From Invariant Curves to Strange Attractors

Qiudong Wang\textsuperscript{1} and Lai-Sang Young\textsuperscript{2}

Abstract. We prove that simple mechanical systems, when subjected to external periodic forcing, can exhibit a surprisingly rich array of dynamical behaviors as parameters are varied. In particular, the existence of global strange attractors with fully stochastic properties is proved for a class of second order ODEs.

In the history of classical mechanics, dissipative systems received only limited attention, in part because it was believed that in these systems all orbits eventually tended toward stable equilibria (fixed points or periodic cycles). Evidence that second order equations with a periodic forcing term can have interesting behavior first appeared in the study of van der Pol’s equation, which describes an oscillator with nonlinear damping. The first observations were due to van der Pol and van der Mark. Cartwright and Littlewood proved later that in certain parameter ranges, this equation had periodic orbits of different periods [CL]. Their results pointed to an attracting set more complicated than a fixed point or an invariant curve. Levinson obtained detailed information for a simplified model [Ln]. His work inspired Smale, who introduced the general idea of a horseshoe [Sm], which Levi used later to explain the observed phenomena [Li1].

A number of other differential equations with chaotic behavior have been studied in the last few decades, both numerically and analytically. Examples from the dissipative category include the equations of Lorenz [Lo], [G], [Ro], [Ry], [Sp], [T], [W], Duffing’s equation [D], [Ho], Lorentz gases acted on by external forces [CELS], and modified van der Pol type systems [Li2]. For a systematic treatment of the Lorenz and Duffing equations, see [GH]. While some progress has been made, the number of equations for which a rigorous global description of the dynamics is available has remained small.

In this paper, we consider an equation of the form

\[
\frac{d^2\theta}{dt^2} + \lambda \left(\frac{d\theta}{dt} - 1\right) = \Phi(\theta)P_T(t),
\]

where \( \theta \in S^1 \) and \( \lambda > 0 \). If the right side is set identically equal to zero, this equation

\textsuperscript{1}Dept. of Math., University of Arizona, Tucson, AZ 85721, email dwang@math.arizona.edu. This research is partially supported by a grant from the NSF

\textsuperscript{2}Courant Institute of Mathematical Sciences, 251 Mercer St., New York, NY 10012, email lsy@cims.nyu.edu. This research is partially supported by a grant from the NSF
represents the motion of a particle subjected to a constant external force which causes it to decelerate when its velocity exceeds one and to accelerate when it is below one. Independent of initial condition, the particle approaches uniform motion in which it moves with velocity equal to one. To this extremely simple dynamical system, we add another external force in the form of a pulse: $\Phi$ is an arbitrary function, $P_T$ is time-periodic with period $T$, and for $t \in [0, T)$, it is equal to 1 on a short interval and 0 otherwise. We learned after this work was completed that a similar equation has been studied numerically in the physics literature by G. Zaslavsky.\footnote{Zaslavsky produced in \cite{Z1} numerical evidence of strange attractors. He also discussed in \cite{Z2} how this model can be viewed as a strong idealization of the turbulence problem.}

We prove that the system above exhibits, for different values of $\lambda$ and $T$, a very rich array of dynamical phenomena, including

(a) invariant curves with quasi-periodic behavior,
(b) gradient-like dynamics with stable and unstable equilibria,
(c) transient chaos caused by the presence of horseshoes, with almost every trajectory eventually tending to a stable equilibrium, and
(d) strange attractors with SRB measures and fully stochastic behavior.

These results are new for the equation in question. As abstract dynamical phenomena, (a)–(c) are fairly well understood, and their occurrences in concrete models have been noted; see \cite{GH}. The situation with regard to (d) is very different. The analysis that allows us to handle attractors of this type was not available until recently. To our knowledge, this is the first time a concrete differential equation has been proved analytically to have a global nonuniformly hyperbolic attractor with an SRB measure.\footnote{Levi proved in \cite{L1} the occurrence of scenario (c) for his modified van der Pol systems, not scenario (d) as is sometimes incorrectly reported.} We regard Theorem 3, which discusses the strange attractor case, as the main result of this paper.

Our proof of Theorem 3 is based on \cite{WY}, in which we built a dynamical theory for a (general) class of attractors with one direction of instability and strong dissipation. In \cite{WY}, we identified a set of conditions which guarantees the existence of strange attractors with strong stochastic properties. The properties in question include most of the standard mathematical notions associated with chaos: positive Lyapunov exponents, positive entropy, SRB measures, exponential decay of correlations, symbolic coding of orbits, fractal geometry etc. The occurrence of scenario (d) above is proved by checking the conditions in \cite{WY}. For the convenience of the reader, we will recall these conditions as well as the package of results that follows once these conditions are checked.

Our purpose in writing this paper is not only to point out the range of phenomena that can occur when simple second order equations are periodically forced, but to bring to the foreground the techniques that have allowed us to reach these conclusions.
in a relatively straightforward manner. These techniques are clearly not limited to the systems considered here. It is our hope that they will find applications in other dynamical systems, particularly those that arise naturally from mechanics or physics.

1 Statement of Results

1.1 Setting and assumptions

Consider the differential equation

\[
\frac{d^2 \theta}{dt^2} + \lambda \frac{d\theta}{dt} = \mu + \Phi(\theta)P_T(t)
\]

where \(\theta \in S^1\), \(\lambda, \mu > 0\) are constants, \(\Phi : S^1 \to \mathbb{R}\) is a smooth function, and \(P_T\) has the following form: for some \(t_0 < T\), \(P_T\) satisfies

\[
P_T(t) = P_T(t + T) \quad \text{for all } t
\]

and

\[
P_T(t) = \begin{cases} 
1 & \text{for } t \in [0, t_0] , \\
0 & \text{for } t \in (t_0, T) .
\end{cases}
\]

As discussed in the introduction, (1) describes a simple mechanical system consisting of a particle moving in a circle subjected to an external time-periodic force. With \(r = \frac{d\theta}{dt} - \frac{\mu}{\lambda}\), (1) is equivalent to

\[
\frac{d\theta}{dt} = r + \frac{\mu}{\lambda} , \\
\frac{dr}{dt} = -\lambda r + \Phi(\theta)P_T(t) .
\]

Let \(F_T\) denote the time-\(T\)-map of (2), that is, the map that transforms the phase space \(S^1 \times \mathbb{R}\) from time 0 to time \(T\). Unless explicitly stated otherwise, when we write \(F_T\), it will be assumed that \(T\) is the period of the forcing.

We set \(\mu = \lambda\) for simplicity, and normalize the forcing term as follows: Given a function \(\Phi_0 : S^1 \to \mathbb{R}\), we let \(\Phi = \frac{1}{t_0} \Phi_0\), that is to say, the magnitude of this part of the force is taken to be inversely proportional to the duration of its action, and the proportionality constant is taken to be 1 for simplicity. Our analysis will proceed as follows:

* The function \(\Phi_0\) is fixed throughout. With the exception of Theorem 2(b) (where more is assumed), the only requirements are that \(\Phi_0\) is of class \(C^4\) and all of its critical points are nondegenerate.
* We assume \( t_0 < \frac{1}{10} \min \{ \lambda^{-1}, K_0^{-2} \} \) where \( K_0 = \max \{ \| \Phi_0 \|_{C^4}, 1 \} \). Further restrictions on \( t_0 \) are imposed in each case as needed. (We do not regard \( t_0 \) as an important parameter and will assume it is as small as the arguments require.)

* The two important parameters are \( \lambda \) and \( T \). We will prove that (i) the properties of (1) are intrinsically different for \( \lambda \) small and for \( \lambda \) large, and (ii) for fixed \( \lambda \), the properties of (1) depend quite delicately on the value of \( T \).

To interpret our results correctly, the reader should keep in mind that the dynamical pictures described below are not the only ones that can occur, and it is possible to have combinations of them, such as sinks and strange attractors, on different parts of the phase space. Our aim here is to identify several important pure dynamics types, to indicate the nature and approximate locations of the parameter sets on which they occur, and to convey a sense of prevalence, meaning that these phenomena occur naturally and not as a result of mere coincidence.

### 1.2 Statements of theorems

The setting of Sect. 1.1 is assumed throughout. We consider the discrete-time system defined by the Poincaré map \( F_T \). Precise meanings of some of the technical terms are given after the statements of the theorems. Theorem 3 is our main result. The scenarios presented in Theorems 1 and 2 are also integral parts of the picture.

**Theorem 1 (Existence of invariant curves)** Let \( \lambda \geq 4K_0 \) and \( T \geq t_0 + \frac{3}{2} \). Then there is a simple closed curve \( \Omega \) of class \( C^4 \) to which all the orbits of \( F_T \) converge. Moreover, we have the following dichotomy:

(a) (Quasi-periodic attractors) Let \( \Delta_0 = \{ T : \rho(T) \in \mathbb{R} \setminus \mathbb{Q} \} \) where \( \rho(T) \) is the rotation number of \( F_T|\Omega \). Then (i) \( \Delta_0 \) intersects every unit interval in \( \left[ \frac{3}{2}, \infty \right) \) in a set of positive Lebesgue measure, and (ii) the following hold for \( T \in \Delta_0 \): \( F_T|\Omega \) is topologically conjugate to an irrational rotation, and for every \( z \in S^1 \times \mathbb{R} \),

\[
\frac{1}{n} \sum_{j=0}^{n-1} \delta_{F_T^i z} \text{ converges weakly to } \mu \text{ where } \mu \text{ is the unique invariant probability measure on } \Omega.
\]

(b) (Periodic sinks and saddles) There is an open and dense subset \( \Delta_1 \) of \( [t_0 + \frac{3}{2}, \infty) \setminus \Delta_0 \) such that for \( T \in \Delta_1 \), \( F_T \) has a finite number of periodic sinks and saddles on \( \Omega \). Every orbit of \( F_T \) converges to one of these periodic orbits.

Theorem 2 is elementary; it uses standard techniques, and \( \Phi_0 \) is required only to be \( C^2 \). We include this result because the dynamical pictures described occur for a nontrivial set of parameters.
Theorem 2 (Convergence to stable equilibria)

(a) (Gradient-like dynamics) \( \exists \lambda_0 < \max |\Phi_0| \) such that \( \forall \lambda > \lambda_0 \), if \( t_0 \) is sufficiently small, then there are open intervals of \( T \) for which \( F_T \) has a finite number of periodic points all of which are saddles or sinks, and every orbit not on the stable manifold of a saddle tends to a sink.

(b) (Transient chaos) Assume \( \Phi_0 \) has exactly two critical points. Then there exist intervals of \( \lambda \) accumulating at 0 such that for each of these \( \lambda \), if \( t_0 \) is sufficiently small, then there are open intervals of \( T \) for which \( F_T \) has a periodic sink and a “horseshoe”, i.e. a uniformly hyperbolic invariant set \( \Lambda \) such that \( F_T|\Lambda \) is conjugate to a shift of finite type with positive topological entropy. Lebesgue-a.e. \( z \in S^1 \times \mathbb{R} \) is attracted to the sink as \( n \to \infty \).

Remarks  (i) The picture in Theorem 2(a) is more general than that in Theorem 1(b); there are no simple closed invariant curves in general (see Proposition 4.1).
(ii) We describe the scenario in Theorem 2(b) as “transient chaos” for the following reasons: \( \Lambda \) being an invariant set, points near it tend to stay near it for some period of time, mimicking the dynamics on \( \Lambda \). This chaotic behavior, however, is transient, because \( \Lambda \) has Lebesgue measure zero, and for a typical initial condition, the orbit eventually leaves \( \Lambda \) behind and heads for a sink.

Our next result deals with a notion of chaos that is sustained through time. A compact, \( F_T \)-invariant set \( \Omega \subset S^1 \times \mathbb{R} \) is called a global attractor for \( F_T \) if for every \( z \in S^1 \times \mathbb{R} \), \( \text{dist}(F^n_T z, \Omega) \to 0 \) as \( n \to \infty \). In order not to interrupt the flow of ideas, we postpone the technical definitions of some of the terms used in Theorems 2 and 3 to after the statements of both results.

Here is our main result:

Theorem 3 (Strange attractors) For the parameters specified below, \( F = F_T \) has a strange attractor a description of which follows.

Relevant parameter set: There exist \( \bar{\lambda}, \bar{t}_0 > 0 \) such that for every \( \lambda < \bar{\lambda} \) and \( t_0 < \bar{t}_0 \), there is a positive Lebesgue measure set \( \Delta = \Delta(\lambda, t_0) \) in \( T \)-space for which the results of this theorem hold; \( \Delta \subset [T_0, \infty) \) for some large \( T_0 \), and meets every subinterval of \([T_0, \infty)\) of length \( \mathcal{O}(\lambda) \) in a set of positive Lebesgue measure.

Dynamical characteristics: Let \( \lambda < \bar{\lambda}, t_0 < \bar{t}_0, \) and \( T \in \Delta(\lambda, t_0) \). Then \( F = F_T \) has a global attractor \( \Omega \) with the following dynamical properties:

1. Hyperbolic behavior \( F|\Omega \) is nonuniformly hyperbolic with an identifiable set \( C \subset \Omega \) which is the source of all nonhyperbolic behavior. More precisely:
(a) $\mathcal{C} = \bigcup_i \mathcal{C}_i$ where $\mathcal{C}_i$ is a Cantor set located near $(\theta, r) = (c_i, 0)$, $c_i$ being the critical points of $\Phi_0$; at each $z \in \mathcal{C}$, stable and unstable directions coincide, i.e. there is a vector $v$ with $\|DF^n(z)v\| \to 0$ exponentially fast as $n \to \pm \infty$.

(b) Away from $\mathcal{C}$ the dynamics is uniformly hyperbolic. More precisely, let

$$\Omega_\varepsilon := \{z \in \Omega : d_{\mathcal{C}}(F^n z) \geq \varepsilon \ \forall n \in \mathbb{Z}\}$$

where $d_{\mathcal{C}}(\cdot)$ is a notion of distance to $\mathcal{C}$. Then $\Omega$ is the closure of $\bigcup_{\varepsilon > 0} \Omega_\varepsilon$, $\Omega_\varepsilon$ is a uniformly hyperbolic invariant set for each $\varepsilon > 0$, and the hyperbolicity of $F|\Omega_\varepsilon$ deteriorates (e.g. minimum $\angle(E^u, E^s) \to 0$) as $\varepsilon \to 0$.

(2) Statistical properties

(a) $F$ admits a unique SRB measure $\mu$ supported on $\Omega$.

(b) With the exception of a Lebesgue measure zero set of initial conditions, the asymptotic behavior of every orbit of $F$ is governed by $\mu$. More precisely, for Lebesgue-a.e. $z \in S^1 \times \mathbb{R}$, if $\varphi : S^1 \times \mathbb{R} \to \mathbb{R}$ is a continuous function, then $\frac{1}{n} \sum_{i=0}^{n-1} \varphi(F^i z) \to \int \varphi d\mu$ as $n \to \infty$.

(c) $(F, \mu)$ is ergodic, mixing, and Bernoulli.

(d) For every observable $\varphi : \Omega \to \mathbb{R}$ of H"older class, the sequence

$$\varphi, \varphi \circ F, \varphi \circ F^2, \ldots, \varphi \circ F^n, \ldots$$

viewed as a stochastic process with underlying probability space $(\Omega, \mu)$ has exponential decay of correlations and obeys the Central Limit Theorem.

(3) Symbolic coding and other geometric properties

(a) Kneading sequences are well defined for all critical orbits, i.e. all orbits emanating from $\mathcal{C}$.

(b) With respect to the partition defined by the fractal sets $\mathcal{C}_i$, the coding of orbits in $\Omega$ is well defined and essentially one-to-one. More precisely, if $\sigma$ is the shift operator, then there is a closed subset $\Sigma \subset \prod_{s=1}^{\infty} \{1, \ldots, s\}$ with $\sigma(\Sigma) \subset \Sigma$ and a continuous surjection $\pi : \Sigma \to \Omega$ such that $\pi \circ \sigma = F \circ \pi$; moreover, $\pi$ is one-to-one except over $\bigcup_{c_i} F^n \mathcal{C}$ where it is two-to-one. (In general, $(\Sigma, \sigma)$ is not a shift of finite type.)

(c) Let $h_{\text{top}}(F)$ denote the topological entropy of $F$, $N_n$ the number of cylinder sets of length $n$ in $\Sigma$ above, and $P_n$ the number of fixed points of $F^n$. Then

$$h_{\text{top}}(F) = \lim_{n \to \infty} \frac{1}{n} \log N_n = \lim_{n \to \infty} \frac{1}{n} \log P_n .$$

Moreover, $F$ has an invariant measure of maximal entropy.
For a more detailed description of the dynamics on these strange attractors, see [WY]. We review below the definitions and related background information for some of the technical terms used in the theorems. For more information on this material, see [KH] and [Y1].

A compact $F$-invariant set $\Lambda$ is called uniformly hyperbolic if the following hold: (1) The tangent space at every $x \in \Lambda$ splits into $E^u(x)+E^s(x)$ with $\min_{x \in \Lambda} \angle(E^u, E^s) > 0$; (2) there exist $C \geq 1$ and $\sigma < 1$ such that for all $x \in \Lambda$ and $n \geq 0$, $\|DF^n(x)v\| \leq C_\sigma^n \|v\|$ for all $v \in E^s(x)$, $\|DF^{-n}(x)v\| \leq C_\sigma^n \|v\|$ for all $v \in E^u(x)$.

In Theorem 3(1)(b), not only does $\min \angle(E^u, E^s) \rightarrow 0$ as $\varepsilon \rightarrow 0$, we have $C \rightarrow \infty$ as well. This means the smaller $\varepsilon$, the longer it takes for the geometry of hyperbolic behavior to take hold.

An $F$-invariant Borel probability measure $\mu$ is called an SRB measure if $F$ has a positive Lyapunov exponent $\mu$-a.e. and the conditional measures of $\mu$ on unstable manifolds are equivalent to the Riemannian volume on these leaves. SRB measures are of physical relevance because they can be observed: in dissipative dynamical systems, all invariant probability measures are necessarily singular, but ergodic SRB measures with nonzero Lyapunov exponents have the property that there is a positive Lebesgue measure set of points $z$ for which $\frac{1}{n} \sum_{i=0}^{n-1} \varphi(F^i z) \rightarrow \int \varphi \, d\mu$ as $n \rightarrow \infty$ for every continuous function $\varphi$.

Referring to the set of points $z$ above as the measure-theoretic basin of $\mu$, Theorem 3(2)(b) says that the measure-theoretic basin here is not just a positive Lebesgue measure set, it is, modulo a set of Lebesgue measure zero, the entire phase space.

By a decomposition theorem for SRB measures with no zero exponents ([Le]), the uniqueness of $\mu$ implies that it is ergodic, and the mixing and Bernoulli properties are equivalent to $(F^n, \mu)$ being ergodic for all $n \geq 1$.

We say the dynamical system $(F, \mu)$ has exponential decay of correlations for Hölder continuous observables if given a Hölder exponent $\eta$, there exists $\tau = \tau(\eta) < 1$ such that for all $\varphi \in L^\infty(\mu)$ and $\psi : \Omega \rightarrow \mathbb{R}$ Hölder with exponent $\eta$, there exists $K = K(\varphi, \psi)$ such that

$$\left| \int (\varphi \circ F^n) \psi \, d\mu - \int \varphi \psi \, d\mu \right| \leq K(\varphi, \psi) \tau^n$$

for all $n \geq 1$. Finally, we say the Central Limit Theorem holds for $\varphi$ with $\int \varphi \, d\mu = 0$ if $\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \varphi \circ F^i$ converges in distribution to the normal distribution, and the variance is strictly positive unless $\varphi \circ F = \psi \circ F - \psi$ for some $\psi$. 

7
1.3 Illustrations

Figure 1 below shows the approximate location and shape of the invariant curve or strange attractor (corresponding to different values of \( \lambda \) and \( T \)) for the time-\( T \)-map \( F_T : S^1 \times \mathbb{R} \to S^1 \times \mathbb{R} \).

![Diagram](image)

(a) Invariant curves:
\[
\lambda > 1
\]

(b) Strange attractor:
\[
\lambda << 1
\]

Figure 1

Figure 2 explains the mechanisms behind the changes in the dynamical picture as \( \lambda \) decreases. The straight line in (a) represents \( \{ r = 0 \} \) in \((\theta, r)\)-coordinates, and the subsequent pictures show the images of this line (or circle) at various times under the flow. Figure 2(b) shows the effect of the forcing; observe that it need not constitute a large perturbation. For \( t \in (t_0, T] \), the forcing is turned off, and the system relaxes to a limit cycle with contraction rate \( e^{-\lambda} \). Figure 2(d) shows the image of \( \{ r = 0 \} \) for \( \lambda > 1 \) and \( e^{-\lambda T} \) reasonably contractive; these parameters correspond to the existence of invariant curves. As \( \lambda \) decreases, the effect of the shear term in (2) becomes more prominent, as shown in (e). As \( \lambda \) decreases further, one sees a phenomenon resembling “the breaking of the wave” which accompanies the break-up of the invariant circle. Finally, in Figure 2(g), a tubular neighborhood of \( \{ r = 0 \} \) is folded and mapped into itself, leading to the formation of horseshoes and/or strange attractors.
Figure 2. Image of \( \{r = 0\} \) at time \( t \)
2 Preliminary Information on the ODE

2.1 Singular limits

Let

\[
\begin{pmatrix}
\theta(t) \\
r(t)
\end{pmatrix} = \begin{pmatrix}
\theta(\theta_0, r_0; t) \\
r(\theta_0, r_0; t)
\end{pmatrix}
\]

denote the solution of (2) with \(\theta(0) = \theta_0\) and \(r(0) = r_0\). Then a simple exercise gives

\[
F_T : \begin{pmatrix}
\theta_0 \\
r_0
\end{pmatrix} \mapsto \begin{pmatrix}
\theta(T) \\
r(T)
\end{pmatrix} = \begin{pmatrix}
\theta(t_0) + (T - t_0) + \frac{r(t_0)}{r(t_0)e^{-\lambda(T-t_0)}}
\end{pmatrix}
\]

where the value of \(\theta(T)\) above is to be interpreted as mod 1 or on \(S^1\). We let \(a = \{T - t_0\}\) be the fractional part of \(T - t_0\), \(b = e^{-\lambda n}\) where \(n = [T - t_0]\) is the integer part of \(T - t_0\), and let \(T_{a,b} = F_T\). Then

\[
T_{a,b} : \begin{pmatrix}
\theta_0 \\
r_0
\end{pmatrix} \mapsto \begin{pmatrix}
\theta(t_0) + a + \frac{r(t_0)}{e^{-\lambda}r(t_0)}
\end{pmatrix}.
\]

(3)

(The appearance of "T" in both \(F_T\) and \(T_{a,b}\) is unfortunate; we hope it is not confusing. We wish eventually to make a connection to [WY] and this notation is used there.)

We first fix \(t_0\) and \(\lambda\), and let \(T \to \infty\). Clearly, \(b \to 0\) as \(T \to \infty\). The limit of \(F_T\) as \(T \to \infty\) does not exist. However, \(T_{a,b}\) has the following well defined singular limit as \(b \to 0\):

\[
T_{a,0} : \begin{pmatrix}
\theta_0 \\
r_0
\end{pmatrix} \mapsto \begin{pmatrix}
\theta(t_0) + a
\end{pmatrix}.
\]

(4)

Let \(T^1_{a,0}\) denote the first component of \(T_{a,0}\). We will show in Sect. 2.3 that as \(t_0 \to 0\), \(T^1_{a,0} \to \hat{T}^1_{a,0}\) where

\[
\hat{T}^1_{a,0} : \begin{pmatrix}
\theta_0 \\
r_0
\end{pmatrix} \mapsto \theta_0 + \frac{r_0}{\lambda} \Phi_0(\theta_0) + a.
\]

(5)

In later sections, we will also work with two families of circle maps \(f_a\) and \(\hat{f}_a\) obtained by restricting \(T^1_{a,0}\) and \(\hat{T}^1_{a,0}\) respectively to \(\{r_0 = 0\}\), i.e.

\[
f_a(\theta_0) = \theta(\theta_0, 0; t_0) + \frac{r(\theta_0, 0; t_0)}{\lambda} + a;
\]

\[
\hat{f}_a(\theta_0) = \theta_0 + \frac{1}{\lambda} \Phi_0(\theta_0) + a.
\]

While our results are not confined to these limiting situations, the relation between \(F_T\) and the objects that appear in equation (1), namely, \(\Phi_0, \lambda\) and \(t_0\), can be made transparent by comparing first \(T_{a,b}\) and \(T_{a,0}\) and then \(T^1_{a,0}\) and \(\hat{T}^1_{a,0}\). This is how we will go about obtaining information on \(F_T\).
2.2 The time-$t_0$-map

In this subsection we consider the solution of (2) for $t \in [0, t_0]$ and record some derivative estimates. We first write

$$r(t) = u(t)e^{-\lambda t},$$

$$\theta(t) = v(t) + t - \frac{1}{\lambda} u(t)e^{-\lambda t}.$$  \hspace{1cm} (6)

Differentiating (6) and plugging into (2), we obtain

$$u(t) = u_0 + \frac{1}{t_0} \int_0^t \Phi_0(\theta(\tau)) e^{\lambda \tau} d\tau,$$

$$v(t) = v_0 + \frac{1}{\lambda t_0} \int_0^t \Phi_0(\theta(\tau)) d\tau.$$  \hspace{1cm} (7)

Substituted back into (6), this gives $u_0 = r_0, v_0 = \theta_0 + \frac{r_0}{\lambda}$, and

$$r(t) = \left( r_0 + \frac{1}{t_0} \int_0^t \Phi_0(\theta(\tau)) e^{\lambda \tau} d\tau \right) e^{-\lambda t},$$

$$\theta(t) = \theta_0 + t + \frac{r_0}{\lambda} (1 - e^{-\lambda t}) + \frac{1}{\lambda t_0} \left\{ \int_0^t \Phi_0(\theta(\tau)) d\tau - e^{-\lambda t} \int_0^t \Phi_0(\theta(\tau)) e^{\lambda \tau} d\tau \right\}.$$  \hspace{1cm} (8)

We assume in the rest of this subsection that $|r_0| < 1$.

Recall that $K_0 = \max \{ \| \Phi_0 \|_{C^4}, 1 \}$. We remark that in most of our bounds involving $K_0$, it is in fact not necessary to use the $C^4$-norm. For example, $K_0$ in Lemmas 2.1 and 2.4 can be replaced by the $C^1$-norm of $\Phi_0$.

**Lemma 2.1** (i) $|\theta(t_0) - \theta_0| < 5K_0 t_0$;

(ii) $\frac{\max}{\max \text{ } t \leq t_0} \left| \frac{\partial \theta(t)}{\partial \theta_0} - 1 \right| < 4K_0 t_0$;  \hspace{1cm} $\frac{\max}{\max \text{ } t \leq t_0} \left| \frac{\partial \theta(t)}{\partial r_0} \right| < 2t_0$;

(iii) $\left| \frac{\partial r(t_0)}{\partial \theta_0} \right| < 2K_0$;  \hspace{1cm} $\left| \frac{\partial r(t_0)}{\partial r_0} \right| \leq 2$.

**Proof:** (i) From (8),

$$|\theta(t) - \theta_0| < t + \frac{|r_0|}{\lambda} (1 - e^{-\lambda t}) + \frac{1}{\lambda t_0} \left| \int_0^t \Phi_0(1 - e^{\lambda \tau}) d\tau \right|$$

$$+ \frac{(1 - e^{-\lambda t})}{\lambda t_0} \left| \int_0^t \Phi_0 e^{\lambda \tau} d\tau \right|.$$
Using the inequalities $1 - e^{-x} \leq x$ and $e^x - 1 \leq xe^x$ for $x > 0$ and the fact that $\lambda t_0 < \frac{1}{10}$, we see immediately that the four terms above add up to $< 5K_0t_0$.

(ii) \[
\frac{\partial \theta}{\partial \theta_0} = 1 + A + B
\]

where

\[
A = \frac{1}{\lambda t_0} \int_0^t \Phi'_0 \frac{\partial \theta}{\partial \theta_0} (1 - e^{\lambda \tau}) d\tau , \\
B = \frac{1}{\lambda t_0} (1 - e^{-\lambda t}) \int_0^t \Phi'_0 \frac{\partial \theta}{\partial \theta_0} e^{\lambda \tau} d\tau .
\]  

(9)

Letting $\Theta_1 = \max_{t \leq t_0} |\frac{\partial \theta(t)}{\partial \theta_0} - 1|$ and recalling that $t_0K_0 < \frac{1}{10} K_0^{-1} \leq \frac{1}{10}$, we obtain

\[
\Theta_1 \leq |A| + |B| \leq \frac{5}{2} K_0 t_0 (\Theta_1 + 1) \leq \frac{1}{4} \Theta_1 + \frac{5}{2} K_0 t_0 ,
\]

which implies that $\Theta_1 < 4K_0t_0$. Similarly, writing

\[
\frac{\partial \theta}{\partial r_0} = \frac{1}{\lambda} (1 - e^{-\lambda t}) + \tilde{A} + \tilde{B}
\]

where

\[
\tilde{A} = \frac{1}{\lambda t_0} \int_0^t \Phi'_0 \frac{\partial \theta}{\partial r_0} (1 - e^{\lambda \tau}) d\tau , \\
\tilde{B} = \frac{1}{\lambda t_0} (1 - e^{-\lambda t}) \int_0^t \Phi'_0 \frac{\partial \theta}{\partial r_0} e^{\lambda \tau} d\tau
\]

and reasoning as above, we arrive at the desired bound for $\frac{\partial \theta}{\partial r_0}$.

(iii) follows from (ii) and a straightforward computation. \hfill \qed

We will also need estimates on higher derivatives.

**Lemma 2.2** For $i = 2$ and 3,

\[
\max_{0 \leq t \leq t_0} \left| \frac{\partial^i \theta(t)}{\partial \theta_0^i} \right| \leq 20 K_0 t_0 .
\]

**Proof:** Letting $A$ and $B$ be as in (9), we have

\[
\frac{\partial^2 \theta}{\partial^2 \theta_0} = \frac{\partial A}{\partial \theta_0} + \frac{\partial B}{\partial \theta_0}
\]

12
where
\[
\frac{\partial A}{\partial \theta_0} = \frac{1}{\lambda t_0} \int_0^t \left( \Phi''_0 \left( \frac{\partial \theta}{\partial \theta_0} \right) + \Phi'_0 \frac{\partial^2 \theta}{\partial \theta_0^2} \right) (1 - e^{\lambda \tau}) d\tau,
\]
\[
\frac{\partial B}{\partial \theta_0} = \frac{1}{\lambda t_0} \int_0^t \left( \Phi''_0 \left( \frac{\partial \theta}{\partial \theta_0} \right) + \Phi'_0 \frac{\partial^2 \theta}{\partial \theta_0^2} \right) e^{\lambda \tau} d\tau.
\]
Let \( \Theta_2 = \max_{0 \leq t \leq t_0} \left| \frac{\partial^2 \theta(t)}{\partial \theta_0^2} \right| \). Similar reasoning as before gives
\[
\Theta_2 \leq \frac{5}{2} K_0 t_0 (1 + \Theta_1)^2 + \frac{1}{4} \Theta_2.
\]
and thus \( \Theta_2 < 8 K_0 t_0 \). The proof for \( i = 3 \) is similar. \( \square \)

### 2.3 Relating \( F_T \) to \( \Phi_0, \lambda \) and \( t_0 \)

Let \( T_{a,0}^1 \) and \( \hat{T}_{a,0}^1 \) be as in Sect. 2.1. Reading off \( \theta(t_0) \) and \( r(t_0) \) from equation (8), we have
\[
T_{a,0}^1(\theta_0, r_0) = \theta_0 + \frac{r_0}{\lambda} + \frac{1}{\lambda t_0} \int_0^{t_0} \Phi_0(\theta(t)) dt + a.
\]
From this, it follows immediately that \( |T_{a,0}^1 - \hat{T}_{a,0}^1| \leq \frac{K_0 t_0}{\lambda} \). We record the following estimates for future use.

**Lemma 2.3** (i) 
\[
\frac{4}{5\lambda} < \frac{\partial T_{a,0}^1}{\partial r_0} < \frac{6}{5\lambda};
\]
(ii) There is a numerical constant \( M_0 \) such that for \( i = 1, 2, 3 \),
\[
\left| \frac{\partial T_{a,0}^1}{\partial \theta_0} - \frac{\partial \hat{T}_{a,0}^1}{\partial \theta_0} \right| \leq \frac{M_0 K_0^2 t_0}{\lambda}.
\]

**Proof:** (i) Since
\[
\frac{\partial T_{a,0}^1}{\partial r_0} = \frac{1}{\lambda} + \frac{1}{\lambda t_0} \int_0^{t_0} \Phi'_0 \frac{\partial \theta}{\partial r_0} dt,
\]
it suffices to observe that the second term on the right has absolute value bounded above by \( \frac{1}{\lambda} K_0 (2t_0) \) (Lemma 2.1(ii)), which is \( \frac{1}{5\lambda} \).

(ii)
\[
\left| \frac{\partial T_{a,0}^1}{\partial \theta_0} - \frac{\partial \hat{T}_{a,0}^1}{\partial \theta_0} \right| = \frac{1}{\lambda t_0} \left| \int_0^{t_0} \Phi'_0(\theta(t)) \left( \frac{\partial \theta}{\partial \theta_0} - 1 \right) dt + \int_0^{t_0} \left( \Phi'_0(\theta(t)) - \Phi'_0(\theta_0) \right) dt \right|
\]
\[
\leq \frac{K_0}{\lambda} (4 K_0 t_0 + 5 K_0 t_0) = \frac{9 K_0^2 t_0}{\lambda}
\]

13
by Lemma 2.1(ii) and (i). For the second derivative,

\[
\frac{\partial^2 T_{a,0}^1}{\partial \theta_0^2} - \frac{\partial^2 \hat{T}_{a,0}^1}{\partial \theta_0^2} = \frac{1}{\lambda t_0} \int_0^{t_0} \left( \Phi_0'' \left( \frac{\partial \theta}{\partial \theta_0} \right)^2 + \Phi_0' \frac{\partial^2 \theta}{\partial \theta_0^2} \right) dt - \frac{1}{\lambda} \Phi_0''(\theta_0)
\]

\[
= \frac{1}{\lambda t_0} \int_0^{t_0} \Phi_0'' \left( \left( \frac{\partial \theta}{\partial \theta_0} \right)^2 - 1 \right) dt + \frac{1}{\lambda t_0} \int_0^{t_0} \Phi_0' \frac{\partial^2 \theta}{\partial \theta_0^2} dt
\]

\[
+ \frac{1}{\lambda t_0} \int_0^{t_0} (\Phi_0''(\theta(t)) - \Phi_0''(\theta_0)) dt .
\]

Observe that each of the functions of the integrals is bounded by constant \( K_0^2 t_0 \).
The third derivative is estimated similarly. \( \square \)

We are now ready to estimate the derivative of \( F_T \). Let \( \hat{T}_{a,0} = (\hat{T}_{a,0}^1, 0) \). Then

\[
D \hat{T}_{a,0} = \left( \begin{array}{cc} 1 + \frac{1}{\lambda} \Phi_0' & \frac{1}{\lambda} \\
0 & 0 \end{array} \right).
\]

Lemma 2.4

\[
DF_T = \left( \begin{array}{cc} 1 + \frac{1}{\lambda} \Phi_0' & \frac{1}{\lambda} \\
0 & 0 \end{array} \right) + \left( \begin{array}{cc} A_1 & B_1 \\
C_1 & D_1 \end{array} \right)
\]

where

\[
|A_1| \leq \frac{9K_0^2 t_0}{\lambda} + \frac{2K_0}{\lambda} e^{-\lambda(T-t_0)}, \quad |B_1| \leq \frac{2K_0}{\lambda} + \frac{2}{\lambda} e^{-\lambda(T-t_0)},
\]

\[
|C_1| \leq 2K_0 e^{-\lambda(T-t_0)}, \quad |D_1| \leq 2 e^{-\lambda(T-t_0)}.
\]

Proof: Comparing (3) and (4), and using Lemma 2.1, we see that

\[
DT_{a,b} = DT_{a,0} + \left( \begin{array}{cc} A_2 & B_2 \\
C_2 & D_2 \end{array} \right)
\]

where \( |A_2| \leq \frac{2K_0}{\lambda} e^{-\lambda(T-t_0)}, \ |B_2| \leq \frac{2}{\lambda} e^{-\lambda(T-t_0)}, \ |C_2| \leq 2K_0 e^{-\lambda(T-t_0)}, \) and \( |D_2| \leq 2 e^{-\lambda(T-t_0)} \). By Lemma 2.3,

\[
DT_{a,0} = D \hat{T}_{a,0} + \left( \begin{array}{cc} A_3 & B_3 \\
C_3 & D_3 \end{array} \right)
\]

where \( |A_3| \leq \frac{9K_0^2 t_0}{\lambda}, \ |B_3| \leq \frac{2K_0}{\lambda} \) and \( C_3 = D_3 = 0. \) \( \square \)
2.4 Absorbing sets

For dynamical systems with noncompact phase spaces, it is convenient to know that the action takes place in compact regions. An absorbing set for $F_T$ is an open set $A$ with compact closure such that $F_T(A) \subset A$ and for all $z \in S^1 \times \mathbb{R}$, there exists $n = n(z)$ such that $F^n_T z \in A$.

**Lemma 2.5** Assume $\lambda(T - t_0) \geq 1$. Then

$$A := \{(\theta, r) \in S^1 \times \mathbb{R} : |r| < 4K_0e^{-\lambda(T-t_0)}\}$$

is an absorbing set for $F_T$.

**Proof:** Write $(\theta_n, r_n) = F^n_T(z)$ for $z = (\theta_0, r_0)$. By (3) and (8) we have

$$|r_n| < e^{-\lambda(T-t_0)}|r_{n-1}| + 2K_0e^{-\lambda(T-t_0)}.$$  

With $e^{-\lambda(T-t_0)} < \frac{1}{2}$, this proves $F_T(A) \subset A$. The other condition follows since inductively we have

$$|r_n| < e^{-n\lambda(T-t_0)}|r_0| + 2K_0 \sum_{i=1}^{n} e^{-\lambda(T-t_0)} < 4K_0e^{-\lambda(T-t_0)}.$$

\[\square\]

3 A view from the singular limit

Recall that $f_a$ is the restriction of $T^1_{a,0}$ to $\{r_0 = 0\}$ (see Sect. 2.1). We have thus defined, for each choice of $\Phi_0$, $\lambda$ and $t_0$, a one-parameter family of circle maps which represents the behavior of equation (1) as $T \to \infty$. Conversely, one can recover information on the system defined by (1) if its singular limit $\{f_a\}$ is known: for large $T$ or, equivalently, small $b > 0$, $F_T = T_{a,b}$ can be thought of as a perturbation of $f_a$ or an “unfolding” of $f_a$ to a small neighborhood of $\{r_0 = 0\}$.

In this section, we look at the problem from the point of view of the singular limit. Forgetting temporarily their connection to equation (1), we think of $f_a$ as abstract circle maps. The following is a brief review of several types of behaviors that are known to be “typical” and a general discussion of existing methods for transporting these one-dimensional behaviors to two dimensions.

A. The invertible case: circle diffeomorphisms

The classical theory of Poincaré and Denjoy is well known (see any elementary text). We point out a striking resemblance between $\hat{f}_a$ and the well known family of circle maps first studied by Arnold [A]:

$$g_{\mu,\varepsilon} : x \mapsto x + \mu + \varepsilon \cos(2\pi x), \quad \varepsilon \geq 0.$$
A dichotomy of behavior was observed for this family: “resonant wedges” in the 
$(\mu, \varepsilon)$-plane corresponding to rational rotation numbers, and the “devil’s staircase”
defined by $\mu \mapsto \rho(\mu, \varepsilon)$. These ideas are very much behind our results in Theorem 1.

For us, an important question is how to bring these results for $f_a$ to $T_{a,b}$ for $b > 0$.
KAM techniques (using the intersection property) come to mind for the persistence of
invariant circles with Diophantine rotation numbers, but they will not be used here.
Because of strong normal contraction, invariant curves are shown to exist independent of
rotation number using techniques from hyperbolic theory. The situation is then
reduced to one dimension.

B. Smooth non-invertible circle maps

For general information on one-dimensional maps, see e.g. [dMvS]. Two types of
dynamical behaviors are known to be prevalent. They are (i) maps with attractive
periodic cycles, and (ii) maps with absolutely continuous invariant measures. There
is some evidence that these are the only observable pure dynamics types. For the
quadratic family $Q_a : x \mapsto 1 - ax^2$, (i) and (ii) together account for a set of full
Lebesgue measure in parameter space$^5$ [Lyu2]. We discuss these two cases separately.

(i) Periodic sinks

Continuing to use the quadratic family as a paradigm, we see that period doubling
occurs for $a$ below some $a_0$ (see e.g. [CE]). In this regime, $Q_a$ has a periodic sink
which attracts all points in the interval except for a finite number of unstable periodic
orbits and their pre-images. Above $a_0$, there is an open set of $a$ for which $Q_a$ has
an attractive periodic orbit, but the set of points not attracted to the sink is now
a complicated invariant set on which the map is uniformly expanding. The set of
parameters with this property has been shown to be dense ([GS] and [Lyu1]).

When the dynamical picture of a one-dimensional map is as above, it “unfolds”
into a two-dimensional diffeomorphism satisfying Smale’s Axiom A [Sm]. The passage
of uniform expansion in one dimension to uniform hyperbolicity in two dimensions
is relatively simple due to the robustness or stability of uniform hyperbolic behavior
(see e.g. [Sh]).

(ii) Absolutely continuous invariant measures

There is another type of dynamics that is prevalent in the probabilistic sense.

**Theorem** [J] Let $Q_a : x \mapsto 1 - ax^2$, $a \in [0, 2]$. Then there exists a positive measure
set of $a$ for which $Q_a$ has an absolutely continuous invariant probability measure with
a positive Lyapunov exponent.

The question here is: what does the existence of absolutely continuous invariant
measures for $f_a$ tell us about $T_{a,b}$ for $b$ small? The answer to this question is far
from simple, and the situation was only resolved quite recently. It came as a result of
two different sets of developments. The first is an “abstract” theory of nonuniformly

---

$^5$(i) and (ii) can easily occur on different parts of the phase space for multimodal maps.
hyperbolic systems, which provides a general framework for studying chaotic behavior. The most important idea that has come out of this theory is probably the notion of an *SRB measure* ([Si], [R1], [B], [P], [R2], [Le], [LY], [PS], [Y2]; see also [Y1] for an exposition). The other set of developments is more directly related to small perturbations of one-dimensional maps. Pioneering work in this direction was carried out by Benedicks and Carleson [BC], who achieved an important breakthrough on the Hénon maps. The utility of [BC] in applications, however, is limited by the fact that it relied on computations using explicitly the formulas of the Hénon maps. For other results related to the Hénon maps, see e.g. [BY], [MV].

In a recent preprint [WY], we extended the analysis in [BC] to a more general class of attractors, namely those with strong dissipation, one direction of instability, and well defined singular limits. We also developed the geometric and dynamical pictures of these attractors more fully, merging some of the ideas from [BC] with those from general nonuniform hyperbolic theory. Checkable conditions were given for the first time that guarantee the existence of SRB measures and their stochastic behavior. The properties in the statement of Theorem 3 is a summary of the results in [WY], and this entire package is guaranteed once certain fairly simple conditions are satisfied. These conditions are stated and checked for our equation in Section 5.

4 Proofs of Theorems 1 and 2

4.1 Proof of Theorem 1

We assume throughout this subsection that \( \lambda \) and \( T \) satisfy the hypotheses of Theorem 1, i.e. \( \lambda \geq 4K_0 \) and \( T - t_0 \geq \frac{3}{2} \). This implies in particular that \( e^{-\lambda(T-t_0)} \leq e^{-6} < \frac{1}{100} \). Recall also that a standing assumption throughout is \( t_0 < \frac{1}{100} K_0^2 \leq \frac{1}{10} \) (see Sect. 1.1). Let \( F = F_T \).

4.1.1 Existence of invariant circles

Identifying the tangent space of \( z \in S^1 \times \mathbb{R} \) with the \((\theta, r)\)-plane, we introduce the following cones:

\[
K^c := \{(\theta, r) : |r| < \frac{1}{4} |\theta| \} \\
K^s := \{(\theta, r) : |r| > |\theta| \}.
\]

**Lemma 4.1**

(a) For \( z \in \{|r| < 1\} \), \( v \in K^c \implies DF_z(v) \in K^c \) and \( |DF_z(v)| > \frac{1}{3} |v| \).

(b) For \( z \) with \( F^{-1}z \in \{|r| < 1\} \), \( v \in K^s \implies DF^{-1}_z(v) \in K^s \) and \( |DF^{-1}_z(v)| > 10|v| \).

**Proof:** Write

\[
DF_z = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.
\]
Substituting the admissible values of \( \lambda, T \) and \( t_0 \) into Lemma 2.4, we obtain
\[
|A - 1| < \frac{1}{2} ; \quad |B| < \frac{1}{3} , \quad \text{and} \quad |C|, |D| < \frac{1}{50} .
\] (10)
(The estimate for \( |C| \) uses \( K_0 e^{-6K_0} < e^{-4} \).) Let \( s(v) \) denote the slope of a vector \( v \in \mathbb{R}^2 \). Then
\[
s(DF(v)) = \frac{C + Ds(v)}{A + Bs(v)} .
\]
To verify \( DF(\mathcal{K}^e) \subset \mathcal{K}^e \), for example, we choose \( v \) with \( |s(v)| < \frac{1}{4} \), and substituting in the numbers from (10), we obtain
\[
|s(DF(v))| < \frac{\frac{1}{50} + \frac{1}{10}}{\frac{1}{2} - \frac{1}{3}} < \frac{1}{4} .
\]
The other claims are checked similarly.

We have thus identified a family of stable cones \( \mathcal{K}^s \) and a family of center cones \( \mathcal{K}^c \). We call \( \mathcal{K}^c \) “center cones” because while vectors in \( \mathcal{K}^c \) may be expanded or contracted by \( DF \), they are not contracted as strongly as vectors in \( DF^{-1}(\mathcal{K}^s) \). This domination implies uniform hyperbolicity on the projective level, a property relied upon heavily in the proof of the next lemma.

Recall that \( A = \{ |r| < 4K_0 e^{-\lambda(T-t_0)} \} \) is an absorbing set of \( F \) (Lemma 2.5).

**Lemma 4.2** There is an \( F \)-invariant curve \( \Omega \) in \( A \) such that

(a) \( \Omega \) is the graph of a \( C^4 \) function \( g : S^1 \to \mathbb{R} \) with \( |g'| < 1/4 \);

(b) for every \( z \in A \), \( \text{dist}(F^n z, \Omega) \to 0 \) as \( n \to \infty \).

**Proof:** By standard arguments from hyperbolic theory, it follows from Lemma 4.1 that there is a stable foliation \( W^s \) defined everywhere on \( A \). Tangent vectors to the leaves of \( W^s \) satisfy \( |s(v)| > 1 \), so that each \( W^s \)-leaf is a \( C^1 \) segment joining the two boundary components of \( A \). Moreover, \( F \) maps each \( W^s \)-leaf strictly into a \( W^s \)-leaf, contracting length by a factor \( < \frac{1}{15} \). It follows from this that \( \Omega := \cap_{n>0} F^n(A) \) is a compact set which meets each \( W^s \)-leaf in exactly one point. Part (b) of Lemma 4.2 follows immediately.

Let \( \gamma_0 \) be the curve \( \{ r = 0 \} \). Then the images \( \gamma_n := F^n \gamma_0 \) converge in the Hausdorff metric to \( \Omega \), the center manifold of \( F \). By Lemma 4.1(a), the tangent vectors to \( \gamma_n \) have slopes between \( \pm 1/4 \) for all \( n \). This proves that \( \Omega \) is the graph of a Lipschitz function \( g \) with Lipschitz constant \( \leq 1/4 \). That \( g \) is \( C^4 \) follows from the fact that \( F \) is \( C^4 \) and standard graph transform arguments involving the Fiber Contraction Theorem. We refer the reader to [HPS].

\[ \square \]
4.1.2 Dynamics on invariant circles

For each $T$, let $\Omega_T$ be the simple closed curve left invariant by $F_T$. We introduce a family of maps $h_T : S^1 \to S^1$ as follows: For $\theta_0 \in S^1$, let $z$ be the unique point in $\Omega_T$ whose $\theta$-coordinate is $\theta_0$. Then $h_T(\theta_0) = \theta_1$ where $\theta_1$ is the $\theta$-coordinate of $F_T(z)$. Let $\rho(h_T)$ denote the rotation number of $h_T$. Since

$$\frac{d\theta_1}{dT} > 1 - e^{-\lambda(r-t_0)}|r(t_0)| > \frac{99}{100},$$

it is an easy exercise to see that $T \mapsto \rho(h_T)$ is a continuous nondecreasing function with $\rho(h_{T+1}) \approx \rho(h_T) + 1$.

**Case 1** $\rho(h_T) \in \mathbb{R} \setminus \mathbb{Q}$. By Denjoy theory, $h_T$ is topologically conjugate to the rigid rotation by $\rho(h_T)$, which is well known to admit only one invariant probability measure. This together with Lemmas 2.5 and 4.2(b) imply immediately the unique ergodicity of $F_T$. To prove that $\Delta_0$ in Theorem 1 has positive Lebesgue measure, we appeal to the following theorem of Herman:

**Theorem** [He] Let $\text{Diff}^r_+(S^1)$ denote the space of $C^r$ orientation-preserving diffeomorphisms of $S^1$. Let $s \mapsto h_s \in \text{Diff}^3_+(S^1)$ be $C^1$ and suppose that for some $s_0 < s_1$, $\rho(h_{s_0}) \neq \rho(h_{s_1})$. Then $\{s \in [s_0, s_1] : \rho(h_s) \in \mathbb{R} \setminus \mathbb{Q}\}$ has positive Lebesgue measure.

**Case 2** $\rho(h_T) \in \mathbb{Q}$. We fix $p, q \in \mathbb{Z}^+$, $p, q$ relatively prime, and let $I$ be a connected component of $\{T : \rho(h_T) = \frac{p}{q}\}$ with nonempty interior. From (11), it follows that

$$\frac{d}{dT}(h_T^q(\theta_0)) > \frac{99}{100}$$

for every $\theta_0$. Standard transversality arguments give an open and dense subset $\tilde{I}$ of $I$ such that for $T \in \tilde{I}$, the graph of $h_T^q$ is transversal to the diagonal of $S^1 \times S^1$. For $T \in \tilde{I}$, the fixed points of $h_T^q$ (in the order in which they appear on $S^1$) are alternately strictly repelling and strictly contracting. With the contraction normal to $\Omega_T$, they correspond to saddles and sinks respectively for $F_T$.

This completes the proof of Theorem 1.

4.2 Proof of Theorem 2

Our analysis will proceed as follows. Referring the reader to Sect. 2.1 for definitions and notation, we will argue that uniformly expanding invariant sets of $f_a$ translate directly into uniformly hyperbolic invariant sets of $T_{a,b}$ for $b$ sufficiently small. That being the case, to produce the phenomena described in Theorem 2, it suffices to produce the corresponding behaviors for $f_a$. Furthermore, since uniformly expanding invariant sets are stable under perturbations, and $f_a$ is a small perturbation of $\hat{f}_a$ for $t_0 << \lambda$ (Lemma 2.3), it suffices to work with $\hat{f}_a$. Recall that

$$\hat{f}_a(s) = s + \frac{1}{\lambda} \Phi_0(s) + a.$$

19
4.2.1 Gradient-like dynamics

Let $m_0 = - \min \Phi_0$. Then $f_a$ is a circle diffeomorphism if and only if $\lambda > m_0$. Fix $\lambda > m_0$. Varying $a$ (which corresponds to moving the graph of $f_a$ up and down), we see that there is an open set of $a$ for which $f_a$ has a finite number of fixed points which are alternately repelling and attracting. For these $a$, it is a simple exercise to show that for sufficiently small $t_0$ and $b$, $F_T = T_{a,b}$ has the gradient-like dynamics described in Theorem 2. More generally, if $\rho(f_a) = \frac{2}{\lambda}$, then the discussion above applies to $f_a^q$ unless $f_a^q = id$.

Gradient-like dynamics, in general, persist when $\lambda$ drops below $m_0$. Intuitively, no simple closed invariant curve exists beyond this point because the unstable manifold of the saddle “turns around”. We provide a rigorous proof in a restricted context.

**Proposition 4.1** Suppose $\Phi_0$ has exactly two critical points and negative Schwarzian derivative. Then there exist intervals of $\lambda$, $t_0$ and $T$ for which $F_T$ has gradient-like dynamics but there are no smooth simple closed invariant curves.

**Proof** Let $c_1$ and $c_2$ denote the critical points of $\Phi_0$. There is an interval of $a_0$ such that if $\Phi_0 = a_0 + \Phi_0$, then $\Phi_0$ has exactly two zeros, at say $p_1$ and $p_2$. Fix such an $a_0$. Without loss of generality, we assume $p_1 < c_1 < p_2 < c_2 < p_1 + 1 = p_1$, and $\Phi_0(p_1) > 0$, $\Phi_0(p_2) < 0$. In the rest of the proof, for each $\lambda$ we consider, let $f = f_a$ where $a = -\frac{m}{\lambda}$ mod 1, so that $f(s) = s + \frac{1}{\lambda} \Phi_0(s)$. Observe that $p_1$ is a repelling fixed point of $f$, $p_2$ is an attractive fixed point of $f$, and $f'(c_1) = f'(c_2) = 1$. This discussion is valid for all $\lambda$.

For large $\lambda$, $f$ maps $(c_1, c_2)$ strictly into itself. (See Figure 3(a).) This continues to be the case for some interval of $\lambda$ below $m_0$. Since $\Phi'_0 < 0$ on $(c_1, c_2)$, we have $1 - \frac{m}{\lambda} < f' < 1$ on $(c_1, c_2)$, so there exist $\varepsilon$, $\varepsilon' > 0$ and an interval $L$ of $\lambda$ below
$m_0$ for which \( f(c_1 + \varepsilon, c_2 - \varepsilon) \subseteq (c_1 + 2\varepsilon, c_2 - 2\varepsilon) \) and \( |f'(c_1 + \varepsilon, c_2 - \varepsilon)| < 1 - \varepsilon' \). (See Figure 3(b).) Thus every point in \((c_1 + \varepsilon, c_2 - \varepsilon)\) tends to \( p_2 \), and since every point in \( S^1 \setminus (c_1 + \varepsilon, c_2 - \varepsilon) \) eventually enters \((c_1 + \varepsilon, c_2 - \varepsilon)\), we conclude that \( f \) and hence \( F = T_{a,b} \) have gradient-like dynamics for \( a \) as above and \( t_0 \) and \( b \) suitably small.

Let \( \tilde{p}_1 \) and \( \tilde{p}_2 \) denote the saddle and sink of \( F \) respectively. To prove the proposition, suppose \( F \) leaves invariant a smooth simple closed curve \( \Omega \). Since it is not possible for all the points in an invariant circle to converge to the same point, \( \Omega \) must intersect the stable manifold of \( \tilde{p}_1 \). This implies \( \tilde{p}_1 \in \Omega \), and hence \( W^u \), the unstable manifold of \( \tilde{p}_1 \), must be contained in \( \Omega \). Fix an orientation on \( \Omega \), and let \( \tau \) be a positively oriented tangent field on \( W^u \). To derive a contradiction, we will produce, for every \( \varepsilon_1 > 0 \), two points \( z, z' \in W^u \) such that \( d(z, z') < \varepsilon_1 \) and \( \tau(z) \) and \( \tau(z') \) point in opposite directions.

By the negative Schwarzian property of \( \Phi_0 \), \( f' = 0 \) at exactly two points \( x_1 < x_2 \) in \((c_1, c_2)\). Move \( \lambda \) if necessary so \( x_i \neq p_2 \), \( i = 1, 2 \). Without loss of generality, we may assume \( x_1 \in (c_1, p_2) \). The following two statements, which we claim are valid for suitable choices of \( t_0, a \) and \( b \), clearly lead to the desired contradiction.

(1) The right branch of \( W^u \) is roughly horizontal until about \( f(x_1) \), where it makes a sharp turn and doubles back for a definite distance, creating two roughly parallel segments with opposite orientation (see Figure 4).

(2) There exist pairs of points on these parallel segments joined by stable curves.

![Figure 4](image)

Claims (1) and (2) follow from Lemma 4.3, which is a general result valid for any \( \lambda \) and any \( \Phi_0 \) (and not just the ones considered in this subsection). It is similar in spirit to Lemma 4.2 and has the same proof, which will be omitted.

**Lemma 4.3** Given \( f_a \) and constants \( \delta, \varepsilon > 0 \), \( \exists \overline{b} = \overline{b}(\Phi_0, \lambda, \delta, \varepsilon) \ll \delta \) such that the following hold for \( F = T_{a,b} \) with \( b < \overline{b} \). Let \( z = (r, \theta) \in A \) (which depends on \( b \)) be such that \( |f'_a(\theta)| > \delta \). Then:

(a) \( |s(v)| = \mathcal{O}(\frac{\delta}{\varepsilon}) \implies |s(DF_zv)| = \mathcal{O}(\frac{\varepsilon}{\delta}) \) and \( |DF_zv| > (1 - \varepsilon)\delta|v| \);

(b) there exists \( C = C(\Phi_0, \lambda) \) such that \( |s(DF_zv)| > C\delta \implies |s(v)| > C\delta \)

and \( \frac{|DF_zv|}{|v|} = \mathcal{O}(\frac{\varepsilon}{\delta}) \).

Claim (1) follows immediately from Lemma 4.3(a). Part (b) of this lemma implies that if a region of \( A \) misses the two rectangles \( \{(r, \theta) : |f'(\theta)| < \delta\} \) in all of its forward
iterates, then it is foliated by stable curves. Since $f'(p_2) \neq 0$, Claim (2) is easily arranged by choosing $\delta$ sufficiently small.

\[ \text{4.2.2 Transient chaos} \]

We return to the family $\hat{f}_a$ where $\lambda$ is now assumed to be small. Let $c_1$ and $c_2$ be the critical points of $\Phi_0$. Then $\hat{f}_a$ has exactly two critical points $s_1$ and $s_2$ near $c_1$ and $c_2$. Let $a$ be fixed for now. As $\lambda$ is varied, the critical values $\hat{f}_a(s_1)$ and $\hat{f}_a(s_2)$ move at rates $\sim \frac{1}{\lambda}$ in opposite directions. There exists, therefore, a sequence of $\lambda$ for which they coincide. Observe that this sequence is independent of $a$. We now fix each of these $\lambda$ and adjust $a$ so that $\hat{f}_a(s_1) = s_1$ where $s_1$ is the critical point with the property that $|\Phi''_0(c_1)| \leq |\Phi''_0(c_2)|$. We will show that for the $(\lambda, a)$-pairs selected above, $f = \hat{f}_a$ has the following properties: (i) it has a sink, and (ii) when restricted to the set of points that are not attracted to the sink, $f$ is uniformly expanding.

By design, we have $f(s_1) = s_1$, which is therefore a sink, and $f(s_2) = s_1$. For $i = 1, 2$, let $\alpha_i = \frac{\sqrt{1.4}}{|\Phi''_0(c_i)|} \lambda$ and $I_i = [s_i - \alpha_i, s_i + \alpha_i]$.

**Lemma 4.4** Assume $\lambda$ is sufficiently small. Then

(a) for $s \not\in I_1 \cup I_2$, we have $|f'(s)| > \sqrt{1.4}$;

(b) for $s \in I_1 \cup I_2$, we have $f^n s \to s_1$ as $n \to \infty$.

**Proof:** (a) We may assume for $s \not\in I_1 \cup I_2$ that $|f'(s)| \geq |f'(s_i \pm \alpha_i)|$ for some $i$. Since this is $\frac{1}{\lambda} |\Phi''_0(c_i)| \alpha_i$ for some $\xi_i \in I_i$, it is $> \sqrt{1.4}$.

(b) First we check $f(I_i) \subset I_1$, $i = 1, 2$:

$$|f(s_i \pm \alpha_i) - f(s_i)| = \frac{1}{2\lambda} |\Phi''_0(\xi_i)| \alpha_i^2 \leq \frac{1}{2\lambda} |\Phi''_0(\xi_i)| \cdot \frac{1.5}{|\Phi''_0(c_i)|^2} \lambda^2 \leq \frac{\lambda}{|\Phi''_0(c_i)|} \leq \frac{\lambda}{|\Phi''_0(c_1)|} < \alpha_1.$$  

A similar computation shows that $f$ restricted to $I_1$ is a contraction. \qed

Let $F = T_{a,b}$ where $\lambda$ and $a$ are near the ones selected above and $t_0$ and $b$ are sufficiently small. Let $B_i$, $i = 1, 2$, be the two components of $A \setminus \{(\theta, r) : \theta \in I_1 \cup I_2\}$. With $\lambda$ sufficiently small, $F$ wraps each $B_i$ around $A$ (in the horizontal direction) at least once, with $F(B_i)$ crossing completely $B_j$ every time they meet. This, on the topological level, is the standard construction of a horseshoe. Let

$$\Lambda := \{z \in A : F^n(z) \in B_1 \cup B_2 \forall n \in \mathbb{Z}\}.$$  

With $b$ sufficiently small, the uniform hyperbolicity of $F|\Lambda$ follows from Lemma 4.3.

This completes the proof of Theorem 2.
5 Proof of Theorem 3

5.1 Conditions from [WY] for strange attractors

As explained in the introduction, the proof of Theorem 3 is obtained largely via a direct application of [WY] – provided the conditions in Sect. 1.1 of [WY] are verified. For the convenience of the reader, we give a self-contained discussion of these conditions here, modifying one of them to improve its checkability and adding a new one, (C4), to guarantee mixing. The notation in this section is that in [WY].

We consider a family of maps $T_{a,b} : A = S^1 \times [-1,1] \to A$ where $a \in [a_0,a_1] \subseteq \mathbb{R}$ and $b \in B_0 \subseteq \mathbb{R}$, $B_0$ being any subset with 0 as an accumulation point.\(^6\)

In this setup, $b$ is a measure of dissipation; our results hold for $b$ sufficiently small. We explain the role of the parameter $a$: For systems that are not uniformly hyperbolic, a scenario that competes with that of strange attractors and SRB measures is the presence of periodic sinks. In general, arbitrarily near systems with SRB measures, there are open sets of maps with sinks; proving directly the existence of an SRB measure for a given dynamical system requires information of arbitrarily high precision. We get around this problem by considering one-parameter families, in our case $a \to T_{a,b}$, and by showing that if a family satisfies certain reasonable conditions, then a positive measure set of parameters with SRB measures is guaranteed. We now state our conditions on these families.

(C1) Regularity conditions

(i) For each $b \in B_0$, the function $(x,y,a) \mapsto T_{a,b}(x,y)$ is $C^3$; and as $b \to 0$, these functions converge in the $C^3$ norm to $(x,y,a) \mapsto T_{a,0}(x,y)$.

(ii) For each $b \neq 0$, $T_{a,b}$ is an embedding of $A$ into itself, whereas $T_{a,0}$ is a singular map with $T_{a,0}(A) \subset S^1 \times \{0\}$.

(iii) There exists $K > 0$ such that for all $a,b$ with $b \neq 0$,

$$\frac{|\det DT_{a,b}(z)|}{|\det DT_{a,b}(z')|} \leq K \quad \forall z, z' \in S^1 \times [-1,1].$$

As before, we refer to $T_{a,0}$ as well as its restriction to $S^1 \times \{0\}$, i.e. the family of one-dimensional maps $f_a : S^1 \to S^1$ defined by $f_a(x) = T_{a,0}(x,0)$, as the singular limit of $T_{a,b}$. The rest of our conditions are imposed on the singular limit alone.

The second condition in [WY] is:

(C2) There exists $a^* \in [a_0,a_1]$ such that $f = f_{a^*}$ satisfies the Misiurewicz condition.

The Misiurewicz condition (see [M]) encapsulates a number of properties some of which are hard to check or not needed in full force. We propose here to replace it by

\(^6\)In [WY], $B_0$ is taken to be an interval but the formulation here is all that is used.
(C2’) a set of conditions that is more directly checkable (although a little cumbersome to state). That the results in [WY] are valid when (C2) is replaced by (C2’) below is proved in Lemma A.1 in the Appendix.

(C2’) Existence of a sufficiently expanding map from which to perturb
There exists a* ∈ [a0, a1] such that f = f a* has the following properties: There are numbers c1 > 0, N1 ∈ Z+, and a neighborhood I of the critical set C such that
(i) f is expanding on S¹ \ I in the following sense:
   (a) if x, fx, ..., f^n x ∉ I, n ≥ N1, then |(f^n)x| ≥ ε^n;
   (b) if x, fx, ..., f^n x ∉ I and f^n x ∈ I, any n, then |(f^n)x| ≥ ε^n;
(ii) f^n x ∉ I ∀ x ∈ C and n > 0;
(iii) in I, the derivative is controlled as follows:
   (a) |f^n| is bounded away from 0;
   (b) by following the critical orbit, every x ∈ I \ C is guaranteed a recovery
time n(x) ≥ 1 with the property that f^j x ∉ I for 0 < j < n(x) and
      |(f^(n(x))x)| ≥ ε^n(x).

Next we introduce the notion of smooth continuations. Let C a denote the critical set of f a. For x = x(a*) ∈ C a*, the continuation x(a) of x to a near a* is the unique critical point of f a near x. If p is a hyperbolic periodic point of f a*, then p(a) is the unique periodic point of f a near p having the same period. It is a fact that in general, if p is a point whose f a-orbit is bounded away from C a*, then for a sufficiently near a*, there is a unique point p(a) with the same symbolic itinerary under f a.

(C3) Conditions on f a* and T a*,0
(i) Parameter transversality For each x ∈ C a*, let p = f(x), and let x(a) and
    p(a) denote the continuations of x and p respectively. Then
    \[ \frac{d}{da} f_a(x(a)) \neq \frac{d}{da} p(a) \quad \text{at } a = a*. \]

(ii) Nondegeneracy at “turns”
    \[ \frac{\partial}{\partial y} T_{a*,0}(x, 0) \neq 0 \quad \forall x \in C_{a*}. \]

The following fact often facilitates the checking of condition (C3)(i):

Lemma 4.1 ([TTY], Section VII) Let f = f a*, and suppose \( \sum_{n \geq 0} \frac{1}{|f^n|} \frac{1}{|f x|} < \infty \) for all x ∈ C. Then
\[
\sum_{k=0}^{\infty} \frac{[(\partial_a f_a)(f^k x)]_{a=a*}}{(f^k)'(fx)} = \left[ \frac{d}{da} f_a(x(a)) - \frac{d}{da} p(a) \right]_{a=a*}. \]
The main conditions in [WY] are contained in (C1)–(C3) (or, equivalently, (C1), (C2') and (C3)). The conclusions of Theorem 3, however, are more specific than those of [WY], which allow the co-existence of multiple ergodic SRB measures. We now introduce a fourth condition,\(^7\) which along with (C1)–(C3) implies the uniqueness of SRB measures and their mixing properties. This implication is proved in Lemma A.2 in the Appendix.

(C4) **Conditions for mixing**

(i) \( e^{c_1} > 2 \) where \( c_1 \) is in (C2').

(ii) Let \( J_1, \ldots, J_r \) be the intervals of monotonicity of \( f_a \), and let \( P = (p_{i,j}) \) be the matrix defined by

\[
p_{i,j} = \begin{cases} 
1 & \text{if } f(J_i) \supset J_j, \\
0 & \text{otherwise}.
\end{cases}
\]

Then there exists \( N_2 > 0 \) such that \( P^{N_2} > 0 \).

The discussion in this subsection can be summarized as follows:

**Theorem 3'** Assume \( \{T_{a,b}\} \) satisfies (C1), (C2'), (C3) and (C4) above. Then for all sufficiently small \( b > 0 \), there is a positive measure set of \( a \) for which \( T_{a,b} \) has the properties in (1), (2) and (3) of Theorem 3.

We remark that [WY] contains a more detailed description of the dynamical picture than the statement of Theorem 3 and refer the interested reader there for more information.

In the rest of this section the discussion pertains to the differential equation (1) defined in Sect. 1.1. All notation is as in Sect. 2.1. To prove Theorem 3, it suffices to verify that for the parameters in question, \( T_{a,b} \) satisfies the conditions above. This is carried out in the next three subsections.

### 5.2 Verification of (C2'): expanding properties

Among the conditions to be checked, (C2'), which guarantees a suitable environment from which to perturb, is arguably the most fundamental of the four. It is also the one that requires the most work. In this subsection, we will – after placing some restrictions on \( \lambda \) and \( t_0 \) – show that (C2') is valid for all \( f_a \) for which (C2')(ii) is satisfied. The existence of \( a \) satisfying (C2')(ii) is the topic of the next subsection.

---

\(^7\)Condition (*) in Sect. 1.2 of [WY], the only condition in [WY] not implied by (C1)–(C3), is clearly contained in (C4)
Let \( \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{k_1} \) be the critical points of \( \Phi_0 \), and let \( k_2 = \min \{1, \frac{1}{2} \min_i |\Phi''_0(\bar{x}_i)|\} \).

We fix \( \varepsilon = \varepsilon(\Phi_0) > 0 \) with the property that \( |\bar{x}_i - \bar{x}_j| > 4\varepsilon \) for \( i \neq j \) and \( |\Phi''_0| > k_2 \) on \( \cup_i (\bar{x}_i - 2\varepsilon, \bar{x}_i + 2\varepsilon) \), and claim that by choosing \( \lambda \) and \( t_0 \) sufficiently small, we may assume the following about \( f_\alpha \). Let \( C \) denote the critical set of \( f_\alpha \), and let \( C_\varepsilon \) denote the \( \varepsilon \)-neighborhood of \( C \). Then

(i) \( C = \{x_1, \ldots, x_{k_1}\} \) with \( |x_i - \bar{x}_i| < \varepsilon \);

(ii) \( \text{on } C_\varepsilon, \ |f''_\alpha| > \frac{k_2}{\lambda} \).

To justify these claims, observe first that by taking \( \lambda \) small enough, the critical set of \( f_\alpha \) can be made arbitrarily close to that of \( \Phi_0 \). Second, by choosing \( t_0 \) sufficiently small (independent of \( \lambda \)), we can make \( \|f_\alpha - \tilde{f}_\alpha\|_{C^3} < \frac{\varepsilon_1}{4} \) for \( \varepsilon_1 \) as small as we please (Lemma 2.3). These observations together with \( \tilde{f}_\alpha'' = \frac{1}{\lambda} \Phi''_0 \) imply (i) and (ii).

A number of other conditions will be imposed on \( \lambda \); they will be specified as we go along. Some of these conditions are determined via an auxiliary constant \( K > 1 \) which depends only on \( \Phi_0 \) and which will be chosen to be large enough for certain purposes. Let \( \sigma := 2k_7^{-1}K^3\lambda \). We assume \( \frac{1}{2}\sigma < \varepsilon \), so that \( |f''_\alpha(x)| > K^3 \) for \( x \in C_\varepsilon \setminus C_{\frac{1}{2}\sigma} \). We also assume \( \lambda \) is small enough that \( |f''_\alpha| > K^3 \) outside of \( C_\varepsilon \). Together these imply

(iii) \( |f''_\alpha| > K^3 \) outside of \( C_{\frac{1}{2}\sigma} \).

For simplicity of notation, we write \( f = f_\alpha \) in the rest of this subsection.

**Lemma 5.1** Let \( c \in C \) be such that \( f^n(c) \notin C_{\sigma} \) \( \forall n > 0 \). Consider \( x \) with \( |x - c| < \frac{1}{2}\sigma \), and let \( n(x) \) be the smallest \( n \) such that \( |f^n(x) - f^n(c)| > \frac{1}{3K_0}K^3\lambda \). Then \( n(x) > 1 \) and \( |(f^n(x))'| \geq k_3K^n(x) \) for some \( k_3 = k_3(K_0, k_2) \).

Before giving the proof of this lemma, we first prove a distortion estimate.

**Sublemma 5.1** Let \( x, y \in S^1 \) and \( n \in \mathbb{Z}^+ \) be such that \( \omega_i \), the segment between \( f^ix \) and \( f^iy \), satisfies \( |\omega_i| < \frac{1}{3K_0}K^3\lambda \) and \( \text{dist}(\omega_i, C) > \frac{1}{2}\sigma \) for all \( i \) with \( 0 \leq i < n \). Then

\[
\frac{(f^n)'x}{(f^n)'y} \leq 2.
\]

**Proof:**

\[
\log \frac{(f^n)'x}{(f^n)'y} = \sum_{i=0}^{n-1} \log \frac{f'(f^i x)}{f'(f^i y)} \leq \sum_{i=0}^{n-1} \frac{|f'(f^i x) - f'(f^i y)|}{|f'(f^i y)|}.
\]

\[
\leq \sum_{i=0}^{n-1} \frac{(1 + \frac{K_0}{K}) |f^i x - f^i y|}{K^3} < (1 + \frac{K_0}{K}) \left( \sum_{i=0}^{n-1} \frac{1}{K^3} \right) |f^{n-1} x - f^{n-1} y|.
\]

Assuming that \( \frac{1}{\lambda} \) and \( K \) are sufficiently large, this is \( < \frac{1}{7} \). \( \Diamond \)
Proof of Lemma 5.1 First we show $n(x) > 1$. Given the location of $x$, we have

$$K > |f'(x)| = |f''(\xi)| |x - c|$$

for some $\xi$ between $x$ and $c$. This implies

$$|f(x) - fc| = \frac{1}{2} |f''(\zeta)||x - c|^2 < \frac{1}{2} \frac{|f''(\zeta)|}{|f''(\xi)|^2} K^2$$

which we may assume is $< \frac{1}{3K_0} K^3 \lambda$. For $n(x) = 2$, use

$$|(f^2)'(x)| \cdot |x - c| \geq \text{const} \ K^3 \lambda \quad \text{and} \quad |x - c| < \text{const} \ K \lambda .$$

We assume from here on that $n = n(x) \geq 3$, and estimate $|(f^n)'(x)|$ as follows. Since $|f^n x - f^n c| > \frac{1}{3K_0} K^3 \lambda$, it follows from Sublemma 5.1 that for some $\xi_1$,

$$\frac{1}{2} |f''(\xi_1)||x - c|^2 \cdot 2|(f^{n-1})'(fc)| > \frac{1}{3K_0} K^3 \lambda . \quad (12)$$

Reversing the inequality at time $n - 1$ and using Sublemma 5.1 again, we have

$$\frac{1}{2} |f''(\xi_2)||x - c|^2 \cdot \frac{1}{2})|(f^{n-2})'(fc)| < \frac{1}{3K_0} K^3 \lambda . \quad (13)$$

Substituting the estimate for $|(f^{n-1})'(fc)|$ from (12) into

$$|(f^n)'(x)| \geq |f''(\zeta)||x - c| \cdot \frac{1}{2} |(f^{n-1})'(fc)|,$$

we obtain

$$|(f^n)'| \geq \frac{1}{2} \frac{|f''(\zeta)|}{|f''(\xi_1)|} \frac{1}{2K_0} K^3 \lambda \frac{1}{|x - c|} .$$

Now plug the estimate for $|x - c|$ from (13) into the last inequality and use the lower bounds for $|f''(\xi_2)|$ and $|(f^{n-2})'(fc)|$ from (ii) and (iii) earlier on in this subsection. We arrive at the estimate

$$|(f^n)'| > \frac{1}{2} \frac{|f''(\zeta)|}{|f''(\xi_1)|} \frac{1}{3K_0} K^3 \lambda \sqrt{\frac{\lambda}{4 \frac{1}{5K_0} K^3 \lambda}} = \text{const} \ K^{\frac{3}{4}(n-2)+\frac{3}{2}} .$$

The power to which $K$ is raised is $\geq n$ for $n \geq 3$. This completes the proof of Lemma 5.1.

We have proved the following: Suppose $f_\sigma$ has the property that each of its critical points $c$ satisfies $f_\sigma'(c) \not\in C_\sigma$ for all $n > 0$. Then (C2'){(i) and (iii) hold for $f_\sigma$ with $I = C_\sigma$. This follows from properties (ii) and (iii) in the first part of this subsection and from Lemma 5.1.
5.3 Verification of (C2'): "multiple Misiurewicz points"

The goal of this section is to show that for many values of the parameter \( a, f_a \) has the property that its critical orbits (in strictly positive time) stay away from its critical set. Precise statements will be formulated later. We remark that for the quadratic family \( x \mapsto 1 - ax^2 \) or any other family with a single critical point, this is a trivial exercise: there are many periodic orbits or compact invariant Cantor sets \( \Lambda \) disjoint from the critical set, and if changes in parameter correspond to the movement of \( f_a(c) \) in a reasonable way, then there would be many parameters for which \( f_a(c) \in \Lambda \). We call these parameters "Misiurewicz points". For maps with more than one critical point, as circle maps necessarily are, the required condition is that all of the critical orbits are trapped in some invariant set away from \( C \). This is clearly more problematic, especially with \( \Lambda \) having measure zero. We call parameters with these properties "multiple Misiurewicz points". Their existence and \( O(\lambda) \)-density within the family \( \{f_a\} \) is the concern of this subsection.

Recall that \( \sigma = 2k_2^{-1} K^3 \lambda \) and \( C_\sigma \) is the \( \sigma \)-neighborhood of \( C \). Recall also from Sect. 5.2 that outside of \( C_\sigma \), \( |f_a'| > K^3 \). We are looking for a parameter \( a^* \) such that \( f = f_{a^*} \) has the property that for all \( c \in C \), \( f^n c \not\in C_\sigma \) \( \forall n > 0 \). Write \( C = \{x_1, \cdots, x_{k_1}\} \) as before, and let \( \Delta \) be a parameter interval. For \( k = 1, 2, \cdots, k_1 \) and \( i = 1, 2, \cdots \), we introduce the curves of critical points

\[
a \mapsto \gamma_i^{(k)}(a) := f_a^i(x_k), \quad a \in \Delta.
\]

Observe that for all \( k \), \( \frac{d}{da} \gamma_1^{(k)} = 1 \), and and for all \( i \),

\[
\frac{d}{da} \gamma_{i+1}^{(k)}(a) = \frac{d}{da} \gamma_i^{(k)}(a) f_a'(\gamma_i^{(k)}(a)) + 1.
\]

Thus if \( \gamma_j^{(k)}(a) \not\in C_\sigma \) for all \( j \leq i \) and \( K \) is sufficiently large, then

\[
\frac{d}{da} \gamma_{i+1}^{(k)}(a) \approx \frac{d}{da} \gamma_i^{(k)}(a) f_a'(\gamma_i^{(k)}(a)) \quad (14)
\]

and

\[
\frac{d}{da} \gamma_{i+1}^{(k)}(a) \geq \frac{1}{2} K^{3i}. \quad (15)
\]

We also have the following distortion estimate:

**Sublemma 5.2** For \( k = 1, 2, \cdots, k_1 \) and \( n \in \mathbb{Z}^+ \), let \( \Delta \subset [0, 1) \) be such that \( \gamma_{i}^{(k)}(a) \not\in C_\sigma \) for \( i = 1, 2, \cdots, n - 1 \). Assume that \( |\gamma_{n-1}^{(k)}| \leq \frac{1}{3K_0} K^3 \lambda \). Then for all \( a, a' \in \Delta \), we have

\[
\left| \frac{d}{da} \gamma_{n}^{(k)}(a) \right| \leq 2.
\]
Using (14) and (15), we see that the proof is entirely parallel to that of Sublemma 5.1 with slightly weaker estimates. We leave it as an exercise for the reader.

Let \( d \) be the minimum distance between critical points. Choosing \( \lambda \) sufficiently small, we may assume \( 6k_1\sigma << d \). The following is the main result of this subsection.

**Lemma 5.2** Given \( \Delta_0 \subset [0,1] \) with \( |\Delta_0| = 6k_1\sigma \), there exists \( a^* \in \Delta_0 \) such that \( \forall \ c \in C, \ f_{a^n} c \not\in C_\sigma \ \forall n > 0 \).

**Proof:** We describe first an algorithm for selecting a sequence of intervals \( \Delta_0 \supset \Delta_1 \supset \Delta_2 \supset \cdots \) so that \( a^* \in \cap_i \Delta_i \) has the desired property:

At step \( n \), the \((k_1 + 1)\)-tuple \( (\Delta_n; i_{1,n}, i_{2,n}, \cdots, i_{k_1,n}) \) is called an “admissible configuration” if \( \Delta_n \) is a subinterval of \( \Delta_0 \), \( i_{k,n} \leq n \), and the following conditions are satisfied for each \( k \):

(A1) \( \gamma^{(k)}_{i,n} \cap C_\sigma = \emptyset \) for all \( i \leq i_{k,n} \);

(A2) for all \( a, a' \in \Delta_n \), \[
\left| \frac{d}{da} \gamma^{(k)}_{i_{k,n}}(a) \right| \leq 2 .
\]

(A3) (“minimum length condition”) \( |\gamma^{(k)}_{i_{k,n+1}}|_{\Delta_n} \geq 12k_1\sigma \).

Observe that (A3) is about the length of the critical curve one iterate later.

Let us first show that we have an admissible configuration for \( n = 1 \). Let \( i_{k,1} = 1 \) for all \( k \). The parameter interval \( \Delta_1 \) is chosen as follows. Since \( \frac{d}{da} \gamma^{(k)}_1 = 1 \), we have \( |\gamma^{(k)}_1|_{\Delta_0} = 6k_1\sigma \), so that \( \gamma^{(k)}_1 \) meets at most one component of \( C_\sigma \) and \( |(\gamma^{(k)}_1)^{-1} C_\sigma| \leq 2\sigma \). Even in the worst case scenario when all \( k_1 \) intervals \( (\gamma^{(k)}_1)^{-1} C_\sigma \) are evenly spaced, there exists an interval \( \Delta_1 \subset \Delta_0 \) with \( |\Delta_1| = 2\sigma \) such that \( \gamma^{(k)}_1|_{\Delta_1} \cap C_\sigma = \emptyset \) for all \( k \). (A1) and (A2) are trivially satisfied, as is (A3) since \( |\gamma^{(k)}_2|_{\Delta_1} > 2\sigma K^3 \), and \( 2K^3 \) is assumed to be \( > 12k_1 \).

We now discuss how to proceed at a generic step, i.e. step \( n \), assuming we are handed an admissible configuration \( (\Delta_n; i_{1,n}, i_{2,n}, \cdots, i_{k_1,n}) \). First, we divide the set \{1,2,\cdots, k_1\} into indices \( k \) that are “ready to advance”, meaning the situation is right for the \( k \)th curve to progress to the next iterate, and those that are not. Say \( k \in \mathcal{A} \) if

(A4) \( |\gamma^{(k)}_{i_{k,n}}|_{\Delta_n} < \frac{1}{3K_0} K^3 \lambda \) (distortion estimate holds for the next iterate);

(A5) \( |\gamma^{(k)}_{i_{k,n+1}}|_{\Delta_n} < d \). (image of the next iterate meets at most one interval in \( C_\sigma \))
Consider first the case where \( \mathcal{A} \neq \emptyset \). We set \( i_{k,n+1} = i_{k,n} + 1 \) for \( k \in \mathcal{A} \), \( i_{k,n+1} = i_{k,n} \) otherwise, and look for \( \Delta_{n+1} \subset \Delta_n \) so that \( (\Delta_{n+1}; i_{1,n+1}, \ldots, i_{k_1,n+1}) \) is again an admissible configuration.

Let \( k \in \mathcal{A} \). By virtue of (A3) and (A5), we have \( 12k_1\sigma < |\gamma^{(k)}_{i_{k,n+1}}|_{\Delta_n} < d \), so that the fraction of \( \gamma^{(k)}_{i_{k,n+1}}|_{\Delta_n} \) in \( C_\sigma \) is \( \leq \frac{1}{3k_1} \). By virtue of (A4) and Sublemma 5.2, we have good control of the distortion of \( a \mapsto \frac{d}{da} \gamma^{(k)}_{i_{k,n+1}} \). Together this gives

\[
|\gamma^{(k)}_{i_{k,n+1}}|_{\Delta_n} \leq \frac{1}{3k_1} |\Delta_n|.
\] (16)

By the same geometric argument as in the case \( n = 1 \), there exists a subinterval \( \Delta_{n+1} \subset \Delta_n \) of length \( \frac{1}{3k_1} |\Delta_n| \) with the property that \( \gamma^{(k)}_{i_{k,n+1}}|_{\Delta_{n+1}} \cap C_\sigma = \emptyset \) for all \( k \in \mathcal{A} \). For this choice of \( \Delta_{n+1} \), we have (A1) by design, and (A2) is given by (A4) from step \( n \). As for (A3), observe that by the same reasoning as in (16), the pullback of any interval of \( S^1 \) of length \( 2\sigma \) has length \( \leq \frac{1}{3k_1} |\Delta_n| \), so \( \gamma^{(k)}_{i_{k,n+1}}|_{\Delta_{n+1}} \geq 2\sigma \), and one iterate later, it is guaranteed to have length \( > 2K^3\sigma \).

Consider now \( k \notin \mathcal{A} \). (A1) and (A2) are inherited from the previous step, and (A3) is checked as follows: If \( k \notin \mathcal{A} \) because (A4) fails, then

\[
|\gamma^{(k)}_{i_{k,n+1}}|_{\Delta_n} \geq \frac{1}{2} \cdot \frac{1}{3k_1} |\gamma^{(k)}_{i_{k,n}}|_{\Delta_n} \geq cK^3\lambda
\]

where \( c \) is a constant independent of \( K \) of \( \lambda \). Notice that this uses only the distortion estimate from step \( n \). One iterate later, this curve will have length \( > cK^6\lambda \), which we may assume is \( > 12k_1\sigma \). If (A4) holds but (A5) fails, then the distortion estimate holds for the next iterate, and

\[
|\gamma^{(k)}_{i_{k,n+1}}|_{\Delta_n} \geq \frac{1}{6k_1} |\gamma^{(k)}_{i_{k,n+1}}|_{\Delta_n} \geq cd
\]

which we may also assume is \( > 12k_1\sigma \). This completes the construction from step \( n \) to step \( n + 1 \) when \( \mathcal{A} \neq \emptyset \).

If \( \mathcal{A} = \emptyset \), then we let \( \Delta'_n \) be the left half of \( \Delta_n \), and observe that the \((n + 1)\)-tuple \( (\Delta'_n; i_{1,n}, i_{2,n}, \ldots, i_{k_1,n}) \) is again admissible. To verify (A3), we fix \( k \), and argue separately as in the last paragraph the two cases corresponding to (i) the failure of (A4) with respect to \( \Delta_n \) and (ii) the failure of (A5) but not (A4). Repeat this process if necessary until \( \mathcal{A} \neq \emptyset \).

\[ \square \]

### 5.4 Verification of (C1), (C3) and (C4)

We now verify the remaining conditions in Sect. 5.1. Observe from the arguments below that (C1) and (C3)(ii) are quite natural for systems arising from differential equations, while (C3)(i) and (C4) are, to a large extent, consequences of the fact that the maps \( f_a \) are sufficiently expanding.

30
**Verification of (C1):** Let $F_{t_0}$ denote the time-$t_0$-map of (2) (the period of the forcing continues to be $T$). Then (i) follows from the fact that $F_{t_0}$ has bounded $C^3$ norms on $S^1 \times [-1, 1]$; (ii) is obvious, and (iii) is a consequence of the fact that
$$\det(DF_T) = e^{-\lambda(T-t_0)} \det(DF_{t_0}).$$

**Verification of (C3):** For (i), since $(\partial_a f_a)(\cdot) = 1$ and $|(f^k)'(f_x)| \geq K^k$, Lemma 4.2 applies, and the quantity in question has absolute value $\geq 1 - \sum_{i \geq 1} \frac{1}{K^i} > 0$. Part (ii) is Lemma 2.3(i).

**Verification of (C4):** (i) is proved since $e^{c_1} = K > 2$. For (ii), by choosing $\lambda$ sufficiently small depending on $\Phi_0$, it is easily arranged that $p_{i,j} = 1$ for all $i,j$.

This completes the proof of Theorem 3.

**Appendix**

We supply here the proofs of the two lemmas promised in Sect. 5.1. This Appendix has to be be read in conjunction with [WY].

**Lemma A.1** *All the theorems in [WY] remain valid if the Misiurewicz condition in Step I, Sect. 1.1, of [WY] is replaced by condition (C2') in Sect. 5.1 of this paper.*

**Proof** The three most important uses of the Misiurewicz condition in [WY] are:

- the nondegeneracy of the critical points (this is guaranteed by (C2')(iii)(a));
- every critical orbit stays a fixed distance away from $C$ (this is precisely (C2')(ii));
- there exist $c_0, c > 0$ such that for every critical point $x$, $|(f^n)'(f_x)| > c_0 e^{cn}$ (this is guaranteed by (C2')(i) and (ii)).

These three properties aside, the only consequences of the Misiurewicz condition used in [WY] are contained in

**Lemma 2.5 of [WY]** *Let $C_\delta$ denote the $\delta$-neighborhood of $C$. Then there exist $\tilde{c}_0, \tilde{c}_1 > 0$ such that the following hold for all sufficiently small $\delta > 0$: Let $x \in S^1$ be such that $x, f_x, \ldots, f^{n-1}_x \notin C_\delta$, any $n$. Then*

(i) $|(f^n)'x| \geq \tilde{c}_0 \delta e^{\tilde{c}_1 n}$;

(ii) if, in addition, $f^n x \in C_\delta$, then $|(f^n)'x| \geq \tilde{c}_0 e^{\tilde{c}_1 n}$.

We claim that the conclusions of this lemma also follow from (C2'). Let $n_1 < \cdots < n_q, 0 \leq n_1, n_q \leq n$, be the times when $f^{n_i} x \in I$. Then
\[ |(f^{n_1})'x| \geq e^{c_1 n_1} \quad \text{by (C2')(i)(b)}; \]
\[ |(f^{n_{i+1}-n_i})'(f^{n_i}x)| \geq e^{c_1(n_{i+1}-n_i)} \quad \text{by (C2')(iii)(b) followed by (i)(b)}; \]
\[ |(f^{n-n_q})'(f^{n_q}x)| = |f'(f^{n_q}x)| \cdot |(f^{n-(n_q+1)})'(f^{n_q+1}x)| \]
where \[ |f'(f^{n_q}x)| \geq |f''(\xi)| \, d(x, C) \geq c_0 \delta \quad \text{by (C2')(iii)(a)} \]
and \[ |(f^{n-(n_q+1)})'(f^{n_q+1}x)| \geq c_0^q e^{c_1(n-(n_q+1))} \quad \text{by (C2')(i)(a)}. \]
Together these inequalities prove both of the assertions in the lemma. \(\square\)

**Lemma A.2** Let \(\{T_{a,b}\}\) be as in Sect. 5.1 of this paper, and let \(\Delta\) be the set of \((a, b)\) such that \(T = T_{a,b}\) satisfies the conclusions of Theorem 1 in [WY]. Suppose \(\{T_{a,b}\}\) also satisfies (C4), and \(\delta\) is smaller than a number depending on \(c_1\). Then

(i) \(T\) admits at most one SRB measure \(\mu\); 
(ii) \((T, \mu)\) is mixing.

**Proof:** Let \(\{x_1 < \cdots < x_r\}\) be the set of critical points of \(f\). Consider a segment \(\omega \subset \partial R_0\) corresponding to an outermost \(I_{\mu^j}\) at one of the components of \(C^{(0)}\). First we claim there exist \(N \in \mathbb{Z}^+\) and \(\omega' \subset \omega\) such that \(T^i \omega \cap C^{(0)} = \emptyset\) for all \(0 < i < N\) and \(T^N \omega'\) connects two components of \(C^{(0)}\).

This claim is proved as follows. Let \(\omega'\) denote the image of \(\omega\) at the end of its bound period. Then \(\omega'\) has length \(> \delta K^\beta\). We continue to iterate, deleting all parts that fall into \(C^{(0)}\). Then \(i\) steps later, the undeleted part of \(T^i \omega'\) is made up of finitely many segments. Suppose that for all \(i \leq n\), none of these segments is long enough to connect two components of \(C^{(0)}\), so that the number of segments deleted up to step \(i\) is \(\leq 2^i\). We estimate the average length of these segments at time \(n\) as follows: First, the pull-back to \(\omega'\) of all the deleted parts has total measure \(\leq \sum_{i \leq n} 2^i e^{-c_1 i (2\delta)}\) by (C2')(i)(b). Since \(2 < e^{c_1}\) by (C4)(i), we may assume this is \(< \frac{1}{2} \delta K^\beta\) provided \(\delta\) is sufficiently small. The undeleted segments of \(T^n \omega\) add up, therefore, to \(> e^{c_1 n \frac{1}{2} \delta K^\beta}\) in length, and since there are at most \(2^n\) of them, their average length is \(> 2^{-n} e^{c_1 n \frac{1}{2} \delta K^\beta}\). Thus one sees that as \(n\) increases, there must come a point when our claim is fulfilled.

Next we observe that if \(\omega\) is a \(C^2(b)\) segment connecting two components of \(C^{(0)}\), then using (C4)(ii) and reasoning as with finite state Markov chains, we have that for every \(n \geq N_2\) and every \(k \in \{1, \cdots, r\}\), there is a subsegment \(\omega_{n,k} \subset \omega\) such that for all \(i < n\), \(T^i \omega_{n,k} \cap C^{(0)} = \emptyset\) and \(T^n \omega_{n,k}\) stretches across the region between \(x_k\) and \(x_{k+1}\), extending beyond the critical regions containing these two points.

Recall that in [WY], Sects. 8.1 and 8.2, a finite number of ergodic SRB measures \(\{\mu_i; i \leq r'\}\) are constructed, and it is shown in Sect. 8.3 that these are all the ergodic SRB measures \(T\) has. The discussion above shows that starting at any reference set, a segment \(\omega \subset \partial R_0\) as above will spend a positive fraction of time in every reference set, proving that \(r' \leq 1\). Furthermore, starting from any reference set, the return time to it takes on all values greater than some \(N_0\), proving that \(\mu_1\) is mixing. \(\square\)
6 Concluding Remarks

- For area-preserving maps, it is well known that when integrability first breaks down, the phase portrait is dominated by KAM curves. Farther away from integrability, one sees larger Birkhoff zones of instability interspersed with elliptic islands. Continuing to move toward the chaotic end of the spectrum, it is widely believed – though not proved – that most of the phase space is covered with ergodic regions with positive Lyapunov exponents.

This paper deals with the corresponding pictures for strongly dissipative systems. We consider a simple model consisting of a periodically forced limit cycle. Keeping the magnitude of the “kick” constant, we prove that scenarios roughly parallel to those in the last paragraph occur for our Poincaré maps, with attracting invariant circles (taking the place of KAM curves), periodic sinks (instead of elliptic islands), and as the contractive power of the cycle diminishes, we prove that the stage is shared by at least two scenarios occupying parameter sets that are delicately intertwined: horseshoes and sinks, and strange attractors.

By “strange attractors”, we refer to attractors characterized by SRB measures, positive Lyapunov exponents, and strong mixing properties. For the differential equation in question, we prove that the system has global strange attractors of this kind for a positive measure set of parameters.

- Our second point has to do with bridging the gap between abstract theory and concrete problems. Today we have a fairly good hyperbolic theory, yet chaotic phenomena in naturally occurring dynamical systems have continued to resist analysis. One of the messages of this paper is that for certain types of strange attractors, the situation is now improved: For attractors with strong dissipation and one direction of instability, there are now relatively simple, checkable conditions which, when satisfied, guarantee the existence of an attractor with a detailed package of statistical and geometric properties. Our conditions are formulated to give rigorous results, but where rigorous analysis is out of reach, they can also serve as a basis for numerical work to provide justification for various mathematical statements about strange attractors.
References


