Chaotic Attractors of Relaxation Oscillators

John Guckenheimer[1], Martin Wechselberger[2] and Lai-Sang Young[3]

[1]Mathematics Department, Cornell University
[2]School of Mathematics and Statistics, University of Sydney
[3]Courant Institute of Mathematical Sciences, New York University

E-mail: `jmg16@cornell.edu`

Abstract. We develop a general technique for proving the existence of chaotic attractors for three dimensional vector fields with two time scales. Our results connect two important areas of dynamical systems: the theory of chaotic attractors for discrete two-dimensional Henon-like maps and geometric singular perturbation theory. Two dimensional Henon-like maps are diffeomorphisms that limit on non-invertible one dimensional maps. Wang and Young formulated hypotheses that suffice to prove the existence of chaotic attractors in these families. Three dimensional singularly perturbed vector fields have return maps that also are two dimensional diffeomorphisms limiting on one dimensional maps. We describe a generic mechanism that produces folds in these return maps and demonstrate that the Wang–Young hypotheses are satisfied. Our analysis requires a careful study of the convergence of the return maps to their singular limits in the C^k topology for $k \geq 3$. The theoretical results are illustrated with a numerical study of a variant of the forced van der Pol oscillator.

AMS classification scheme numbers: 34C26,34E15,37D45,37E10

Submitted to: *Nonlinearity*

1. Introduction

The discovery of chaotic attractors for low dimensional dynamical systems was a major achievement of dynamical systems theory during the twentieth century. There are many numerical simulations and observations that suggest concrete systems of differential equations have chaotic attractors, but there are few analytical results establishing their existence.‡ The mathematical theory is most tractable for uniformly hyperbolic attractors, but typical numerical examples arising from applications are not uniformly hyperbolic. This paper addresses the problem of locating chaotic attractors in specific families of differential equations by connecting two substantial theories that have been developed recently: (1) the study of Henon-like families of planar diffeomorphisms and

‡ The Lorenz system [16] is a notable exception in which theory and simulation have been connected with verified computation in the study of differential equation attractors.

(2) geometric singular perturbation theory. We show how Henon-like families arise in a generic way within the context of periodically forced relaxation oscillations. Relaxation oscillations were introduced and studied by van der Pol [21] in the 1920's and continue to be used as models for diverse phenomena. Thus this work gives a mathematical analysis of how chaotic attractors arise in the context of a familiar class of models for physical systems.

The study of Henon-like families of planar diffeomorphisms with strong contraction is framed in terms of perturbations from one dimensional mappings with folds. While still complicated, one dimensional theory is considerably better understood than two dimensional theory. Wang and Young [23, 24, 25, 26] have formulated a set of geometric hypotheses that suffice to prove the existence of chaotic attractors with Henon-like characteristics in families of strongly dissipative maps. Their hypotheses relate largely to properties of the limiting one dimensional family.

The collapse of two dimensional diffeomorphisms to one dimensional mappings is a phenomenon that occurs naturally in the context of *slow-fast* systems of differential equations. These singular perturbation problems are systems of differential equations of the form

$$\begin{aligned} \varepsilon\dot{x} &= f(x,z) \\ \dot{z} &= g(x,z) \end{aligned} \tag{1}$$

where $\varepsilon \geq 0$ is a small parameter determining the ratio of time scales. In this paper, $x \in \mathbb{R}$ and $z \in S^1 \times \mathbb{R}$; f and g are C^∞ functions. The singular limit $\varepsilon = 0$ gives a system of differential algebraic equations in which motion is constrained to the *critical manifold* $f = 0$. However, to represent fully the behavior of system (1) in the singular limit, we must allow "jumps" of trajectories from one sheet of the critical manifold to another that follow the direction of trajectories when $\varepsilon > 0$. In the examples we study, jumps parallel to the x-axis occur at folds, where the tangent plane to the critical manifold includes this direction. As we will show, the geometric picture above gives rise to a singular limit that reproduces the setting studied in the theory of Henon-like maps, in that return maps for the flow are two dimensional diffeomorphisms for $\varepsilon > 0$ that converge to one-dimensional maps for $\varepsilon = 0$. Moreover, the limit maps can have critical points, a central feature of the theory of Henon-like maps.

This paper is organized as follows: In Sections 2 and 3, we identify some properties of the system in (1) that lead potentially to Henon-like attractors. In Section 4, the precise conditions in Wang and Young [23, 24, 25, 26] are reviewed. In Section 5, we verify these conditions for a specific family of forced relaxation oscillations. Part of this verification is numerical and is non-rigorous.

2. A Class of Forced Relaxation Oscillators

This section introduces the class of dynamical systems that we study. They are slow-fast systems of the form

$$\begin{aligned} \varepsilon\dot{x} &= f(x,y,\theta) \\ \dot{y} &= g(x,y,\theta) \\ \dot{\theta} &= \omega \end{aligned} \tag{2}$$

where $(x,y,\theta) \in \mathbb{R} \times \mathbb{R} \times S^1$, f and g are C^∞ functions, $\omega > 0$ is the slow driving frequency, and $\varepsilon \ll 1$ is the singular perturbation parameter. A few geometric assumptions will be imposed on this family; they are described below.

Our first assumption is that the critical manifold S, defined by $f = 0$, is a "cubic" shaped surface with a pair of folds (see Figure 1):

Assumption 1 *The critical manifold* $S = S_a^- \cup L^- \cup S_r \cup L^+ \cup S_a^+$ *where* $S_a^+ \cup S_a^- := \{(x,y,\theta) \in S : f_x(x,y,\theta) < 0\}$ *are attracting upper and lower branches,* $S_r := \{(x,y,\theta) \in S : f_x(x,y,\theta) > 0\}$ *is a repelling branch, and* $L^+ \cup L^- := \{(x,y,\theta) \in S : f_x(x,y,\theta) = 0\}$ *are fold-curves (circles). Furthermore, we assume* $f_{xx}(x,y,\theta) \neq 0$ *on* $L^\pm$*, and* $f_y(x,y,\theta) \neq 0$ *on all of* S*. The latter implies that the cubic-shaped manifold* S *is the graph of a function* $y = \varphi(x,\theta)$, $\theta \in S^1$ *and* $x \in \mathbb{R}$.

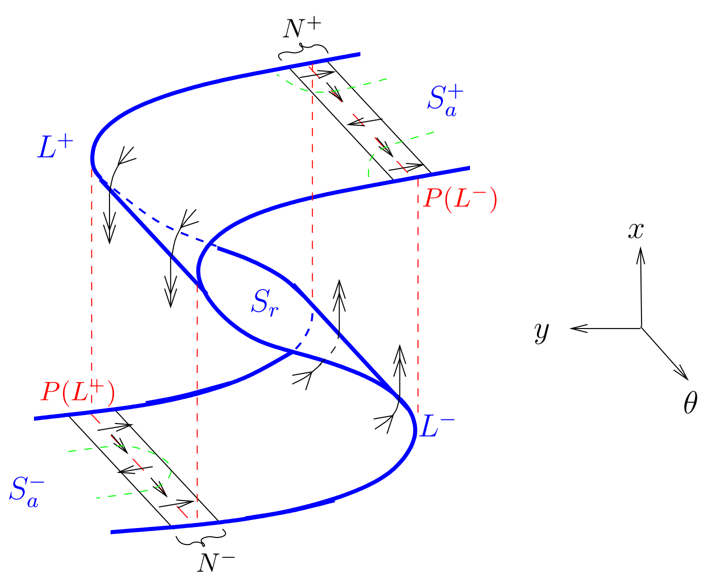

Figure 1. The critical manifold S and fold-curves $L^\pm$.

We use the representation $y = \varphi(x,\theta)$ to define a projection of the system (2) onto S. Differentiating $y = \varphi(x,\theta)$ with respect to time and using the implicit function theorem gives the relationship $f_x\dot{x} = -(f_y g + f_\theta\omega)$. This equation for $\dot{x}$ is singular on the folds, so we rescale the system to obtain the *reduced flow* on the critical manifold:

$$\begin{aligned} \dot{x} &= (f_y g + f_\theta\omega) \\ \dot{\theta} &= -\omega f_x \end{aligned} \tag{3}$$

This system has the same phase portrait as the singular limit of (2), but the orientation of trajectories is reversed on S_r.

Our second assumption is that all trajectories flow into the folds for the slow flow:

Assumption 2 *All points $p \in L^{\pm}$ of the fold-curves $L^{\pm}$ are jump points, i.e. the normal switching condition [15]*

$$(f_y\, g + f_\theta\, \omega)|_{p \in L^{\pm}} \neq 0 \tag{4}$$

is satisfied and the reduced flow near the fold-curves $L^{\pm}$ is directed toward the fold-curves $L^{\pm}$.

In the singular limit of system of (2), trajectories arrive at the folds from both S_a and S_r and the existence of trajectories breaks down. When $\varepsilon > 0$ is small, the trajectories of (2) execute fast "jumps" when they reach the vicinity of the fold-curves. We would like to capture this behavior and extend the definition of the slow flow from the critical manifold to all of $\mathbb{R} \times \mathbb{R} \times S^1$ in a way that embodies the limiting behavior of trajectories for (2). On the complement of S, the limiting direction of (2) is parallel to the x-axis. We define the map $P : \mathbb{R} \times \mathbb{R} \times S^1 - S \to S_a$ by projection along the x-axis. For many points $z \in \mathbb{R} \times \mathbb{R} \times S^1 - S$, the line parallel to the x-axis through z meets S_a in exactly one point, so there is no ambiguity in how $P(z)$ is defined. Where the line above meets S_a twice, P is defined so that the segment joining z and $P(z)$ does not intersect S_r. Thus z and $P(z)$ lie on the same side of S_r on a line parallel to the x-axis. Note that $L^{\pm}$ is not regarded as part of S_a, and $P(L^{\pm}) \subset S_a^{\mp}$ (see Figure 1). Heuristically, the fast trajectory segments that connect points to their images under P are instantaneous on the slow time scale. More formally, we use Benoit's concept of *candidates* [4, 20] to define trajectories of the reduced system.

Definition 1 *A trajectory of the reduced system consists of a continuous curve γ of the form $\gamma_0 \cup \alpha_1 \cup \beta_1 \cup \alpha_2 \cup \beta_2 \cup \cdots$ where*

- *γ_0 is a segment (perhaps trivial) of a line parallel to the x-axis that does not intersect S_r and terminates on S_a.*
- *α_i is a trajectory of the slow flow on the critical manifold terminating at a fold-curve $L^{\pm}$.*
- *β_i is a segment of a line parallel to the x-axis connecting a point on $L^{\pm}$ to a point of the critical manifold $S_a^{\mp}$.*

It is readily seen that there is a unique trajectory of the reduced system from any point in $\mathbb{R} \times \mathbb{R} \times S^1 - S$. Points off the critical manifold S move to S along γ_0. Points on S follow the slow flow until they reach the fold-curves. When they do, they jump along a segment β_i to the opposite sheet of S_a. While each point is the initial point of a unique trajectory, the trajectories depend discontinuously on initial conditions lying in S_r, reflecting the fact that trajectories of (2) separate on the fast time scale. The termination points of the curves β_i will play an important role in our analysis.

The third assumption we make about the slow flow is that all of the trajectories with initial conditions in neighborhoods of $P(L^{\pm}) \subset S_a^{\mp}$ reach the fold-curves.

Assumption 3 *There exist neighborhoods $N^{\pm} \subset S_a^{\pm}$ of $P(L^{\mp})$ with the property that all trajectories of the slow flow with initial conditions in $N^{\pm} \subset S_a^{\pm}$ reach the fold-curve $L^{\pm}$ (in finite time). The associated maps $P(L^{\mp}) \subset N^{\pm} \mapsto L^{\pm}$ are well defined and are surjective.*

Note that Assumption 3 implies that there are no equilibrium points of the slow flow on $S_a^{\pm}$ between $N^{\pm}$ and the fold-curves and that Assumption 2 implies that there are no equilibrium points on the fold-curves.§ Figure 1 depicts a reduced system that satisfies Assumptions 1-3.

To analyze how the dynamics of (2) approach those of the reduced system as $\varepsilon \to 0$, we introduce Poincaré return maps for the two systems. We want cross-sections that remain uniformly transverse to the reduced system as $\varepsilon \to 0$. Since the vector field points in opposite directions along the x-axis on the two sides of the critical manifold, we pick cross-sections that do not intersect S. Specifically, we choose two cross sections $\Sigma^{\pm}$ orthogonal to the x-axis, so that when ε is sufficiently small, all trajectories which leave from L^+ intersect Σ^+ before reaching S_a^- and all trajectories which leave from L^- intersect Σ^- before reaching S_a^+. We assume further that these cross-sections have distances that are $O(1)$ from the critical manifold S (see Figure 2). For $\varepsilon > 0$, we define the maps $H_\varepsilon^+ : \Sigma^+ \to \Sigma^-$ and $H_\varepsilon^- : \Sigma^- \to \Sigma^+$ as the transition maps following trajectories of the ε-flow, that is to say, $H_\varepsilon^{\pm}(z)$ is the first point of intersection of the trajectory of the ε-flow with initial condition z with the cross-section $\Sigma^{\mp}$. The corresponding singular limit maps $H_0^{\pm}$ are defined on $\Sigma^{\pm}$ by jumping to the slow manifold S, i.e. applying the map P, then following trajectories of the reduced system, and so on. The Poincaré return map from the cross-section Σ^- to itself is given by $\Pi_\varepsilon = H_\varepsilon^- \circ H_\varepsilon^+$, $\varepsilon \geq 0$. We will omit the subscript ε in $H_\varepsilon^{\pm}$ and Π_ε in statements that apply to all $\varepsilon \geq 0$.

Theorem 1 *[15, 20] Consider system (2) satisfying Assumptions 1-3. Then the Poincaré map $\Pi : \Sigma^- \to \Sigma^-$ induced by the flow of system (2) on a suitable transverse section Σ^- to the fast vector field is well defined for sufficiently small ε. The map is given by*

$$\Pi \begin{pmatrix} y \\ \theta \end{pmatrix} = \begin{pmatrix} R(y,\theta,\varepsilon) \\ G(y,\theta,\varepsilon) \end{pmatrix}, \tag{5}$$

where $G(y,\theta,\varepsilon) = G_0(y,\theta) + O(\varepsilon^{2/3})$. There is a constant $c > 0$ so that $|R(y,\theta,\varepsilon)| < \exp(-c/\varepsilon)$ and the function $G_0(y,\theta)$ describes the return map induced by the reduced flow (3).

If Assumptions 1-3 are satisfied, the projections of the fold-curves $P(L^{\pm})$ can be represented as graphs $x = \psi_1^{\pm}(\theta)$, $\theta \in S^1$, for system (3). The reduced flow is transversal

§ Isolated points on the fold-curves which violate the normal switching condition (4) are called *folded singularities*. The folded singularities are equilibrium points of the slow flow (3). Canard trajectories that follow portions of the unstable critical manifold S_r emanate from generic folded saddles and folded nodes.

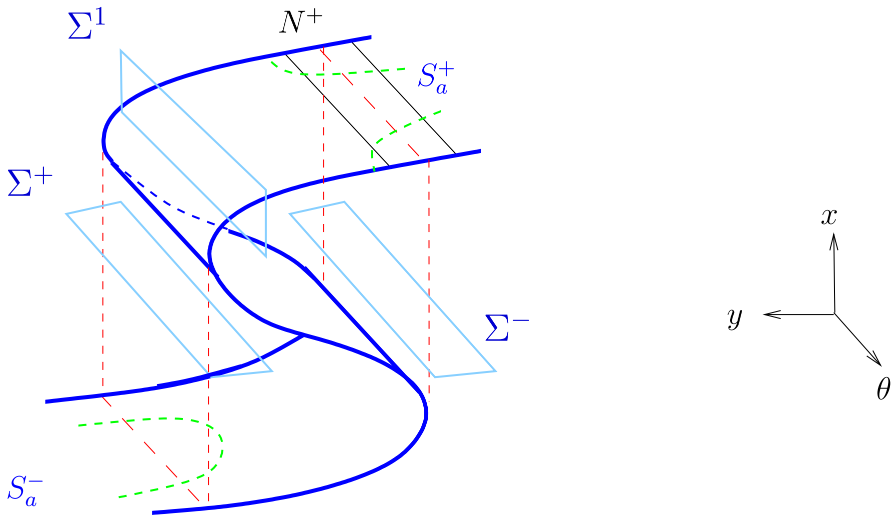

Figure 2. Poincaré section Σ^- of the map $\Pi : \Sigma^- \to \Sigma^-$ and auxiliary sections Σ^+ and Σ^1

to the projection-curves $P(L^\pm)$ if the (transversality) condition

$$l^*(p) := \begin{pmatrix} 1 \\ \psi_1' \end{pmatrix} \cdot \begin{pmatrix} f_y\, g + \omega\, f_\theta \\ -\omega \end{pmatrix}\Bigg|_{p\in P(L^\pm)} \neq 0 \tag{6}$$

is satisfied.

Systems for which the transversality condition (6) holds at all points of $P(L^\pm)$ were studied by Szmolyan and Wechselberger in [20]. They showed that the the derivative $\|(\partial/\partial\theta)G_0(y,\theta)\|$ has a positive lower bound (independent of ε) and that the Poincaré map of these systems possesses an invariant slow manifold with associated stable foliation:

Proposition 2.1 *Consider system (2) satisfying Assumptions 1-3. Further assume that (6) holds at all points of $P(L^\pm)$. Then system (2) possesses a unique invariant torus. The associated Poincaré map (5) possesses a unique invariant slow manifold with associated stable foliation.*

It follows from Proposition 2.1 that the dynamics of system (2) can be analyzed by studying the circle diffeomorphism $\theta \mapsto G_0(0,\theta)$. The dynamics of these driven relaxation oscillators is determined by the properties of these circle diffeomorphisms; in particular their attractors are periodic orbits or two dimensional quasi-periodic invariant tori. The forced van der Pol oscillator [6] with forcing amplitude $a < 1$ is a prominent example which possesses an invariant torus.

The dynamics of systems for which the transversality condition (6) fails at some points of $P(L^\pm)$ are more complicated. Assumption 3 allows trajectories of the reduced flow to intersect $P(L^\pm)$ more than once and to have tangencies with these curves, as long as they reach $L^\mp$ in finite time. In this case, as $\varepsilon \to 0$, the Poincaré map (5) converges to a circle map that may not be a diffeomorphism. This is precisely the type of behavior we analyze in this paper.

Assumption 4 *There exist isolated critical points $p_i^* \in P(L^\pm)$, $i \geq 1$, which violate the transversality condition (6), i.e.*

$$l^*(p_i^*) = 0\,. \tag{7}$$

Furthermore these critical points p_i^* are non-degenerate, i.e.

$$l^{*\prime}(p_i^*) := \frac{d}{d\theta} l^*(p_i^*) \neq 0\,. \tag{8}$$

The reduced flow is tangent to $P(L^\pm)$ at the critical points p_i^*. The non-degeneracy condition (8) guarantees that the circle map G_0 has turning points at trajectories that pass through p_i^*. Trajectories close to these turning points intersect $P(L^\pm)$ more than once. This produces folds in the Poincaré map Π_0; see Figures 3 and 5. For small $\varepsilon > 0$, the strong contraction of the flow near $S_a^\pm$ pulls the sheets of these folds "exponentially close" to each other. The folding of the Poincaré map Π_ε precludes the existence of an invariant manifold with associated stable foliation, as the orientation of strong stable manifolds is reversed on the two sheets on the opposite sides of a fold. Therefore system (2) under Assumptions 1-4 does not possess an invariant torus.

The folding behavior of the Poincaré map Π_ε is typical for Henon-like maps. That is why if Π_ε is sufficiently stretching in the θ-direction, one may expect to find horseshoes and strange attractors.

3. Convergence to Singular Limit

To apply the theory of Wang and Young [23, 24, 25] to the systems described in Section 2, we need to prove the convergence of Π_ε to Π_0 in the C^3 topology on the space of two dimensional maps. Since the asymptotic form of Π in Theorem 1 is singular in ε, this convergence is not apparent. We have been unsuccessful in locating a suitable reference for the convergence properties we require, even for the hyperbolic portion of the flow along the critical manifold.

Recall that $\Pi_\varepsilon = H_\varepsilon^+ \circ H_\varepsilon^-$ where $H_\varepsilon^- : \Sigma^- \to \Sigma^+$ and $H_\varepsilon^+ : \Sigma^+ \to \Sigma^-$. For concreteness, we consider H^-; the arguments for H^+ are identical. This section is devoted to proving the following theorem.

Theorem 2 *For any integer $k > 0$, the maps $H_\varepsilon^\pm$ converge to their rank one singular limits $H_0^\pm$ in the C^k topology as $\varepsilon \to 0$.*

We first choose normalizing coordinates $(u, v, \theta) = (\hat{u}(x, y, \theta), \hat{v}(y, \theta), \theta)$ near the fold-curve of system (2) so that L^+ becomes the curve $u = v = 0$ and the critical manifold near L^+ is given by $v = u^2$ [1]. After a time rescaling, system (2) has the representation

$$\begin{aligned} \varepsilon u' &= v - u^2 \\ v' &= \hat{f}(u, v, \theta, \varepsilon) \\ \theta' &= \hat{g}(u, v, \theta, \varepsilon) \end{aligned} \tag{9}$$

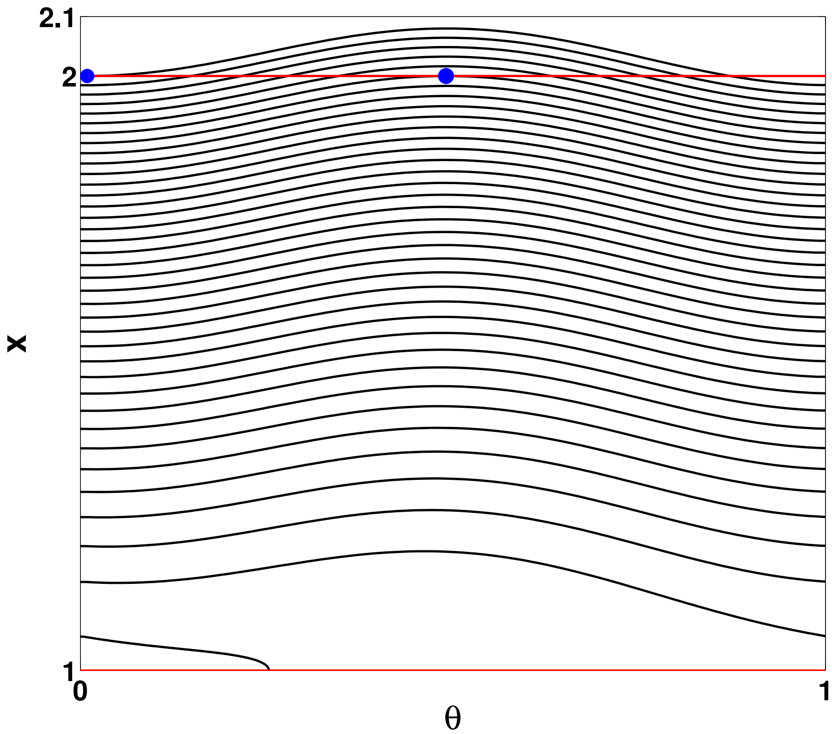

Figure 3. A trajectory of the slow flow for the example studied in Section 5. The slow flow on S_a^+ is projected onto the (θ, x) plane; θ is a periodic variable with period 1. The fold curve L^+ (see Figure 1) corresponds to the circle $x = 1$; $P(L^-)$ corresponds to $x = 2$. To understand the geometry of the map Π_0, we look at the Poincaré map $\hat{H}$ from $x = 2$ to $x = 1$ following the slow flow: $\hat{H}$ has two critical points located where the slow vector field is tangent to $x = 2$; they are marked by blue dots on the figure. The trajectory with initial condition at the left hand dot is plotted until it reaches $x = 1$. All points of intersection of this trajectory with $x = 2$ have the same image under $\hat{H}$; see Figure 5.

By virtue of the normal switching condition, we may assume $\hat{f}(0, 0, \theta, 0) < 0$ in this neighborhood. We will work in these coordinates for the rest of the proof.

To study the convergence properties of H_ε^-, we decompose the map into four segments by introducing three additional cross-sections near L^+:

- Σ^1 is defined by $v = \delta$ for a small number $\delta > 0$ independent of ε (see Figure 2);
- Σ^2 and Σ^3 are ε-dependent cross-sections just before and after the jump at L^+; they are defined by $v = c_2\varepsilon^{2/3}$, $c_2 > 0$, and $u = -c_3\varepsilon^{1/3}$, $c_3 > 0$, respectively.

The Poincaré map from Σ^- to Σ^1 is denoted by $H^{-,1}$; the one from Σ^1 to Σ^2 is denoted by $H^{1,2}$, and so on. We analyze each phase of the motion separately, using normalizing systems of coordinates adapted to the different phases.

To prove the C^k convergence of H_ε^-, one can prove the C^k convergence of each of the 4 maps to be composed. Alternately, one can prove (i) $H_\varepsilon^- \to H_0^-$ in the C^0

topology, and (ii) $\{H_\varepsilon^-, \varepsilon > 0\}$ is uniformly bounded in the C^{k+1} norm. This is sufficient because any C^{k+1} bounded set is compact in C^k, and from (i), every limit point of every subsequence of H_ε^- in the C^k metric must in fact be H_0^-. We will elaborate, but C^0 convergence is in fact relatively simple; it follows largely from Gronwall's inequality. Much of the work to follow is about C^k convergence or boundedness for $k \geq 1$.

I. The normally hyperbolic phase $H^{-,1} : \Sigma^- \to \Sigma^1$

Proposition 3.1 *For any integer $k \geq 0$, $H_\varepsilon^{-,1} \to H_0^{-,1}$ in the C^k topology as $\varepsilon \to 0$.*

Our starting point is the existence theorem for invariant slow manifolds of a slow-fast system and their stable and unstable manifolds, frequently referred to as "Fenichel theory" [14]. In the current context, this theory states the following.

Theorem 3 *Assume that system (2) satisfies Assumption 1. Let r be a positive integer and $\bar{S}_0 \subset S_a$ be a closed domain in the stable part of its critical manifold. Then there is a C^r family of manifolds S_ε defined for $\varepsilon \geq 0$ sufficiently small so that S_ε is an overflowing invariant manifold.*

Here, S_ε is a manifold with boundary and *overflowing* means that trajectories that enter or leave S_ε do so through its boundary. The set $\bar{S}_0$ relevant for the proof of Proposition 3.1 is the region in S_a from Σ^1 to N^+ (see Figure 2).

Using this theorem, we make an ε-dependent C^k coordinate change to $(w, v, \theta, \varepsilon) = (\hat{w}(u, v, \theta, \varepsilon), v, \theta, \varepsilon)$ so that in the relevant region the subspace $w = 0$ is an invariant slow manifold for all small $\varepsilon \geq 0$. This can be done by taking $\hat{w}(u, v, \theta, \varepsilon) = u - \gamma(v, \theta, \varepsilon)$ where the graph of $(v, \theta) \mapsto \gamma(v, \theta, \varepsilon)$ is S_ε. Clearly, Σ^1 is not affected by this change of coordinates. Using a bump function, it is easy to arrange it so that Σ^- is also unchanged. It suffices to prove the C^k convergence of $H_\varepsilon^{-,1}$ in these coordinates. The assertion in the proposition then follows since the slow manifolds of the ε-flow converge in C^k to S_a by Theorem 3.

Since the flow is normally hyperbolic along the portion of S_a^+ in question, we conclude that $\varepsilon \dot{w} = w h(w, z, \theta, \varepsilon)$ with $h < 0$. We may thus rescale time by $-h$ to obtain a system of the form

$$\begin{aligned} \varepsilon w' &= -w \\ v' &= \bar{f}(w, v, \theta, \varepsilon) \\ \theta' &= \bar{g}(w, v, \theta, \varepsilon) \end{aligned} \tag{10}$$

without affecting the map $H_\varepsilon^{-,1}$. Equation (10) yields

$$w(t) = w(0) \exp(-t/\varepsilon)$$

and a system of two equations not involving the w-coordinate (except for the appearance of $w(0)$ in the argument of $\bar{f}$ and $\bar{g}$):

$$\begin{aligned} v' &= \bar{f}(w(0) \exp(-t/\varepsilon), v, \theta, \varepsilon) \\ \theta' &= \bar{g}(w(0) \exp(-t/\varepsilon), v, \theta, \varepsilon) \end{aligned} \tag{11}$$

These equations make sense for $\varepsilon > 0$. We let $w(t,\varepsilon), v(t,\varepsilon)$ and $\theta(t,\varepsilon)$ denote their solutions. For $\varepsilon = 0$, solutions for $t > 0$ are defined for initial condition $z \in \Sigma^-$ as follows: $w(t,0) = 0$ for all $t > 0$, and $v(t,0)$ and $\theta(t,0)$ are defined using the reduced flow on S_a with $P(z)$ as initial condition. We claim that on any time interval bounded away from 0, $w(t,\varepsilon), v(t,\varepsilon)$ and $\theta(t,\varepsilon)$ converge uniformly in C^k to $w(t,0), v(t,0)$ and $\theta(t,0)$ respectively as $\varepsilon \to 0$. (This convergence is *not* uniform on time intervals of the form $(0,t_0)$ because of the jump in $w(\cdot,0)$ at $t = 0$.)

Let $t_1 > 0$ (to be thought of as roughly equal to the transition time from Σ^- to Σ^1) be fixed. A formal argument for the C^0 convergence of the time-t_1-map of (10) goes as follows. Let $z = (v,\theta) \in \Sigma^-$, and let $\zeta_\varepsilon(t)$ and $\eta_\varepsilon(t)$ denote the solutions of (10) and (11) with initial condition z. We will show that for any given $\alpha > 0$, $|\zeta_\varepsilon(t_1) - \zeta_0(t_1)| < \alpha$ for ε sufficiently small (independent of z). By Gronwall's inequality, there exist $\beta > 0$ and $\varepsilon(\beta) > 0$ such that for all $t_0 \in (0,t_1)$ and $\varepsilon < \varepsilon(\beta)$, if $|\zeta_\varepsilon(t_0) - \zeta_0(t_0)| < \beta$, then $|\zeta_\varepsilon(t_1) - \zeta_0(t_1)| < \alpha$. Next we choose $t_0 > 0$ small enough that for all small ε, $|\eta_\varepsilon(t_0) - \eta_0(t_0)| < \beta/2$. Here we have used the bounded C^0 norms of $\bar{f}$ and $\bar{g}$ in the relevant region of phase space to limit the total movement of η_ε in the time interval $[0,t_0]$. Finally, shrink ε if necessary so that $|\zeta_\varepsilon(t_0) - \zeta_0(t_0)| < \beta$.

To prove the C^k convergence of these time-t-maps, we differentiate (10) and (11) with respect to z to obtain a hierarchy of variational equations. These are again nonautonomous vector fields in which the time dependence is through terms of the form $\exp(-t/\varepsilon)$. Solutions of the variational equations give the derivatives of the flow map of the system (10) and (11). Therefore, the same estimates that we have used for the convergence of time-t-maps establish that the derivatives of these maps converge as $\varepsilon \to 0$.

To complete the proof of Proposition 3.1, it remains only to observe that the functions T_ε, where $T_\varepsilon(z)$ is the time it takes for the ε-flow to reach Σ^1 starting at $z \in \Sigma^-$, converge in C^k to T_0 as $\varepsilon \to 0$. This is a direct consequence of the argument above.

Remark. One might expect that difference between $H_\varepsilon^{-,1}$ and $H_0^{-,1}$ would be exponentially small, but this is not true, even if the only ε dependence of the system (10) appears through the terms $\exp(-t/\varepsilon)$. This is evident from simple examples such as

$$\begin{aligned} \varepsilon w' &= -w \\ z' &= 1 \\ \theta' &= w \end{aligned} \tag{12}$$

The solutions of system (12) are given by

$$(w(t), z(t), \theta(t)) = (w(0)\exp(-t/\varepsilon), z(0)+t, \theta(0)+\varepsilon w(0)(1-\exp(-t/\varepsilon)) \tag{13}$$

Thus the Poincaré map H_ε for the section from $w = 1$ to $z = 1$ is given by $H_\varepsilon(z,\theta) = (\exp(-(1-z)/\varepsilon), \theta + \varepsilon(1 - \exp(-(1-z)/\varepsilon))$ while $H_0(z,\theta) = (0,\theta)$. The difference between these two maps is $O(\varepsilon)$ but not $o(\varepsilon)$. This is due to the fact that the change in θ during the initial fast convergence to the slow manifold is $O(\varepsilon)$. Classical results of Vasil'eva [22] give asymptotic expansions in ε for the solutions of system (10).

We record the following in anticipation of the blow-up analysis to follow. Returning to (u, v, θ)-coordinates, let $u = R(v, \theta, \varepsilon)$ and $\theta = G(v, \theta, \varepsilon)$ be the two components of the map $H^{-,1}$. Let S_ε be the slow manifolds in Theorem 3, and note that $S_0 \cap \Sigma^1 = \{u = \sqrt{\delta}\}$.

Lemma 1 *(i)* $|R - \sqrt{\delta}| = O(\varepsilon)$;
(ii) $|\frac{\partial R}{\partial v}| = O(\varepsilon)$, $|\frac{\partial R}{\partial \theta}| = O(\varepsilon)$.

Proof: It follows from Theorem 3 that $S_\varepsilon \cap \Sigma^1$ is the graph of $u = \sqrt{\delta} + \varepsilon \bar{\gamma}(\theta, \varepsilon)$ where $\bar{\gamma}$ is uniformly bounded in C^k as a function of θ. Let $\hat{R}(v, \theta, \varepsilon)$ denote the w-component of $H^{-,1}$ in (w, v, θ)-coordinates. Then

$$R(v, \theta, \varepsilon) = \hat{R}(v, \theta, \varepsilon) + \sqrt{\delta} + \varepsilon \bar{\gamma}(G(v, \theta, \varepsilon), \varepsilon). \tag{14}$$

We have shown that $\hat{R}$ and its derivatives are $O(\exp(-c/\varepsilon))$. (i) follows immediately, and (ii) follows from the fact that $|\frac{\partial \tilde{\gamma}}{\partial \theta}|, |\frac{\partial G}{\partial v}|$, and $|\frac{\partial G}{\partial \theta}|$ are all $O(1)$. □

II. Blow-up at fold curve and $H^{2,3} : \Sigma^2 \to \Sigma^3$

We first deal with this crucial part, and adjust the rest of the analysis around it. To get past the fold-curve L^+, we use the blow-up in [20], which is carried out in new coordinates

$$E = \varepsilon^{1/3}, \quad EU = u, \quad E^2 V = v \quad \text{and} \quad \mathrm{E}^2\mathrm{T} = \mathrm{t}.$$

In these coordinates, $\Sigma^2 : V = c_2$ and $\Sigma^3 : U = -c_3$ are independent of ε, and the transformed system is given by

$$\begin{aligned} \frac{dU}{dT} &= V - U^2 \\ \frac{dV}{dT} &= \hat{f}(EU, E^2 V, \theta, E^3) \\ \frac{d\theta}{dT} &= E^2 \hat{g}(EU, E^2 V, \theta, E^3) \end{aligned} \tag{15}$$

Note that system (15) is a Riccati equation in (U, V)-space when $E = 0$. (See Figure 4.) This is a regularly perturbed system as $E \to 0$. Hence the Poincaré maps from Σ^2 to Σ^3 are smooth and vary smoothly with E.

Transitions between $(u, v, \theta, \varepsilon)$ and (U, V, θ, E)-coordinates are made via the blow-up map $\Phi^1_\varepsilon : \Sigma^1 \to \Sigma^1$ given by $\Phi^1_\varepsilon(u, \theta) = (\varepsilon^{-\frac{1}{3}} u, \theta) = (U, \theta)$ and the blow-down map $\Phi^+_\varepsilon : \Sigma^+ \to \Sigma^+$ given by $\Phi^+(V, \theta) = (E^2 V, \theta) = (v, \theta)$. That is to say, we view H^- as

$$H^- = \Phi^+ \circ \tilde{H}^{3,+} \circ \tilde{H}^{2,3} \circ \tilde{H}^{1,2} \circ \Phi^1 \circ H^{-,1} \tag{16}$$

where the ˜ in $\tilde{H}$ means the map in question is to be seen in (U, V, θ)-coordinates. Clearly, Φ^+ does no harm in terms of C^k boundedness or convergence. As for Φ^1, $\partial U / \partial u = \varepsilon^{-\frac{1}{3}}$, but observe that when differentiating $\Phi^1 \circ H^{-,1}$, $\partial U / \partial u$ is followed without exception

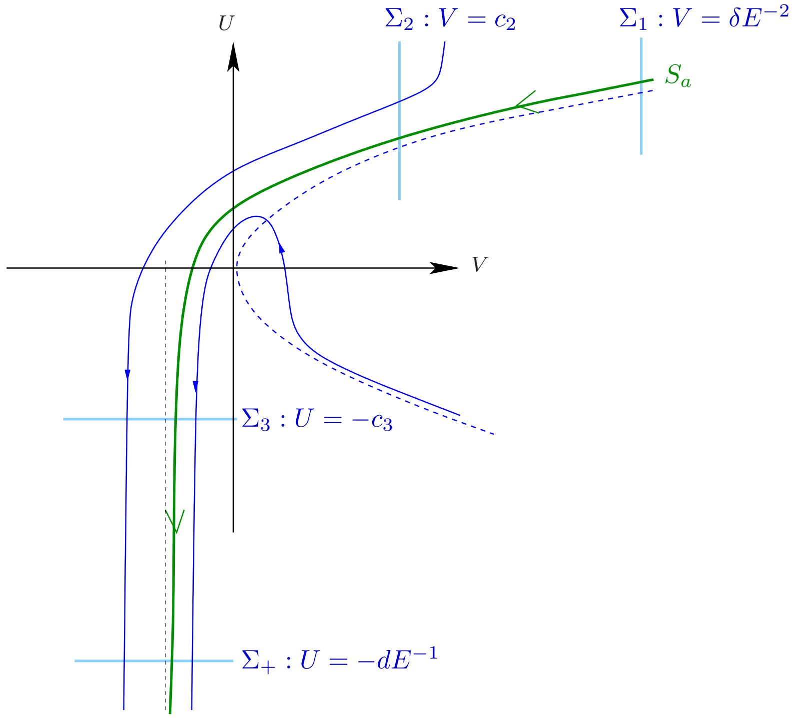

Figure 4. Sections in the rescaled system (15) in (U, V) phase space, which shows solutions of the Riccati equation (given by $E = 0$ in (15)). The special solution (green curve) of the Riccati equation represents the extension of the manifold S_a past the fold-curve

by either $\frac{\partial R}{\partial v}$ or $\frac{\partial R}{\partial \theta}$ where R is the u-component of $H^{-,1}$. By Lemma 1, both of these terms are $O(\varepsilon)$.

To summarize, we have proved, in the notation of (16), the C^k convergence of $\Phi^1 \circ H^{-,1}$ and $\tilde{H}^{2,3}$ for every integer $k > 0$.

III. Before the jump, i.e. $\tilde{H}^{1,2} : \Sigma^1 \to \Sigma^2$

This region is a transition between the two phases of dynamics studied in parts I and II. In (u, v, θ)-coordinates, the minimum angle between the fast direction and the slow manifold goes to 0 as $\varepsilon \to 0$, making it impossible to deduce uniform estimates from standard normal hyperbolic theory. In (U, V, θ)-coordinates, this lack of uniformity is exchanged for unbounded domains and unbounded times: Σ^1 is given by $V = \delta E^{-2}$, and the time T_0 it takes to go from Σ^1 to Σ^2 is $O(E^{-2})$.

Proposition 3.2 *The first k derivatives of $\tilde{H}^{1,2}$ are uniformly bounded for all $E > 0$.*

For simplicity we normalize the speed in the V-direction by rewriting (15) as

$$\frac{dU}{dT} = (V - U^2)X \ , \qquad \frac{dV}{dT} = -1 \ , \qquad \frac{d\theta}{dT} = E^2 Y \tag{17}$$

where $X = -\hat{f}$ and $Y = -\hat{f}/\hat{g}$. The map $\tilde{H}^{1,2}$ is then given by the time-T_0-map where $T_0 = \delta E^{-2} - c_2$. Note that X is positive and bounded away from 0. The next lemma describes the region in which all the action takes place.

Lemma 2 *There exists $C > 0$ such that the following holds for all $E > 0$: between Σ^1 and Σ^2, the U-coordinates of all trajectories starting from Σ^- satisfy $\sqrt{V} < U < \sqrt{V}+C$.*

Proof: By Lemma 1(i), all trajectories meet Σ^1 in $\{|U-\sqrt{V}| < O(E^2)\}$. That $U > \sqrt{V}$ is because $\{V = U^2\}$ bends downward while $dU/dT > 0$ if $U < \sqrt{V}$. Finally we claim that $\sqrt{V} < U < \sqrt{V} + C$ is a trapping region for $C > 0$ large enough, because at $U = \sqrt{V} + C$,

$$(V - U^2)X = -C(U + \sqrt{V})X < -\frac{1}{2\sqrt{V}}.$$

□

Let (ξ, η) denote the (U, θ)-coordinates of a point in Σ^1, and let $U(t) = U(\xi, \eta, t)$ and $\theta(t) = \theta(\xi, \eta, t)$ denote the solution of (17) with $U(\xi, \eta, 0) = \xi$ and $\theta(\xi, \eta, 0) = \eta$. Differentiating (17) and letting

$$\begin{aligned} h_1 &= -2UX + E(V - U^2)X_U, \\ h_2 &= (V - U^2)X_\theta, \end{aligned}$$

we obtain the first variational equations

$$\begin{array}{llll} (i) & \dfrac{dU_\xi}{dT} = h_1 U_\xi + h_2\theta_\xi, & & U_\xi(0) = 1 \\ (ii) & \dfrac{d\theta_\xi}{dT} = E^3 Y_U U_\xi + E^2 Y_\theta \theta_\xi, & & \theta_\xi(0) = 0 \\ (iii) & \dfrac{dU_\eta}{dT} = h_1 U_\eta + h_2\theta_\eta, & & U_\xi(0) = 0 \\ (iv) & \dfrac{d\theta_\eta}{dT} = E^3 Y_U U_\eta + E^2 Y_\theta \theta_\eta, & & \theta_\eta(0) = 1. \end{array}$$

We claim that on the time interval of interest, $|U_\xi|, |U_\eta|, |\theta_\xi|, |\theta_\eta| = O(1)$ (in fact, $|\theta_\xi| = O(E)$). In the relevant region of phase space, $|V - U^2| \leq \text{const}\cdot U$ (Lemma 2), so h_1 is dominated by the term $-2UX$. To estimate $|U_\eta|$ and $|\theta_\eta|$, for example, suppose $|\theta_\eta(s)| = O(1)$ for all $s < t$. Then $|U_\eta(t)| = O(1)$ because if $|U_\eta(s)|$ is large enough, then the first term on the right side of (iii) will dominate the second. On the other hand, (iv) says that if $|U_\eta(s)| = O(1)$ for all $s < t$, then $|\theta_\eta(t)|$ can change at a maximum rate of $O(E^2)$.

Higher order variational equations are obtained by differentiating (i)–(iv). Let α denote a j-tuple of ξ and η, $1 < j \leq k$, and let $\partial^\alpha U$ and $\partial^\alpha\theta$ denote the corresponding partials of U and θ. Proposition 3.2 asserts that $\partial^\alpha U$ and $\partial^\alpha\theta$ are uniformly bounded for all $E > 0$. Observe that for any α, $\partial^\alpha U$ and $\partial^\alpha\theta$ satisfy equations having the same form as (i)–(iv) with additional terms in (known) functions of t:

$$\begin{array}{ll} \dfrac{d(\partial^\alpha U)}{dT} = h_1\partial^\alpha U + h_2\partial^\alpha\theta + \{\ \cdots\ \}, & \partial^\alpha U(0) = 0 \\ \dfrac{d(\partial^\alpha \theta)}{dT} = E^3 Y_U\partial^\alpha U + E^2 Y_\theta\partial^\alpha\theta + \{\ \cdots\ \}, & \partial^\alpha\theta(0) = 0. \end{array}$$

Here the expressions inside the brackets are sums of terms that are products of (i) partials of X and Y and (ii) partials of U, θ, h_1 and h_2 of order $< |\alpha|$. Partials of the kind in (i) are bounded by definition. Those in (ii) are shown inductively to be bounded, except for partials of h_2, which contain terms of size $\leq$ const$\cdot U$ (e.g. $\partial h_2/\partial\eta = -2UU_\eta X_\theta + (V-U^2)X_{\theta U}EU_\eta + (V-U^2)X_{\theta\theta}\theta_\eta$.) As explained earlier, for equations of the form above, as long as all terms are $\leq$ const$\cdot U$, the right side is dominated by $h_1\partial^\alpha U$ once $\partial^\alpha U$ becomes sufficiently large.

IV. After the jump, i.e. $\tilde{H}^{3,+}:\Sigma^3\to\Sigma^+$

We make a final coordinate transformation by setting $W = U/(1-U)$ to obtain

$$\begin{aligned}
W' &= (1+W)^2V - W^2\\
V' &= \hat{f}(EW/(1+W), E^2V, \theta, E^3)\\
\theta' &= E^2\hat{g}(EW/(1+W), E^2V, \theta, E^3)
\end{aligned} \tag{18}$$

The section $\Sigma^3 : U = -c_3$ is now given by $W = -c_3/(1+c_3) < 0$ with $-1 < -c_3/(1+c_3) < 0$, and the section $\Sigma^+ : U = -dE^{-1}$ is given by $W = -d/(d-E)$ which gives in the limit $E \to 0$ the section $W = -1$. System (18) varies smoothly with E in the region between the cross-sections Σ^3 and Σ^+. It is also apparent that $W' < 0$ throughout this region (equivalent to $V < U^2$ in (15)). Therefore, the flow from Σ^3 to Σ^+ requires finite time, so smooth dependence on initial conditions for solutions to systems of ordinary differential equations implies that $\tilde{H}^{3,+}$ converges in the C^k topology as $E \to 0$. These maps are unaffected as we transform back to (U, V, θ)-coordinates.

To recapitulate, we have proved C^k convergence (for arbitrary k) in I,II and IV, and C^k boundedness in III. The C^0 convergence of H_ε^- from Σ^1 to Σ^2 is trivial in (u, v, θ)-coordinates. The proof of Theorem 2 is therefore complete.

4. Henon-like Maps

The existence of chaotic attractors in mappings that are not uniformly hyperbolic has been investigated intensively for over twenty years, beginning with numerical investigations of Flaherty and Hoppensteadt [7], the pioneering work of Hénon [12], and the work of Jakobson [13] on quadratic maps of the interval. Henon [12] illustrated the existence of chaotic attractors in a two dimensional diffeomorphism of the plane defined by quadratic functions. Benedicks and Carleson [2] developed powerful techniques for analyzing Henon maps close to the singular limit (in which the map reduces to a map of the interval). Benedicks and Young [3] constructed SRB measures on these attractors. The results of [2] were extended by Mora and Viana [18], and further extended and refined by Wang and Young [23, 24, 25, 26]. One of the contributions of Wang and Young was to replace the formula of the Hénon maps by a list of concrete, geometric properties that are sufficient to prove the existence of chaotic attractors with SRB measures and positive Lyapunov exponents. This section reviews these geometric conditions and their implications. In the next section, we will discuss these conditions in the context of return maps for a family of forced relaxation oscillations.

Let $M = S^1 \times I$ where $I \subset \mathbb{R}$ is a closed interval. Coordinates in M are denoted by (θ, y). We consider a family of maps $H_{a,\varepsilon} : M \to M$ parametrized by $a \in [a_0, a_1] \subset \mathbb{R}$ and $\varepsilon \in (0, \varepsilon_0] \subset \mathbb{R}$. Conditions (C0) and (C1) below give the overall setup, describing $H_{a,\varepsilon}$ as small perturbations of rank one maps.

(C0) Regularity Conditions

- For each $\varepsilon > 0$, the function $(\theta, y, a) \to H_{a,\varepsilon}(\theta, y)$ is C^3.
- Each $H_{a,\varepsilon}$ is an embedding of M into itself.

(C1) The singular limit There exist rank one maps $H_{a,0} : M \to M$ such that as functions of (θ, y, a), $H_{a,\varepsilon}(\theta, y)$ converge in the C^3 norm to $H_{a,0}(\theta, y)$ as $\varepsilon \to 0$. The image sets of M under $H_{a,0}$, i.e. $H_{a,0}(M)$, are assumed to be diffeomorphic to S^1.

Via small changes of coordinates, we may assume $H_{a,0}(M)$ is independent of a. We denote this set by γ and regard $H_{a,0}$ restricted to γ as a map of S^1 to itself, denoted by h_a. The rest of the conditions involve only the singular limit maps $H_{a,0}$ and h_a. Expansion on the attractor is derived from the corresponding properties of h_a, which we formulate as (C2).

(C2) Expansion in 1D maps There exists $a^* \in [a_0, a_1]$ so that $h_{a^*} = h$ has the following properties. There are $c > 1$, $N \in \mathbb{Z}^+$, and a neighborhood I of the critical set C (the set where $h' = 0$) in γ such that

- if $\xi, h(\xi), \cdots, h^{n-1}(\xi) \notin I$ and $h^n(\xi) \in I$, then $(h^n)'(\xi) \geq c^n$.
- if $\xi, h(\xi), \cdots, h^{n-1}(\xi) \notin I$ and $n \geq N$, then $(h^n)'(\xi) \geq c^n$.
- if $\xi \in I$ is not a critical point, there is $n = n(\xi)$ such that $h(\xi), \cdots, h^{n-1}(\xi) \notin I$ and $(h^n)'(\xi) \geq c^n$.
- $\min(|h''|) > 0$ on I
- if ξ is a critical point, then $h^n(\xi) \notin I$ for all $n > 0$.

The next two conditions ensure the absence of "coincidences" that may obstruct the analysis. The first relies upon the concept of smooth continuations. Since we assume that the critical points of h_{a^*} are non-degenerate, they vary smoothly with a near a^*. Also, points p whose h_{a^*} trajectories avoid the set I have unique continuations as curves of points with a constant symbolic itinerary in $\gamma \setminus I$.

(C3) Parameter transversality For each critical point ξ of h_{a^*}, let $p = h_{a^*}(\xi)$, and let ξ_a and p_a denote the continuations of ξ and p. Then

$$\frac{d}{da} h_a(\xi_a) \neq \frac{d}{da} p_a \qquad \text{at } a = a^*$$

(C4) Nondegeneracy at "turns" At each critical point ξ of h_{a^*},

$$\frac{\partial}{\partial y} H_{a^*,0}(\xi, 0) \neq 0$$

Finally, we include a condition used to deduce additional mixing properties.

(C5) Mixing

- The constant c in (C2) is larger than 2
- Let $J_1, \cdots, J_r$ be intervals of monotonicity of h_{a^*} and let $P = (p_{i,j})$ be the 0-1 matrix with $p_{i,j} = 1$ if and only if $J_j \subset h_{a^*}(J_i)$. Then P is power positive; i.e., there is an n such that all entries of P^n are positive.

Wang and Young proved that families $H_{a,\varepsilon} : M \to M$ satisfying (C0)–(C5) have chaotic attractors with strong stochastic properties. This is expressed through the idea of SRB measures. We review the definition and implications of this important idea:

For $H = H_{a,\varepsilon}$, an H-invariant Borel probability measure μ is called an *SRB measure* if (i) μ-a.e. H has at least one positive Lyapunov exponent; (ii) the conditional measures of μ on unstable manifolds have densities with respect to the Riemannian volume on these manifolds.

Taking the view that positive Lebesgue measure sets correspond to observable events, one regards an invariant measure as *physically relevant* if its properties are reflected on a positive Lebesgue measure set. In dissipative systems, attractors typically have Lebesgue measure zero, and *a priori* there may not be any physically relevant invariant measures. This is why SRB measures are important: A result from general (nonuniform) hyperbolic theory says that if μ is an ergodic SRB measure with no zero Lyapunov exponents, then the set of points whose orbits have asymptotic distributions given by μ, i.e. the set of points z with the property that for every continuous function φ on M,

$$\frac{1}{n}\sum_{i=0}^{n-1}\varphi(H^i(z)) \ \to \ \int \varphi d\mu \qquad \text{as } n \to \infty,$$

has positive Lebesgue measure.

The result above is obtained by showing that the set of points z which lie on stable manifolds of μ-typical points has positive Lebesgue measure. Since orbits starting from such points have positive Lyapunov exponents, it follows that in the presence of an SRB measure of the kind above, positive Lyapunov exponents are observed on a positive Lebesgue set. This is an important characteristic of chaotic attractors.

SRB measures were discovered for Axiom A attractors by Sinai, Ruelle and Bowen. Not all attractors (outside of the Axiom A category) have SRB measures, however. For more information, see the review article [27].

We formulate next a version of the results of Wang and Young suitable for slow-fast systems.

Theorem 4 *[23, 24] Assume the family $H_{a,\varepsilon}$ satisfies (C0)–(C4). Then for each sufficiently small $\varepsilon > 0$, there is a positive measure set of parameters a for which $H_{a,\varepsilon}$ has an SRB measure $\mu_{a,\varepsilon}$. As a consequence, there is a positive Lebesgue measures set $A_{a,\varepsilon} \subset M$ with the property that for every $z \in A_{a,\varepsilon}$,* *(i) the orbit with initial condition z has a positive Lyapunov exponent;*

(ii) the asymptotic distribution of the orbit with initial condition z is given by $\mu_{a,\varepsilon}$.

If in addition (C5) is satisfied, then with respect to $\mu_{a,\varepsilon}$, $H_{a,\varepsilon}$ is mixing; in fact, it has exponential decay of correlations for all Hölder continuous test functions.

Proof: References for these assertions are as follows:

The existence of an SRB measure μ is Theorem 1.3(1) in [23] – except that the "Misiurewicz condition" in Sect. 1.1 of [23] is replaced by (C2) in the Theorem above. That the results in [23] continue to hold with (C2) in lieu of the "Misiurewicz condition" is proved in Lemma A.4 in the Appendix of [24]. The reason for this replacement is to improve applicability: Negative Schwarzian, which is part of the "Misiurewicz condition", is often not valid or hard to check in applications.

We may assume μ is ergodic (by taking one of its ergodic components if it is not). It has no zero Lyapunov exponents because by definition, one of the Lyapunov exponents is positive, and by virtue of the fact that the $H_{a,\varepsilon}$ can be approximated by rank-one maps (see (C1)), the other is negative. Properties (i) and (ii) follow from the general theory of SRB measures as explained earlier.

If μ is mixing, then the assertion on correlation decay is given by Theorem 1.4 in [23]. That (C5) implies the existence of a mixing SRB measure is proved in Lemma A.5 in the Appendix of [24].‖

Remark 1 We explain the roles of the two parameters a and ε. The latter is a measure of how close the map is to its singular limit; see (C1). The reason for introducing a is that in general, to determine if a given system has an SRB measure requires knowledge of the system to infinite precision. One way to formulate checkable conditions is to introduce a notion of *probability*, expressed here in the form of a parameter, on the space of dynamical systems. The situation is similar to that of the quadratic family in one dimension: For the family $f_a(x) = 1 - ax^2,\ x \in [-1, 1], a \in [0, 2]$, it has been shown that there is an open and dense set of parameters for which the map has a stable periodic orbit [8, 17], while for a positive measure set of parameters, the map has an invariant density with a positive Lyapunov exponent [13]. This phenomenon is expected to carry over to our setting for ε small, i.e. in the complement of the parameters a identified in Theorem 4, it is expected that there are many parameters with stable periodic orbits.

Remark 2 We point out an essential difference between slow-fast systems and systems with a single time scale in the application of Wang-Young theory. In systems with a single time scale, there is usually some form of comparability of determinants, a version of which is condition (**) in [23]. This condition is relaxed in [24] to the existence of $K > 0$ independent of (a, ε) such that

$$\frac{\det(DH_{a,\varepsilon}(z))}{\det(DH_{a,\varepsilon}(z'))} \le K \qquad \text{for all } z, z' \in M. \tag{19}$$

‖ In [24], a determinant condition that is not part of (C0)–(C5) is assumed. This condition is needed for some of the other results in [24]; it is *not* relevant for the results cited here. See Remark 2.

These conditions, which are used to control more detailed dynamical behavior in the basin away from the attractor, cannot be expected to hold for forced oscillations; see Section 5. Thus some of the results in [23] (namely Theorems 1.2(2) and 1.3(2),(3)) cannot be applied directly to the systems treated in this paper. The properties in Theorem 4 do not require this type of condition.

The two dimensional version of Wang-Young theory summarized in this section is adequate for present purposes. We mention for future reference that this body of results has been extended to rank-one attractors in n-dimensional phase spaces where n is an arbitrary integer ≥ 2. This will be published in [26], with another preprint to follow.

5. Verification of (C0)–(C5) in Forced Oscillations

We now explain how the (general) results on chaotic attractors reviewed in Section 4 may be applied to slow-fast systems in general and forced oscillations in particular. In this context, the maps $H_{a,\varepsilon}$ are Poincaré maps $\Pi_{a,\varepsilon} : \Sigma^- \to \Sigma^-$ (see Section 2) for a family of systems parameterized by a. As we will see, a subset of (C0)–(C5) is enjoyed by all slow-fast systems satisfying the assumptions in Section 2, while the validity of remaining conditions depends on specific characteristics of the system in question. We will first identify this non-system specific part of (C0)–(C5), then demonstrate how to verify the remaining conditions numerically for an example.

I. Which parts of (C0)–(C5) hold for general slow-fast systems?

We consider a 1-parameter family of slow-fast systems with the following properties: (i) For each parameter a, the system is given by equation (2) in Section 2; (ii) the coefficients and their derivatives depend smoothly on a; and (iii) Assumptions 1–4 in Section 2 hold. Continuing to use the notation from earlier and choosing Σ^- independently of a, we let $\Pi_{a,\varepsilon} : \Sigma^- \to \Sigma^-$ denote the Poincaré map from the cross-section Σ^- to itself. This is the family $H_{a,\varepsilon}$ for which we now seek to verify (C0)–(C5).

The two items in (C0) are corollaries of standard theorems for existence, uniqueness and smooth dependence of solutions to the initial value problem for systems of differential equations.

With respect to (C1): For fixed a, the convergence of $H_{a,\varepsilon}(\theta, y)$ as functions of (y, θ) to a rank one singular limit as $\varepsilon \to 0$ was proved in Section 3. Convergence of derivatives involving a is treated similarly. Here γ is the image of the projection (along lines parallel to the x-axis) of L^- onto Σ^-. It is clearly diffeomorphic to S^1. Observe that critical points of h are exactly those points $\xi \in \gamma$ with the property that the slow flow on S_a^+ is tangent to $P(L^-)$ at $P(\xi)$. By Assumption 4, these critical points are nondegenerate as required in (C2).

The other properties of (C2), which assert the existence of a singular limit map h_{a^*} with expansion, do not necessarily hold for all families satisfying the conditions above. This comment applies also to (C3).

(C4), on the other hand, is a consequence of the properties assumed. Let γ be as above, and let $\xi \in \gamma$ be a critical point of h. Consider a line segment ω through ξ parallel to the y-axis, i.e. ω is perpendicular to γ. We claim H_0^- maps γ diffeomorphically onto a segment in Σ^+. To see this, recall that there are three parts to H_0^-: the jump P from Σ^- to S_a^+, the slow flow along S_a^+ to L^+, and finally the jump P from L^+ to Σ^+. First, $P(\omega)$ is transverse to the slow flow on the critical manifold because $\xi \in \gamma$ being a critical point, $P(\gamma)$ is tangent to the flow at $P(\xi)$. It follows that the map along trajectories to L^+ is a diffeomorphism on $P(\omega)$. Finally, P maps L^+ diffeomorphically to Σ^+.

(C5) is stronger than (C2); it is also dependent upon the family in question.

We finish by explaining why (19) fails for forced oscillations in general. This is because the volume element along an orbit that spends time t along the slow manifold is contracted by $\sim \exp(-ct/\varepsilon)$. Since different orbits spend different amounts of time near the slow manifold, the ratios in volume contraction between returns to Σ^-, i.e. $\det(DH)$, is in general unbounded as $\varepsilon \to 0$.

II. An example with expansion

Consider the family of vector fields

$$\begin{aligned} \varepsilon\dot{x} &= y + x - \frac{x^3}{3} \\ \dot{y} &= -x + a(x^2-1)\sin(2\pi\theta) \\ \dot{\theta} &= \omega \end{aligned} \tag{20}$$

with $\varepsilon \geq 0$ small. The system (20) is a forced relaxation oscillation, a modification of the forced van der Pol equation [6] in which the forcing amplitude depends upon x. The numerical calculations in this part of the paper are not rigorous; we have not tried to establish error bounds for them. More rigorous numerical work is clearly within the realm of feasibility (see discussion at the end). Our main interest, however, is to demonstrate how to verify numerically the Wang-Young conditions in the context of relaxation oscillations. Since these systems appear extensively as models of physical systems, a procedure for detecting strange attractors can be useful.

First, we show that (20) meets the assumptions in Section 2. As a singularly perturbed vector field, the system (20) has two slow variables (y, θ) and one fast variable x. The critical manifold is the smooth surface $y = x^3/3 - x$ and the fold-curves are $x = \pm 1,\ y = \mp 2/3$. On the fold-curves, the normal switching condition is satisfied because $\dot{y} \neq 0$. Since $\dot{\theta}$ never vanishes, there are no equilibrium points at all. Assumptions (1-3) are satisfied. In particular, all trajectories near the stable sheets of the critical manifold flow to the fold-curves and make regular jumps to the opposite sheet of the critical manifold. Note that the system is symmetric with respect to $\sigma : (x, y, \theta) \to (-x, -y, \theta + 0.5)$.

In the singular limit $\varepsilon = 0$, the (rescaled) slow vector field is

$$\begin{aligned} x' &= -x + a(x^2-1)\sin(2\pi\theta) \\ \theta' &= (x^2-1)\omega \end{aligned} \tag{21}$$

The images of the fold-curves by the jumps P are the curves $x = \mp 2,\ y = \mp 2/3$, and the slow vector field is tangent to these curves at points where $x' = \pm 2 + 3a\sin(2\pi\theta) = 0$. This equation has a pair of solutions on each fold-curve when $a > 2/3$. The nondegeneracy of these tangencies is also easily checked, verifying Assumption 4.

We introduce cross-sections Σ^+ and Σ^- to system (20) defined by $x = 1,\ y > 0$ and $x = -1,\ y < 0$, respectively. Note that σ interchanges Σ^+ and Σ^- and that these cross-sections are almost orthogonal to the vector field. The "half" return map H is defined by flowing from Σ^+ to Σ^- and then applying σ. Fixed points of H correspond to symmetric periodic orbits of (20), and $H \circ H$ is the return map of Σ^+. Therefore, it suffices to establish the required conditions for the rank one maps obtained from H by letting $\varepsilon \to 0$.

To begin with, we treat both ω and a as parameters, and look for (ω, a) such that $h_{\omega,a}$ has the property that its two critical values are fixed points. Figure 5 displays the graph of the singular limit $h : \gamma \to \gamma$ for

$$(\omega, a) = (11.5090088860044, 41.8581499911231)$$

computed with the Dormand-Prince variable step Runge-Kutta algorithm described in Hairer-Wanner [11] at a mesh of 10000 points on the circle $x = 1,\ y = 2/3$. The map h has two critical points located at the values of $\arcsin(2/(3 * a))/(2 * \pi)$, with both critical values being at a fixed point p located near 0.7531. The increasing segment of the image of h wraps around the circle six times, while the decreasing segment of the image of h wraps around the circle five times.

To give additional evidence for the existence of parameter values with the properties above, we use finite difference calculations to estimate the derivatives of $(p, h(c_1), h(c_2))$ with respect to the parameters ω and a. The results are approximately $(0.0022, 0.7833, -1.0517)$ for ω and $(-0.0449, 3.1287, 3.1279))$ for a. These results demonstrate that the map $(\omega, a) \to (h(c_1), h(c_2))$ is nonsingular and gives strong evidence for the existence of parameter values for which $h(c_1) = h(c_2) = p$.

From here on ω is fixed and a is the special parameter a^* in (C2). Numerical calculations in the last paragraph show clearly also the derivatives of $(h(c_1), h(c_2))$ with respect to a are much larger than those of p; this is (C3). It remains to study expansion properties of h corresponding to this special parameter:

Figure 6 displays the graph of h'. Two pairs of horizontal lines bound the region where $|h'| < 1$ and the region where $|h'| < 2$. We see that outside of relatively small neighborhoods $U_1 = (0.0043, 0.0142)$ and $U_2 = (0.4860, 0.4958)$ of the critical points c_1 and c_2, h is expanding by a factor of at least 2. (C2) and (C5) require that we produce a number $\lambda > 2$ such that for every $x \in (U_1 \cup U_2)$, $x \neq c_1, c_2$, there exists $n = n(x) > 1$ such that $h^i(x) \notin (U_1 \cup U_2)$ for all $0 < i < n$ and $|(h^n)'(x)| > \lambda^n$.

To obtain the desired expansion, we introduce two smaller intervals $V_1 = (0.0079, 0.0105)$ and $V_2 = (0.4893, 0.4921)$ inside U_1 and U_2. The derivatives on V_i and $U_i \setminus V_i$ will be estimated separately. Here are some relevant estimates: Outside of V_1 and V_2, $|h'| > 0.5$. Inside V_1, $-440 < |h''| < -420$ while in V_2, $400 < |h''| < 410$. The images

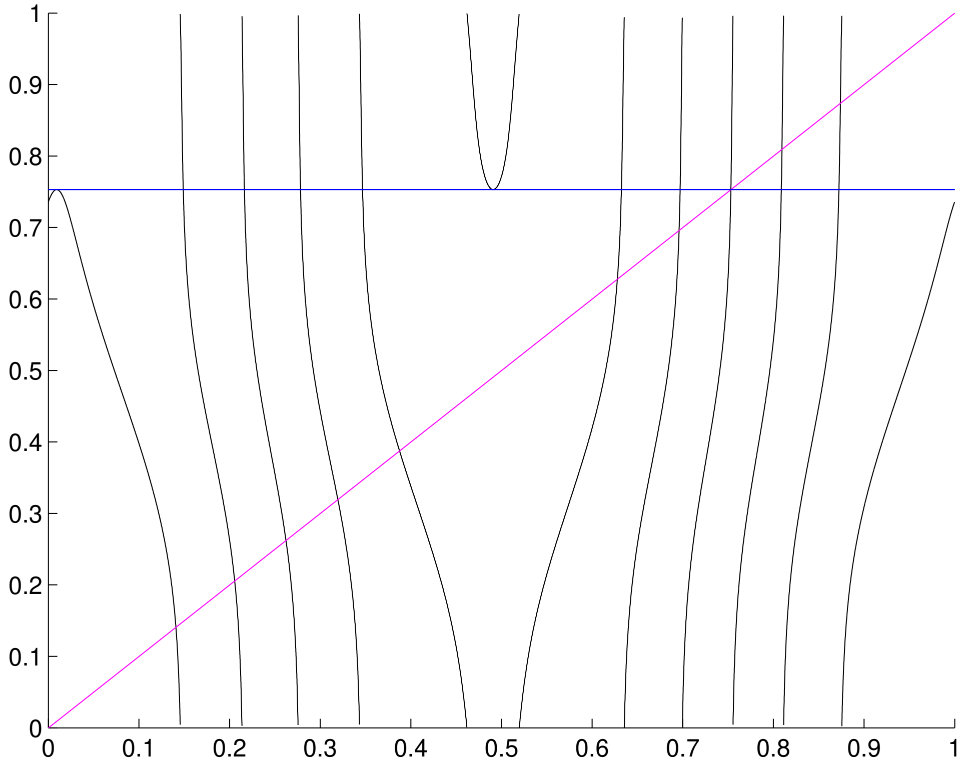

Figure 5. The graph of h

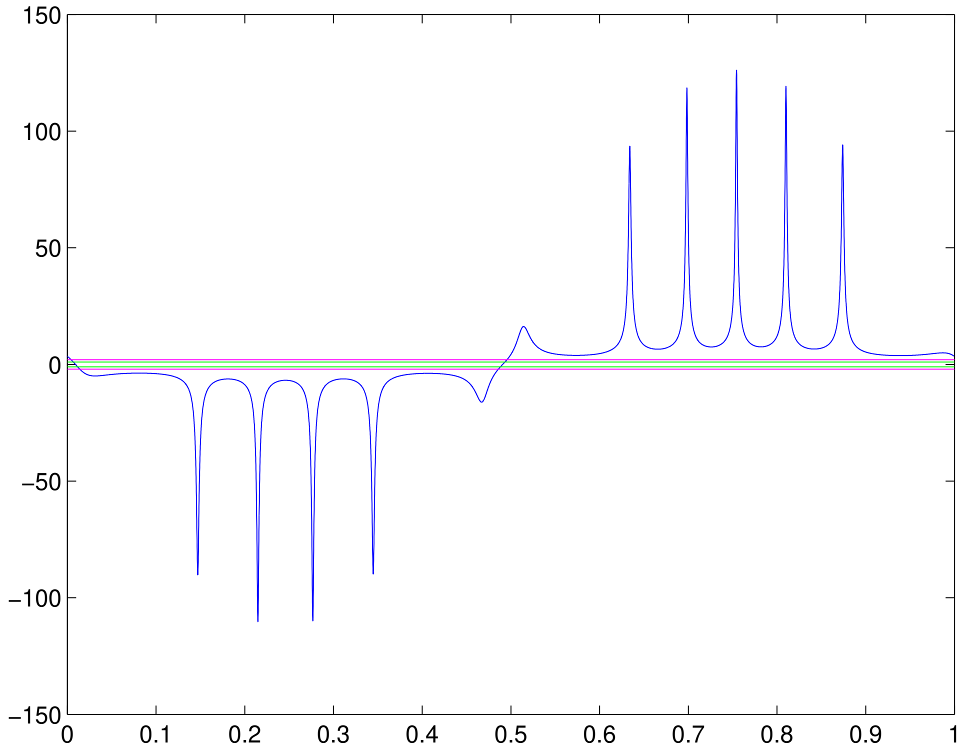

Figure 6. The graph of h'

of V_1 and V_2 are contained in the interval $J = [p - 0.0005, p + 0.0005] \approx [0.7526, 0.7536]$ inside which $56 < |h'| < 100$, while the images of U_1 and U_2 are contained in the interval $[0.747, 0.760]$ inside which $|h'| > 15$. (The derivative $h'(p) > 69$.) Finally, the image of J is contained in the interval $[0.71, 0.81]$ inside which $|h'| > 7.4$.

We assert that these estimates are adequate to prove that there is a $\lambda > 2$ with the required properties. First consider $x \in V_j$, $j = 1$ or 2, and let $k \geq 1$ be such that $h^i(x) \in J$ for all $0 < i \leq k$ and $h^{k+1}(x) \notin J$. This implies $220 \cdot 100^k (x - c)^2 > 0.0005$. Since

$$|x - c| > \sqrt{\frac{0.0005}{220 \cdot 100^k}},$$

$$|(h^{k+1})'(x)| > 400 \cdot 56^k |x - c| > 400 \sqrt{\frac{0.0005}{220}} \left(\frac{56}{\sqrt{100}} \right)^k > 0.6\ (5.6)^k.$$

If $k > 1$, then $|(h^{k+1})'(x)| > 0.6 \cdot (5.6)^k > 2^{k+1}$. For $k = 1$, we use the derivative estimate on $h(J)$ to get $|(h^3)'(x)| > 0.6 \cdot 5.6 \cdot 7.4 > 2^3$. Finally, for $x \in U_j \setminus V_j$, $|(h^2)'(x)| > 0.5 \cdot 15 > 2^2$. We conclude that λ can be chosen > 2, and (C2) and (C5) are satisfied.

This completes the numerical verification of (C2), (C3) and (C5). Our analysis illustrates that specific models of relaxation oscillations fall within the class of systems described by the general theory developed here, but it does not constitute a proof that the system (20) is in this class. Rigorous error bounds for the calculations are feasible in principle. We describe what needs to be done. All of the calculations involve the singular half return map h, obtained by integration of the two dimensional reduced vector field (20) from initial conditions on $x = 2$ to the cross-section $x = 1$. Convergence of Runge-Kutta numerical integration implies that the computed approximation to the map h will converge to h as the precision of the computer arithmetic increases and time steps of the integration decrease. Variational equations can be used to compute the derivatives of h to the same order of accuracy (in step size) as h itself. We believe that interval arithmetic methods applied to the system (20) can be used to prove that it satisfies (C2), (C3) and (C5). Two types of calculations are required: to prove that there are parameter values for which h maps its pair of critical points to fixed points and rigorous verification of the estimates on derivatives that we have given. Our theory of how Henon-like maps can be used to prove the existence of chaotic attractors in singularly perturbed vector fields is independent of the implementation of these calculations. Therefore, we have chosen to leave this matter for future work directed at developing methods for achieving rigorous error estimates for numerical integration. Methods of Guckenheimer and Malo [9, 10] based on transversality are a step in this direction.

Acknowledgments

John Guckenheimer was partially supported by grants from the National Science Foundation and the Department of Energy. Martin Wechselberger was supported at the Mathematical Biosciences Institute, Ohio State University by National Science

Foundation under Agreement No. 0112050. Lai-Sang Young was partially supported by grants from the National Science Foundation

References

[1] Arnold V I, Afrajmovich V S, Il'yashenko Yu S and Shil'nikov L P 1994, *Dynamical Systems V. Encyclopedia of Mathematical Sciences*. (Springer-Verlag)

[2] Benedicks M and Carleson L 1991, The dynamics of the Henon map, *Ann. of Math.* **133** 73–169

[3] Benedicks M and Young L-S 1993, Sinai-Bowen-Ruelle measure for certain Hénon maps, *Invent. Math.* **112** 541–576

[4] Benoît E 1990, Canards et Enlacements, *Publications de l'Inst. des Hautes Etudes Scient.* **72** 63–91

[5] Fenichel N 1979, Geometric Singular Perturbation Theory, *J. Diff Eq.* **31** 53–98

[6] Bold K, Edwards C, Guckenheimer J, Guharay S, Hoffman K, Hubbard J, Oliva R and Weckesser W 2003, The forced van der Pol equation II: Canards in the reduced system, *SIAM J. Appl. Dyn. Syst.* **2** 570–608

[7] Flaherty J and Hoppensteadt F 1978 Frequency entrainment of a forced van der Pol oscillator *Stud. Appl. Math.* **58** 5–15

[8] Graczyk J and Światek G 1998 *The Real Fatou Conjecture*. (Princeton: Princeton University Press)

[9] Guckenheimer J 1995 Phase portraits of planar vector fields: computer proofs *J. Experimental Mathematics* **4** 153–164

[10] Guckenheimer J and Malo S 1996 Computer-generated Proofs of Phase Portraits for Planar Systems, *Int. J. Bifurcation and Chaos* **6** 889–892

[11] Hairer E, Norsett S P and Wanner G 1993, *Solving Ordinary Differential Equations I, 2nd. ed.*, (Springer-Verlag)

[12] Henon M 1976, A two-dimensional mapping with a strange attractor, *Comm. Math. Phys.* **50** 69–77

[13] Jakobson M 1981, Absolutely continuous invariant measures for interval maps, *Comm. Math. Phys.* **81** 39–88

[14] Jones C K R T 1995, Geometric singular perturbation theory, in *Dynamical Systems*, *Springer Lecture Notes Math.* **1609** 44–120

[15] Mishchenko E F, Kolesov Yu S , Kolesov A Yu and Rhozov N Kh 1994, *Asymptotic methods in singularly perturbed systems*, *Monographs in Contemporary Mathematics*, (Consultants Bureau, New York, A Division of Plenum Publishing Corporation)

[16] Lorenz E 1963, Deterministic Nonperiodic Flow, *J. Atmos. Sci*, **20** 130–141

[17] Lyubich M 1997 Dynamics of quadratic polynomials. I, II, *Acta Math.* **178** 185–247, 247–297

[18] Mora L and Viana M 1993, Abundance of strange attractors, *Acta Math.* **171** 1–71

[19] Szmolyan P and Wechselberger M 2001, Canards in $\mathbb{R}^3$, *J. Diff. Eq.* **177** 419–453

[20] Szmolyan P and Wechselberger M 2004, Relaxation Oscillations in $\mathbb{R}^3$, *J. Diff. Eq.* **200** 69–104.

[21] van der Pol B 1926, On relaxation oscillations, *Phil. Mag.* **2** 978–992

[22] Vasil'eva A B 1963, Asymptotic behaviour of solutions of certain problems for ordinary non-linear differential equations with a small parameter multiplying the highest derivatives, *Russian Mathematical Surveys* **18** 13–84 (Russian) *Uspehi Mat. Nauk* **18** 15–86

[23] Wang Q. and Young L-S 2001, Strange Attractors with One Direction of Instability, *Comm. Math. Phys.* **218** 1–97

[24] Wang Q. and Young L-S 2002, From Invariant Curves to Strange Attractors, *Comm. Math. Phys.* **225** 275–304

[25] Wang Q. and Young L-S 2003, Strange Attractors in Periodically-kicked Limit Cycles and Hopf Bifurcations, *Comm. Math. Phys.* **240**, 3, 509–529

[26] Wang Q. and Young L-S 2001, Toward a Theory of Rank One Attractors, *Ann. Math.* to appear.

[27] Young L-S 2002, What are SRB measures, and which dynamical systems have them? *J Stat. Phys.* **108**, 5, 733-754