : b \geq \eta/m\}$; the integer m is chosen so that $z \geq c_{18} > 0$ in $\hat{\Sigma}_k^\varepsilon(\tau)$. Now choose $\hat{\tau}_3 < \tau_3$ so that $y \cdot \hat{s}_3 \in \hat{\Sigma}_k^\varepsilon(\hat{\tau}_3)$, which can easily be done. In the following, we shall study the segment of $L \cdot \hat{s}_3$ which lies in $\hat{\Sigma}_k^\varepsilon(\hat{\tau}_3)$. Let $\hat{y} \in L$ for which $\hat{y} \cdot \hat{s}_3 \in L \cdot \hat{s}_3 \cap \hat{\Sigma}_k^\varepsilon(\hat{\tau}_3)$, and $(\widehat{\delta\theta}(s), \widehat{\delta z}(s), 0)$ be the tangent vector to $L \cdot s$ at $\hat{y} \cdot s$ which takes the value $(0, 1, 0)$ at $s = s_0$; see Fig. 6.4.

LEMMA 6.4. *For any $\hat{y} \in L$, for which $\hat{y} \cdot \hat{s}_3 \in L \cdot \hat{s}_3 \cap \hat{\Sigma}_k^\varepsilon(\hat{\tau}_3)$,*

$$\widehat{\delta\theta}(\hat{s}_3) \geq c_{19} > 0 \tag{6.8}$$

independently of μ.

Proof. From the considerations in the proof of Lemma 6.3, the slope of the tangent vector to $L \cdot s$ at $\hat{y} \cdot s$ is greater than or equal to the slope of vector field there, for $s \in [s_0, \hat{s}_3]$. This implies that

$$\frac{\widehat{\delta z}}{\widehat{\delta\theta}} \geq \frac{z'}{\theta'}$$

or

$$\frac{\widehat{\delta\theta}'}{\widehat{\delta\theta}} \geq \frac{z'}{z}$$

and hence,

$$(\ln \widehat{\delta\theta})' \geq (\ln z)'. \tag{6.9}$$

It easily follows from integrating (6.9) that

$$\widehat{\delta\theta}(\hat{s}_3) \geq z(\hat{s}_3) \geq c_{19},$$

as needed. □

Lemma 6.4 has the important consequence that the curve $L \cdot s$ does not contract too much. We must also study the angle between the tangent vector and the vector field. We denote this angle $\phi(s)$ and prove the following Lemma; see Fig. 6.4.

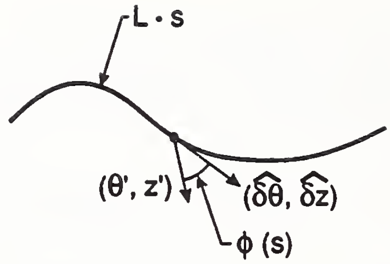

FIG. 6.4. *The tangent vector to $L \cdot s$ and angle made with the vector field.*

LEMMA 6.5. *At $\hat{y} \cdot \hat{s}_3 \in L \cdot \hat{s}_3 \cap \hat{\Sigma}_k^\epsilon(\tau_3)$, we have*

$$|\phi(\hat{s}_3)| \geq c_{20}\varepsilon^{3/2}. \tag{6.10}$$

Proof. Using the notation above, we can write

$$\Gamma = |(\theta', z')||(\widehat{\delta\theta}, \widehat{\delta z})| \sin \phi,$$

each piece being a function of s. From Lemma 6.4, it suffices then to prove the estimate (6.10) for ϕ replaced by Γ.

If $\widehat{\delta z}(\hat{s}_3) > 0$, the Lemma easily holds. Thus we can assume that there is a $s_0 < s_\diamond < \hat{s}_3$ for which $\widehat{\delta z}(s_\diamond) = 0$ and $|s_\diamond - \hat{s}_3| \geq T > 0$. But then in $[s_\diamond, \hat{s}_3]$, $\widehat{\delta\theta}(s)$ is decreasing and $\widehat{\delta\theta}(s) \geq \widehat{\delta\theta}(\hat{s}_3)$. From (6.7), we then have

$$\Gamma' \geq -\tfrac{1}{4}\sqrt{\varepsilon}\Gamma - \varepsilon^{3/2}\widehat{\delta\theta}(\hat{s}_3)B'(\tau), \tag{6.11}$$

and the Lemma easily follows by integrating (6.11) from $s_\diamond$ to $\hat{s}_3$. □

In $\tau = \hat{\tau}_3$, we study the piece of $L_\theta \cdot \hat{s}_3$ near the intersection with $W_{k+1}^{s,L}(\hat{\tau}_3)$. Since at each point $u \in L \cdot \hat{s}_3 \cap \Sigma_k^\epsilon(\hat{\tau}_3)$, the curve $L \cdot \hat{s}_3$ makes an angle of at least $c_{20}\varepsilon^{3/2}$ (by Lemma 6.5) with the vector field and the vector field makes an angle $O(\varepsilon^2)$ in $\Sigma_k^\epsilon(\hat{\tau}_3)$ with the vertical (a fixed), the curve $L \cdot \hat{s}_3$ is locally given by the graph of a function $b = g(a)$. Moreover it must traverse from one vertical side of $\Sigma_k^\epsilon(\hat{\tau}_3)$ to the other (by an application of the Mean Value Theorem). This procedure also applies in the $\tau = \tau_3$ slice as the angle does not decrease between $\hat{\tau}_3$ and τ (nor does $\widehat{\delta\theta}(s)$). Noting that this argument applies when any point $\hat{y} \in L \cdot s_3 \cap \Sigma_k^\epsilon(\tau_3)$, we can conclude that such a point lies on the curve in question, as it would otherwise force another intersection with $W_{k+1}^{s,L}$, contradicting Lemma 6.3.

We now consider how much of $L \cdot s_3 \cap \Sigma_k^\epsilon(\tau_3)$ remains in $\mathcal{B}_k(\tau)$ when $\tau = \tau_4$. If a point has left $\mathcal{B}_k(\tau)$ at $\tau = \tau_4$ (chosen so that $\tau_3 < \tau_4 < \tau_5$ but bounded away from τ_5 by $O(1)$ in ρ-variables), then by applying Lemma 5.3 it must be captured. It is clear that the set of points in $\Sigma_k^\epsilon(\tau_3)$ which stay in $\mathcal{B}_k(\tau)$ up to $\tau = \tau_4$ have $|a| \leq c_{21}\exp(-c_{22}/\varepsilon)$. Set

$$\Sigma_k^\gamma(\tau_3) = \{(a, b) \in \mathcal{B}_k(\tau_3) : |a| \leq c_{21}\exp(-c_{22}/\varepsilon)\}.$$

As above, $L \cdot s_3$ traverses $\Sigma_k^\gamma(\tau_3)$. Let (θ, z_0) and (θ, z_1) be the points in L for which

$$(\theta, z_0) \cdot s_3 \in \Sigma_k^\gamma(\tau_3) \cap \{a = -c_{21} \exp(-c_{22}/\varepsilon)\}$$

and

$$(\theta, z_1) \cdot s_3 \in \Sigma_k^\gamma(\tau_3) \cap \{a = c_{21} \exp(-c_{22}/\varepsilon)\}.$$

The length of the arc $L \cdot s_3 \cap \Sigma_k^\gamma(\tau_3)$ is given by

$$\alpha = \int_{z_0}^{z_1} \left([\widehat{\delta\theta}(s_3)]^2 + [\widehat{\delta z}(s_3)]^2\right)^{1/2} dz$$

where $(\widehat{\delta\theta}(s), \widehat{\delta z}(s), 0)$ is the tangent vector to $L \cdot s$ starting at $(0,1,0)$ and based at $(\theta, z, 0)$. By Lemma 6.4 ,

$$\alpha \geq c_{23}(z_1 - z_0).$$

But by Lemma 6.5 and the width of $\Sigma_k^\gamma(\tau_3)$,

$$\alpha \leq c_{24}\varepsilon^{3/2} \exp(-c_{25}/\varepsilon),$$

and we conclude that

$$|z_1 - z_0| \leq c_{26} \exp(-c_{27}/\varepsilon). \tag{6.12}$$

Eq. (6.12) gives an estimate on the amount of L that is not eventually captured by virtue of staying in $\mathcal{B}_k(\tau)$. Since this occurs in, at most, $O(1/\varepsilon^{3/2})$ $\mathcal{B}_k$'s, the Theorem follows. □

Acknowledgments. We would like to thank M. Levi for helpful suggestions at various stages of the development of this work; B. R. Miller for assistance with elliptic integrals; and A. I. Neishtadt and T. Kaper for several helpful conversations related to this work.

REFERENCES

[1] A. A. Andronov, A. A. Witt, and S. E. Khaikin, *Theory of Oscillators*, Pergamon Press, Oxford, 1966.

[2] V. I. Arnold, *Conditions for the applicability, and estimate of the error, of an averaging method for systems which pass through states of resonance in the course of their evolution*, Sov. Math. Dokl., 6 (1965), pp. 331–334.

[3] V. I. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations*, 2nd. Ed., Springer-Verlag, New York, 1988.

[4] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt, *Mathematical Aspects of Classical and Celestial Mechanics*, in Dynamical Systems III, Encyclopaedia of Mathematical Sciences, Vol. 3, Ed. by V. I. Arnold, Springer-Verlag, New York, 1988.

[5] G. Birkhoff and G.-C. Rota, *Ordinary Differential Equations*, 2nd Ed., Blaisdell, Waltham MA, 1969.

[6] T. J. Burns, *On the rotation of Mercury*, Celestial Mech., 19 (1979), pp. 297–313.
[7] P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Physicists*, Springer-Verlag, Berlin, 1954.
[8] N. Fenichel, *Persistence and smoothness of invariant manifolds for flows*, Indiana Univ. Math. J., 21 (1971), pp. 193–226.
[9] N. Fenichel, *Geometric singular perturbation theory for ordinary differential equations*, J. Diff. Eq., 31 (1979), pp. 53–98.
[10] P. Goldreich and S. J. Peale, *Spin-orbit coupling in the Solar System*, Astron. J., 71 (1966), pp. 425–438.
[11] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York, 1983.
[12] J. K. Hale, *Ordinary Differential Equations*, Wiley-Interscience, New York, 1969.
[13] P. Hartman, *Ordinary Differential Equations*, John Wiley & Sons, New York, 1964.
[14] C. K. R. T. Jones and N. Kopell, *Tracking invariant manifolds with differential forms in singularly perturbed systems*, J. Diff. Eq., to be published.
[15] T. J. Kaper and S. Wiggins, *Lobe area in adiabatic Hamiltonian systems*, Physica D, 51 (1991), pp. 205–212.
[16] K. P. Klaasen, *Mercury's rotation axis and period*, Icarus, 28 (1976), pp. 469–478.
[17] W. T. Kyner, *Passage through resonance*, in *Periodic Orbits, Stability and Resonances*, G. E. O. Giacaglia (ed.), D. Reidel Publishing Co., Dordrecht, Holland, 1970, pp. 501–514.
[18] P. Lochak and C. Meunier, *Multiphase Averaging for Classical Systems*, Springer-Verlag, New York, 1988.
[19] A. I. Neishtadt, *Passage through a separatrix in a resonance problem with a slowly-varying parameter*, J. Appl. Math. Mech., 39 (1975), pp. 594–605.
[20] A. I. Neishtadt, *Passage through resonance in a two-frequency problem*, Sov. Phys. Dokl., 20 (1977), pp. 189–191.
[21] A. I. Neishtadt, *Probability phenomena due to separatrix crossing*, CHAOS, 1 (1991), pp. 42–48.
[22] G. Sansone and R. Conti, *Non-Linear Differential Equations*, Pergamon Press, New York, 1964.
[23] S. Wiggins, *Chaotic Transport in Dynamical Systems*, Springer-Verlag, New York, 1991.