FROM LIMIT CYCLES TO STRANGE ATTRACTORS

WILLIAM OTT AND MIKKO STENLUND

ABSTRACT. We define a quantitative notion of shear for limit cycles of flows. We prove that strange attractors and SRB measures emerge when systems exhibiting limit cycles with sufficient shear are subjected to periodic pulsatile drives. The strange attractors possess a number of precisely-defined dynamical properties that together imply chaos that is both sustained in time and physically observable.

1. INTRODUCTION

This paper is about a mechanism for producing chaos: shear. We are guided by the idea that in the presence of shear, a stable dynamical structure can be transformed into a strange attractor with strong stochastic properties by forcing the structure with a pulsatile drive. The forcing does not overwhelm the intrinsic dynamics. Instead, it acts as an amplifier, amplifying the effects of the intrinsic shear. We focus on one particular dynamical structure of great importance: the limit cycle. Limit cycles are asymptotically stable periodic orbits of flows on Riemannian manifolds.

The application of a periodic pulsatile drive to a flow exhibiting a limit cycle causes deformations to occur. If shear is present in a neighborhood of the limit cycle, if the limit cycle only weakly attracts nearby orbits, and if the time between pulses (the relaxation time) is sufficiently large, then stretch-and-fold geometry emerges in a neighborhood of the limit cycle. Stretch-and-fold geometry suggests that chaotic behavior that is both sustained in time and observable *may* exist. We prove that such chaotic behavior does exist in a certain parameter regime for *any* (generic) forcing function if the shear is sufficiently strong. Moreover, we define a quantity called the shear integral that quantifies the amount of shear that is present in the intrinsic flow in a neighborhood of the limit cycle. We emphasize that the shear integral depends only on the intrinsic system and not on the external forcing. Our result is the first of its kind for general limit cycles. Wang and Young [16, 17] obtain results of a similar flavor for supercritical Hopf bifurcations and certain linear models.

The search for and analysis of stochastic behavior in deterministic dynamical systems have played a major role in guiding dynamical systems research. We discuss a few relevant developments. The theory of uniformly hyperbolic systems is well-developed. Let M be a compact Riemannian manifold and let $f : M \to M$ be a C^2 diffeomorphism of M. An attractor for f is a compact set Ω satisfying $f(\Omega) = \Omega$ for which there exists an open set $U \subset M$ (the basin) such that $f(\bar{U}) \subset U$ and $\Omega = \bigcap_{i=0}^{\infty} f^i(\bar{U})$. An attractor Ω is said to be an Axiom A attractor if the tangent bundle over Ω splits into 2 Df-invariant subbundles E^s and E^u such that vectors in E^s are contracted by Df and vectors in E^u are expanded by Df (we assume E^u is nontrivial). An Axiom A attractor supports a special invariant measure known as a Sinai-Ruelle-Bowen (SRB) measure that describes the asymptotic distribution of the orbit of almost every point in U with respect to Riemannian volume and has strong stochastic properties. In this sense, the chaotic behavior associated with Axiom A systems is observable. It is also sustained in time because of the presence of positive Lyapunov exponent(s). One can, in principle, detect the presence of uniform hyperbolicity in a given system by finding invariant cone families with suitable properties. For example, Tucker uses this approach to prove that the Lorenz equations are chaotic for the classical parameter values studied by Lorenz [13].

Many systems of interest in the biological and physical sciences display some form of hyperbolicity but are not uniformly hyperbolic. A mature theory of nonuniform hyperbolicity has emerged over the last 4 decades. However, the following problem remains a challenge. Given a dynamical system (or a parametrized family of dynamical systems), how can nonuniform hyperbolicity be detected? Numerical techniques include the calculation of Lyapunov exponents and the 0-1 test [4, 5]. This paper addresses the analytical component

of the problem in the context of limit cycles. Our proofs are based on the recently-developed theory of rank one maps [15, 18]. Rank one theory is based on the ideas of Jakobson [8], Benedicks and Carleson [1, 2], and Young [19, 20]. Rank one theory provides checkable conditions that imply the existence of SRB measures with strong stochastic properties in parametrized families of diffeomorphisms.

We conclude the introduction with a remark that the results obtained in this paper are in some sense dual to the phenomenon known as self-induced stochastic resonance (SISR) (see *e.g.* [3]). Our results demonstrate that certain intrinsic characteristics of a deterministic system (shear) can produce stochastic-type behavior when the system is forced in a deterministic way. SISR demonstrates that underlying phase space structures can produce deterministic (coherent) behavior in stochastically-forced systems when the noise level is taken to 0 along certain distinguished limits.

2. STATEMENT OF RESULTS

We state the main results and discuss their relationship to the existing literature. Let $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n$ be a C^5 vector field and consider the differential equation

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(\boldsymbol{x}). \tag{2.1}$$

We assume that (2.1) admits an asymptotically stable hyperbolic periodic solution $\boldsymbol{\eta}$ of length L and period p_0. Let $\boldsymbol{\gamma} : \mathbb{R} \to \mathbb{R}^n$ be a function of the parameter s that parametrizes $\boldsymbol{\eta}$ by length. Define $\Gamma = \{\boldsymbol{\gamma}(s) : s \in [0, L]\}$. Solutions to (2.1) that begin sufficiently close to Γ will converge to Γ at an exponential rate as $t \to \infty$. We are interested in the effects of adding periodic pulsatile forcing to the vector field defining (2.1). For $0 < \rho < T$, define the periodic function $P_{\rho,T} : \mathbb{R} \to \mathbb{R}$ as follows. For $0 \leqslant t \leqslant T$, set

$$P_{\rho,T}(t) = \begin{cases} 1, & \text{if } 0 \leqslant t \leqslant \rho \\ 0, & \text{if } \rho < t < T \end{cases}$$

and then extend periodically to all $t \in \mathbb{R}$ by requiring $P_{\rho,T}(t+T) = P_{\rho,T}(t)$. We study the externally-forced system

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(\boldsymbol{x}) + \varepsilon P_{\rho,T}(t) \boldsymbol{F}(\boldsymbol{x}) \tag{2.2}$$

where $\boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^n$ is a C^4 vector field and the parameter $\varepsilon > 0$ controls the amplitude of the forcing.

In Section 3 we compute a normal form of equation (2.2) that is valid in a tubular neighborhood $\tilde{M} \approx \Gamma \times D$, where D is a closed disk in $\mathbb{R}^{n-1}$ of sufficiently small radius. We are interested in the dynamics of (2.2) in the tubular neighborhood $M \approx \Gamma \times \frac{1}{2}D$. Since the external forcing is periodic with period T, it is natural to study the time-T map induced by (2.2). We write the time-T map as the composition of a **kick map** $H_k : M \to \tilde{M}$ and a **relaxation map** $H_r : \tilde{M} \to \text{int}(M)$. Let H_k be the time-ρ map induced by the flow associated with (2.2). Notice that the external forcing is active during the kick phase because $P_{\rho,T}(t) = 1$ for $0 \leqslant t \leqslant \rho$. For ε sufficiently small, H_k maps M into $\tilde{M}$ diffeomorphically. Let H_r be the time-$(T - \rho)$ map induced by (2.2) with ε set to 0. There exists $T_0 = T_0(\varepsilon)$ such that if $T \geqslant T_0$, then H_r maps $\tilde{M}$ into $\text{int}(M)$. The composition $G_T := H_r \circ H_k$ is the time-T map induced by (2.2).

The dynamical properties of $G_T : M \to \text{int}(M)$ depend on a number of factors. One feature common to every map G_T for $T \geqslant T_0$ is the existence of an attractor Ω defined by

$$\Omega = \bigcap_{i=0}^{\infty} G_T^i(M).$$

We call $U := \text{int}(M)$ the basin of attraction of Ω. For every $\boldsymbol{x} \in U$, $G_T^i(\boldsymbol{x}) \to \Omega$ as $i \to \infty$. Two characteristics of the intrinsic system (2.1) play a key role in determining the structure of Ω and the dynamical properties of G_T: shear and the strength of the limit cycle. We quantify these notions momentarily; for now, imagine that (2.1) exhibits strong shear in M if for most points $\boldsymbol{x} \in \Gamma$, the velocity vector $\boldsymbol{f}(\hat{\boldsymbol{x}})$ varies substantially as $\hat{\boldsymbol{x}}$ moves away from $\boldsymbol{x}$ in directions orthogonal to the limit cycle Γ. Think of the limit cycle Γ as strongly stable if solutions to (2.1) that begin in M converge quickly to Γ. If the shear is weak and the limit cycle is strongly stable, then the attractor Ω associated with G_T will be an invariant closed curve. We are interested in the opposite situation. Suppose that the shear is strong in M and the limit cycle is weakly

stable. The addition of the periodic pulsatile external force $\varepsilon P_{\rho,T}(t)\boldsymbol{F}(\boldsymbol{x})$ will amplify the effect of the shear in the following way: disturbances that are created when $P_{\rho,T} = 1$ will be stretched during the relaxation period (when $P_{\rho,T} = 0$). The stretching effect increases in intensity as T increases. If T is large, then folds will be created in the phase space. If G_T exhibits stretch-and-fold geometry, then G_T potentially exhibits chaotic behavior that is sustained in time and observable.

This paper aims to accomplish the following.

(1) We define a computable quantity called the shear integral that quantifies the shear associated with the intrinsic system (2.1) near the limit cycle Γ.
(2) We prove that if the magnitude of the shear integral is sufficiently large and if the contraction near the limit cycle Γ is sufficiently weak, then the following holds for suitable values of ε. For a typical external vector field $\boldsymbol{F}$, there exists $T_1 > 0$ and a set $\Delta \subset [T_1, \infty)$ of positive Lebesgue measure such that for $T \in \Delta$, the time-T map G_T associated with (2.2) admits a strange attractor Ω and exhibits chaos that is sustained in time and observable.

The quantity T_1 satisfies $T_1 \gg \rho$, ensuring sufficient relaxation time for the stretch-and-fold geometry to emerge. The term ***strange attractor*** refers to a number of precisely defined dynamical and structural properties that represent sustained, observable chaos. For $T \in \Delta$, Ω supports a unique ergodic SRB measure ν. Here the term SRB measure refers to a measure ν with a positive Lyapunov exponent ν almost everywhere and whose conditional measures on unstable manifolds are absolutely continuous with respect to Riemannian volume on these manifolds. The SRB measure ν satisfies the central limit theorem and exhibits exponential decay of correlations for Hölder continuous observables. For Lebesgue almost every $\boldsymbol{x}$ in the basin of attraction U, the orbit of $\boldsymbol{x}$ has a positive Lyapunov exponent and is asymptotically distributed according to ν in the sense that for every continuous function $\varphi : U \to \mathbb{R}$, we have

$$(2.3) \qquad \lim_{m \to \infty} \frac{1}{m} \sum_{i=0}^{m-1} \varphi(G_T^i(\boldsymbol{x})) = \int \varphi \, d\nu.$$

Notice that this statement is substantially stronger than the conclusion of the Birkhoff ergodic theorem. The Birkhoff ergodic theorem implies that (2.3) holds for ν almost every $\boldsymbol{x}$. However, ν is singular with respect to Lebesgue measure (supported on a set of Lebesgue measure zero) because the dynamics are dissipative. We prove that (2.3) holds for Lebesgue almost every $\boldsymbol{x} \in U$. See **(SA1)**–**(SA4)** in Section 4 for a more precise description of the dynamical properties of G_T for $T \in \Delta$.

We now define the shear integral. In Section 3 we derive a normal form of (2.1) that is valid in $\tilde{M}$. The normal form, expressed in the natural $(s, \boldsymbol{z})$-coordinates introduced in Section 3.1, is given by

$$(2.4\text{a}) \qquad \frac{dt}{ds} = \|\boldsymbol{f}(\boldsymbol{\gamma}(s))\|^{-1} + \langle \boldsymbol{\beta}(s), \boldsymbol{z} \rangle + \omega_1(s, \boldsymbol{z})$$

$$(2.4\text{b}) \qquad \frac{d\boldsymbol{z}}{ds} = \mathsf{A}\boldsymbol{z} + \boldsymbol{\omega}_2(s, \boldsymbol{z})$$

Here $\langle \cdot, \cdot \rangle$ denotes the inner product on $\mathbb{R}^{n-1}$. Functions depending on s in (2.4a)–(2.4b) are periodic in s with period $2L$. The matrix A is in Jordan canonical form. The functions ω_1 and $\boldsymbol{\omega}_2$ represent higher order corrections. The function $\boldsymbol{\beta}$ gives the pointwise magnitude and direction of the shear. Define the ***shear integral*** $\boldsymbol{\Sigma}$ by

$$\boldsymbol{\Sigma} = (\Sigma_1, \ldots, \Sigma_{n-1}) := \int_0^{2L} \boldsymbol{\beta}(\tau) \, d\tau$$

and define the ***shear factor*** σ by $\sigma := \|\boldsymbol{\Sigma}\|$.

Having defined the shear integral, we describe the setting of the main theorem. We identify intrinsic parameters (parameters associated with $\boldsymbol{f}$) and external parameters (parameters associated with the external forcing). We fix the normalized shear vector $\frac{\boldsymbol{\Sigma}}{\sigma}$ and view the shear factor σ as the first intrinsic parameter. The second intrinsic parameter quantifies the strength of the contraction near the limit cycle and is derived from A. We assume for the sake of simplicity that A is a diagonal matrix given by $\mathsf{A} = \mathrm{diag}(\lambda_1, \ldots, \lambda_{n-1})$ where $0 > \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_{n-1}$ are the eigenvalues of A. We fix the eigenvalue ratios $\mu_i = \frac{\lambda_1}{\lambda_i}$ for $1 \leqslant i \leqslant n-1$ and we view the weakest eigenvalue λ_1 as an intrinsic parameter. The only external parameter

is ε, the factor that controls the amplitude of the external forcing. We fix $\rho > 0$. A key parameter derived from ε, σ, and λ_1 is the **hyperbolicity factor** $\frac{\varepsilon\sigma}{|\lambda_1|}$.

One additional ingredient is needed. Even if σ is large and $|\lambda_1|$ is small, a strange attractor cannot emerge unless the forcing $\boldsymbol{F}$ acts in direction(s) in which shear is present. We express this idea by introducing a certain function on the circle $\mathbb{S} := \frac{\mathbb{R}}{2L\mathbb{Z}}$. We identify $\mathbb{S}$ with the interval $[0, 2L)$. In Section 3 we derive a normal form of the forced system (2.2) that is valid in $\tilde{M}$ when the forcing is active ($P_{\rho,T} = 1$):

$$\text{(2.5a)} \qquad \frac{dt}{ds} = \|\boldsymbol{f}(\boldsymbol{\gamma}(s))\|^{-1} + \langle \boldsymbol{\beta}(s), \boldsymbol{z} \rangle + \omega_3(s, \boldsymbol{z})$$

$$\text{(2.5b)} \qquad \frac{d\boldsymbol{z}}{ds} = \mathsf{A}\boldsymbol{z} + \varepsilon\boldsymbol{\zeta}(s) + \boldsymbol{\omega}_4(s, \boldsymbol{z})$$

Functions depending on s in (2.5a)–(2.5b) are periodic in s of period $2L$. The functions ω_3 and $\boldsymbol{\omega}_4$ are higher order corrections. The function $\boldsymbol{\zeta}$ is related to the projection of $\boldsymbol{F}$ in directions orthogonal to Γ. For $s_0 \in \mathbb{S}$, define $\tilde{s}$ implicitly by

$$\rho = \int_{s_0}^{\tilde{s}} \|\boldsymbol{f}(\boldsymbol{\gamma}(\tau))\|^{-1} \, d\tau.$$

Define the vector

$$\boldsymbol{d} := \left(\frac{\Sigma_i \mu_i}{\sigma} \right)_{i=1}^{n-1}$$

and define $\Phi : \mathbb{S} \to \mathbb{R}$ by

$$\text{(2.6)} \qquad \Phi(s_0) = \left\langle \boldsymbol{d}, \int_{s_0}^{\tilde{s}} \boldsymbol{\zeta}(\tau) \, d\tau \right\rangle.$$

We say that Φ is a **Morse function** if the critical set $C(\Phi) = \{s \in \mathbb{S} : \Phi'(s) = 0\}$ is finite and if for every $s \in C(\Phi)$, we have $\Phi''(s) \neq 0$. We are now in position to state the main theorem. In Theorem 1, we assume that the radius of M is $\leqslant \kappa_0 \varepsilon$ for some constant $\kappa_0 > 0$.

Theorem 1. *Let G_T denote the time-T map associated with (2.2). Suppose that the function Φ defined by (2.6) is a Morse function. Then there exist a small constant $\kappa_1 > 0$ and a large constant $\kappa_2 > \kappa_1$ such that the following holds. If*

(1) $|\lambda_1| < \kappa_1$,
(2) $\frac{\varepsilon}{|\lambda_1|} < \kappa_1$,
(3) $\frac{\varepsilon\sigma}{|\lambda_1|} > \kappa_2$,

*then there exists $T_1 > 0$ and a set $\Delta \subset [T_1, \infty)$ of positive Lebesgue measure such that for $T \in \Delta$, G_T admits a strange attractor Ω in M and satisfies **(SA1)**–**(SA4)** from Section 4. For every interval $I \subset [T_1, \infty)$ of length 1, $\ell(\Delta \cap I) > 0$, where ℓ denotes the Lebesgue measure on $\mathbb{R}$.*

Remark 2.1. The assumption that Φ is a Morse function is quite mild and should hold for a typical forcing vector field $\boldsymbol{F}$. We do not formulate precise results of this type in this paper, but such results should hold in terms of both topological genericity and prevalence. Prevalence is a measure-theoretic notion of genericity that generalizes the concept of 'Lebesgue almost every' to infinite-dimensional spaces. It provides a powerful framework for describing generic phenomena in a probabilistic way (see *e.g.* [6, 7, 10]).

Remark 2.2. Theorem 1 concludes that G_T exhibits sustained, observable chaos for a set of values of T of positive Lebesgue measure rather than for all $T \in [T_1, \infty)$. This is not a consequence of the nature of the proof. Rather, it is a fundamental consequence of the fact that an alternate scenario competes with the SRB scenario in the space of T-values. For an open set $\mathcal{S}$ of T-values in $[T_1, \infty)$, the basin U contains a G_T-invariant Cantor set on which G_T is uniformly hyperbolic (a horseshoe) and a periodic sink. The trajectory of Lebesgue almost every $\boldsymbol{x} \in U$ converges to the periodic sink. Thus for $T \in \mathcal{S}$, G_T exhibits transient chaos: a typical trajectory in the basin will move erratically for some time due to the presence of the horseshoe before finally converging to the periodic sink.

Remark 2.3. The function Φ does not depend on the parameters λ_1, σ, and ε.

Theorem 1 is related to 2 results obtained by Wang and Young in [17]. Wang and Young consider limit cycles forced by periodic δ-function kicks. First, they prove that any limit cycle, when suitably kicked, can be transformed into a strange attractor. This result is universal but not constructive. An artificially-strong kick is needed if geometric conditions are unfavorable for the creation of nonuniform hyperbolicity. Second, they prove that the Hopf limit cycle that emerges from a supercritical Hopf bifurcation can be transformed into a strange attractor. Here the so-called twist factor plays the role of the shear integral. Unlike the shear integral, the twist factor is local in the sense that it depends only on derivatives of the vector field at the bifurcation parameter.

Many of the quantities in Theorem 1 are required to be sufficiently large or sufficiently small. This is an unavoidable consequence of the perturbative nature of the analytic techniques used in the proof. However, numerical evidence suggests that shear-induced chaos emerges over parameter ranges that far exceed those to which the rigorous analysis applies. For example, Lin and Young [9] conduct numerical studies of a linear shear flow model previously studied by Zaslavsky [21]. The work of Lin and Young also provides numerical evidence that the temporal form of the kicks need not be periodic: temporally-sustained chaotic behavior is observed for random kicks at Poisson-distributed times and for continuous-time forcing by white noise.

3. Derivation of the singular limit

3.1. Derivation of the normal forms.

We derive the normal forms (2.4a)–(2.4b) and (2.5a)–(2.5b) that are valid in a small neighborhood of Γ. For $s \in \mathbb{S}$, let $\{e_i(s)\}_{i=1}^n$ be an orthonormal basis for $\mathbb{R}^n$ such that $e_n(s) = \gamma'(s)$ (where γ' denotes the derivative of γ with respect to s) and e_i is a C^5 function of s for all $1 \leqslant i \leqslant n$. One may choose the first $n-1$ vectors in many ways. For example, if γ is at least C^{n+5} and the first n derivatives of γ are linearly independent, then one may construct the basis by applying the Gram-Schmidt procedure to the first n derivatives of γ. For any $x \in \mathbb{R}^n$ sufficiently close to Γ, there exist unique $s \in \mathbb{S}$ and $y = (y_1, \ldots, y_{n-1})$ such that

$$(3.1) \qquad x = \gamma(s) + \sum_{i=1}^{n-1} y_i\, e_i(s).$$

We use (s, y) as new phase variables.

Define

$$\mathsf{E}(s) = \begin{pmatrix} (e_1(s))^\mathsf{T} \\ (e_2(s))^\mathsf{T} \\ \vdots \\ (e_n(s))^\mathsf{T} \end{pmatrix}$$

Differentiating $\mathsf{E}(s)$ with respect to s, we have $\mathsf{E}'(s) = \mathsf{K}(s)\mathsf{E}(s)$ where $\mathsf{K}(s) = (k_{j,i}(s))$ is a skew-symmetric matrix of generalized curvatures defined by $k_{j,i}(s) = \langle e_j'(s), e_i(s) \rangle$. If the first n derivatives of γ are used to create E, then this differential equation is the classical Frenet-Serret equation from differential geometry. For $1 \leqslant i \leqslant n$, define the vector

$$\boldsymbol{k}_i(s) = \begin{pmatrix} k_{1,i}(s) \\ k_{2,i}(s) \\ \vdots \\ k_{n-1,i}(s) \end{pmatrix}$$

Differentiating (3.1) with respect to t, we obtain

$$(3.2) \qquad \frac{d\boldsymbol{x}}{dt} = \sum_{i=1}^{n-1} \frac{dy_i}{dt}\, \boldsymbol{e}_i(s) + \frac{ds}{dt}\left(\gamma'(s) + \sum_{j=1}^{n-1} y_j\, \boldsymbol{e}_j'(s)\right) = \boldsymbol{f}(\boldsymbol{x}) + \varepsilon P_{\rho,T}(t)\boldsymbol{F}(\boldsymbol{x}).$$

Taking the inner product of (3.2) with respect to $\boldsymbol{e}_i(s)$ for $1 \leqslant i \leqslant n-1$ yields

$$\frac{dy_i}{dt} = \langle \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{e}_i(s) \rangle + \varepsilon P_{\rho,T}(t)\langle \boldsymbol{F}(\boldsymbol{x}), \boldsymbol{e}_i(s) \rangle - \frac{ds}{dt}\langle \boldsymbol{y}, \boldsymbol{k}_i(s) \rangle.$$

Taking the inner product of (3.2) with respect to $\boldsymbol{e}_n(s)$ yields

$$\frac{ds}{dt}(\langle \boldsymbol{y}, \boldsymbol{k}_n(s)\rangle + 1) = \langle \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{e}_n(s)\rangle + \varepsilon P_{\rho,T}(t)\langle \boldsymbol{F}(\boldsymbol{x}), \boldsymbol{e}_n(s)\rangle.$$

Notice that $\langle \boldsymbol{y}, \boldsymbol{k}_n(s)\rangle + 1 \neq 0$ if $\|\boldsymbol{y}\|$ is sufficiently small. Consequently, the system

(3.3a)
$$\frac{ds}{dt} = \frac{1}{1 + \langle \boldsymbol{y}, \boldsymbol{k}_n(s)\rangle}(\langle \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{e}_n(s)\rangle + \varepsilon P_{\rho,T}(t)\langle \boldsymbol{F}(\boldsymbol{x}), \boldsymbol{e}_n(s)\rangle)$$

(3.3b)
$$\frac{dy_i}{dt} = \langle \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{e}_i(s)\rangle + \varepsilon P_{\rho,T}(t)\langle \boldsymbol{F}(\boldsymbol{x}), \boldsymbol{e}_i(s)\rangle - \frac{ds}{dt}\langle \boldsymbol{y}, \boldsymbol{k}_i(s)\rangle$$

is valid in a small neighborhood of Γ.

We now extract the terms of leading order in (3.3a) and (3.3b). For $1 \leqslant j \leqslant n$, define $\psi_j(s, \boldsymbol{y}) = \langle \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{e}_j(s)\rangle$. For $1 \leqslant i \leqslant n-1$, we have

$$\psi_i(s, \boldsymbol{y}) = \langle \boldsymbol{\psi}_i^{(1)}(s), \boldsymbol{y}\rangle + \mathcal{O}_{s,\boldsymbol{y}}(\|\boldsymbol{y}\|^2)$$

where

$$\boldsymbol{\psi}_i^{(1)}(s) = \left.\frac{\partial\langle \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{e}_i(s)\rangle}{\partial \boldsymbol{y}}\right|_{\boldsymbol{y}=\boldsymbol{0}}.$$

Here $\mathcal{O}_{s,\boldsymbol{y}}(\|\boldsymbol{y}\|^2)$ denotes a function of s and $\boldsymbol{y}$ for which there exists a constant $K > 0$ independent of s and $\boldsymbol{y}$ such that $|\mathcal{O}_{s,\boldsymbol{y}}(\|\boldsymbol{y}\|^2)| \leqslant K\|\boldsymbol{y}\|^2$. Expanding $\psi_n(s, \boldsymbol{y})$, we have

$$\psi_n(s, \boldsymbol{y}) = \psi_n^{(0)}(s) + \langle \boldsymbol{\psi}_n^{(1)}(s), \boldsymbol{y}\rangle + \mathcal{O}_{s,\boldsymbol{y}}(\|\boldsymbol{y}\|^2)$$

where

$$\psi_n^{(0)}(s) = \|\boldsymbol{f}(\boldsymbol{\gamma}(s))\|$$

$$\boldsymbol{\psi}_n^{(1)}(s) = \left.\frac{\partial\langle \boldsymbol{f}(\boldsymbol{x}), \boldsymbol{e}_n(s)\rangle}{\partial \boldsymbol{y}}\right|_{\boldsymbol{y}=\boldsymbol{0}}.$$

Set $\phi_j(s, \boldsymbol{y}) = \langle \boldsymbol{F}(\boldsymbol{x}), \boldsymbol{e}_j(s)\rangle$ for $1 \leqslant j \leqslant n$. Writing (3.3a) and (3.3b) in terms of ψ_j and ϕ_j, when the forcing is active ($P_{\rho,T}(t) = 1$) we obtain

(3.4)
$$\begin{cases} \dfrac{dt}{ds} = \dfrac{1}{\psi_n^{(0)}(s) + \varepsilon\phi_n(s, \boldsymbol{y})}\left(1 + \left[\boldsymbol{k}_n(s) - \dfrac{\boldsymbol{\psi}_n^{(1)}(s)}{\psi_n^{(0)}(s) + \varepsilon\phi_n(s, \boldsymbol{y})}\right] \cdot \boldsymbol{y} + \mathcal{O}_{s,\boldsymbol{y}}(\|\boldsymbol{y}\|^2)\right) \\[4mm] \dfrac{dy_i}{ds} = \dfrac{\varepsilon\phi_i(s, \boldsymbol{y})}{\psi_n^{(0)}(s) + \varepsilon\phi_n(s, \boldsymbol{y})} + \left(\dfrac{\boldsymbol{\psi}_i^{(1)}(s)}{\psi_n^{(0)}(s) + \varepsilon\phi_n(s, \boldsymbol{y})} - \boldsymbol{k}_i(s)\right) \cdot \boldsymbol{y} \\[4mm] \qquad + \left(\dfrac{\varepsilon\phi_i(s, \boldsymbol{y})}{\psi_n^{(0)}(s) + \varepsilon\phi_n(s, \boldsymbol{y})}\right)\left[\boldsymbol{k}_n(s) - \dfrac{\boldsymbol{\psi}_n^{(1)}(s)}{\psi_n^{(0)}(s) + \varepsilon\phi_n(s, \boldsymbol{y})}\right] \cdot \boldsymbol{y} + \mathcal{O}_{s,\boldsymbol{y}}(\|\boldsymbol{y}\|^2) \end{cases}$$

When the forcing is off ($P_{\rho,T}(t) = 0$), we have

(3.5)
$$\begin{cases} \dfrac{dt}{ds} = \dfrac{1}{\psi_n^{(0)}(s)}\left(1 + \left[\boldsymbol{k}_n(s) - \dfrac{\boldsymbol{\psi}_n^{(1)}(s)}{\psi_n^{(0)}(s)}\right] \cdot \boldsymbol{y} + \mathcal{O}_{s,\boldsymbol{y}}(\|\boldsymbol{y}\|^2)\right) \\[4mm] \dfrac{dy_i}{ds} = \left(\dfrac{\boldsymbol{\psi}_i^{(1)}(s)}{\psi_n^{(0)}(s)} - \boldsymbol{k}_i(s)\right) \cdot \boldsymbol{y} + \mathcal{O}_{s,\boldsymbol{y}}(\|\boldsymbol{y}\|^2) \end{cases}$$

Define

$$b_0(s) := \frac{1}{\psi_n^{(0)}(s)}$$

$$\boldsymbol{b}_1(s) := \frac{1}{\psi_n^{(0)}(s)}\left(\boldsymbol{k}_n(s) - \frac{\boldsymbol{\psi}_n^{(1)}(s)}{\psi_n^{(0)}(s)}\right)$$

and let $\tilde{\mathsf{A}}(s)$ denote the $(n-1) \times (n-1)$ matrix with i^{th} row given by

$$\left(\frac{\boldsymbol{\psi}_i^{(1)}(s)}{\psi_n^{(0)}(s)} - \boldsymbol{k}_i(s) \right)^{\mathsf{T}}.$$

In terms of b_0, $\boldsymbol{b}_1$, and $\tilde{\mathsf{A}}$, system (3.5) becomes

(3.6)
$$\begin{cases} \dfrac{dt}{ds} = b_0(s) + \langle \boldsymbol{b}_1(s), \boldsymbol{y} \rangle + \mathcal{O}_{s,y}(\|\boldsymbol{y}\|^2) \\ \dfrac{d\boldsymbol{y}}{ds} = \tilde{\mathsf{A}}(s)\boldsymbol{y} + \mathcal{O}_{s,y}(\|\boldsymbol{y}\|^2) \end{cases}$$

Applying the Floquet theorem, there exists a real-valued, periodic $(n-1) \times (n-1)$ matrix $\mathsf{P}(s)$ of period $2L$ such that setting $\boldsymbol{z} = \mathsf{P}^{-1}(s)\boldsymbol{y}$, we transform (3.6) into

(3.7a)
$$\frac{dt}{ds} = b_0(s) + ((\boldsymbol{b}_1(s))^{\mathsf{T}}\mathsf{P}(s))\boldsymbol{z} + h_2(s, \boldsymbol{z})$$

(3.7b)
$$\frac{d\boldsymbol{z}}{ds} = \mathsf{A}\boldsymbol{z} + \boldsymbol{h}_1(s, \boldsymbol{z})$$

This is the normal form of (2.2) on which we will base our analysis of the flow during the relaxation period (when $P_{\rho,T}(t) = 0$). We obtain the normal form of (2.2) during the forcing period (when $P_{\rho,T}(t) = 1$) by writing (3.4) in $(s, \boldsymbol{z})$-coordinates, giving

(3.8a)
$$\frac{dt}{ds} = b_0(s) + ((\boldsymbol{b}_1(s))^{\mathsf{T}}\mathsf{P}(s))\boldsymbol{z} + \mathcal{O}_{s,z}(\varepsilon) + \mathcal{O}_{s,z}(\varepsilon\boldsymbol{z}) + \mathcal{O}_{s,z}(\|\boldsymbol{z}\|^2)$$

(3.8b)
$$\frac{d\boldsymbol{z}}{ds} = \mathsf{A}\boldsymbol{z} + \frac{\varepsilon\mathsf{P}^{-1}(s)\boldsymbol{\phi}(s, \boldsymbol{0})}{\psi_n^{(0)}(s)} + \mathcal{O}_{s,z}(\varepsilon\boldsymbol{z}) + \mathcal{O}_{s,z}(\varepsilon^2) + \mathcal{O}_{s,z}(\|\boldsymbol{z}\|^2)$$

where $\boldsymbol{\phi}(s, \boldsymbol{0}) = (\phi_1(s, \boldsymbol{0}), \ldots, \phi_{n-1}(s, \boldsymbol{0}))^{\mathsf{T}}$.

3.2. **A general form of the singular limit.** Let $\tilde{M} \approx \Gamma \times D$ be a tubular neighborhood of Γ in $\mathbb{R}^n$, where D is a disk of sufficiently small radius so that the normal form (3.8a)–(3.8b) is valid. Let $M \approx \Gamma \times \frac{1}{2}D$. We define flow-induced maps $H_k : M \to \tilde{M}$ and $H_r : \tilde{M} \to \tilde{M}$ as follows. Let H_k be the time-ρ map associated with the forced system (3.8a)–(3.8b). We call H_k the 'kick'. Notice that for ε sufficiently small, H_k maps M into $\tilde{M}$. Let H_r be the time-$(T - \rho)$ map associated with the relaxation system (3.7a)–(3.7b). We call H_r the relaxation map. There exists $T_0 = T_0(\varepsilon)$ such that if $T \geqslant T_0$, then H_r maps $\tilde{M}$ into $\text{int}(M)$. The composition $G_T := H_r \circ H_k$ is the time-T map generated by the flow. Our goal is to show that the family $\{G_T : M \to \text{int}(M), \; T \geqslant T_0\}$ of diffeomorphisms on M has a well-defined singular limit in a certain sense as $T \to \infty$.

Let $(s_0, \boldsymbol{y}_0) \in M$. We write $H_k(s_0, \boldsymbol{y}_0) = (\hat{s}, \hat{\boldsymbol{z}})$ and compute $H_r(\hat{s}, \hat{\boldsymbol{z}})$. Integrating (3.7b), we have

$$\boldsymbol{z}(s) = e^{(s-\hat{s})\mathsf{A}} \left(\hat{\boldsymbol{z}} + \int_{\hat{s}}^{s} e^{-(\tau-\hat{s})\mathsf{A}} \boldsymbol{h}_1(\tau, \boldsymbol{z}(\tau)) \, d\tau \right).$$

Integrating (3.7a), we have

(3.9)
$$T - \rho = \int_{\hat{s}}^{s(T)} b_0(\tau) \, d\tau + \hat{\boldsymbol{z}} \cdot \int_{\hat{s}}^{s(T)} \boldsymbol{b}_1(\tau)^{\mathsf{T}} \mathsf{P}(\tau) e^{(\tau-\hat{s})\mathsf{A}} \, d\tau + \sum_{k=1}^{2} E_k(s(T)),$$

where the error terms are given by

$$E_1(s(T)) = \int_{\hat{s}}^{s(T)} \boldsymbol{b}_1(\tau)^{\mathsf{T}} \mathsf{P}(\tau) e^{(\tau-\hat{s})\mathsf{A}} \int_{\hat{s}}^{\tau} e^{-(\xi-\hat{s})\mathsf{A}} \boldsymbol{h}_1(\xi, \boldsymbol{z}(\xi)) \, d\xi \, d\tau$$

$$E_2(s(T)) = \int_{\hat{s}}^{s(T)} h_2(\tau, \boldsymbol{z}(\tau)) \, d\tau.$$

Letting $T \to \infty$ in (3.9) yields nothing meaningful. However, we use the fact that s is can be computed modulo $2L$ to introduce an auxiliary parameter $a \in \mathbb{S}$ and thereby obtain the singular limit. Recall that p_0 is the period of $\boldsymbol{\eta}$. As a varies from 0 to $2L$, $\boldsymbol{\gamma}$ traverses Γ 2 times. Let $\hat{t} : [0, 2L] \to [0, 2p_0)$ be the strictly

increasing function defined by $\boldsymbol{\eta}(\hat{t}(a)) = \boldsymbol{\gamma}(a)$. For $m \in \mathbb{Z}^+$ and $a \in \mathbb{S}$, set $T = \rho + 2p_0 m + \hat{t}(a)$. Substituting into (3.9), writing $s(\rho + 2p_0 m + \hat{t}(a)) = \hat{s} + 2Lm + \tilde{s}(\rho + 2p_0 m + \hat{t}(a))$, and using the fact that

$$\int_v^{v+2Lm} b_0(\tau)\,d\tau = 2p_0 m$$

for all $v \in \mathbb{R}$, we obtain

$$\hat{t}(a) = \int_{\hat{s}}^{\hat{s}+\tilde{s}(\rho+2p_0 m+\hat{t}(a))} b_0(\tau)\,d\tau + \hat{\boldsymbol{z}} \cdot \int_{\hat{s}}^{s(\rho+2p_0 m+\hat{t}(a))} \boldsymbol{b}_1(\tau)^\mathsf{T} \mathsf{P}(\tau) e^{(\tau-\hat{s})\mathsf{A}}\,d\tau$$

(3.10)

$$+ \sum_{k=1}^2 E_k(s(\rho + 2p_0 m + \hat{t}(a))).$$

Define $G_{a,m^{-1}} : M \to \mathrm{int}(M)$ by $G_{a,m^{-1}}(s_0, \boldsymbol{y}_0) = (s(\rho+2p_0 m+\hat{t}(a)), \boldsymbol{y}(\rho+2p_0 m+\hat{t}(a)))$. It follows from [17, Proposition 3.1] that there exists $s_\infty(s_0, \boldsymbol{y}_0, a)$ such that

$$\lim_{m\to\infty} \hat{s} + \tilde{s}(\rho + 2p_0 m + \hat{t}(a)) = s_\infty(s_0, \boldsymbol{y}_0, a)$$

and $s_\infty(s_0, \boldsymbol{y}_0, a)$ is defined implicitly by taking the $m \to \infty$ limit in (3.10):

(3.11) $$\hat{t}(a) = \int_{\hat{s}}^{s_\infty(s_0,\boldsymbol{y}_0,a)} b_0(\tau)\,d\tau + \left\langle \hat{\boldsymbol{z}}, \int_{\hat{s}}^{\infty} \boldsymbol{b}_1(\tau)^\mathsf{T} \mathsf{P}(\tau) e^{(\tau-\hat{s})\mathsf{A}}\,d\tau \right\rangle + \sum_{k=1}^2 E_k(\infty).$$

The family of maps $\{G_{a,0} : M \to \Gamma \times \{\boldsymbol{0}\}\}_{a \in \mathbb{S}}$ defined by

$$G_{a,0}(s_0, \boldsymbol{y}_0) = (s_\infty(s_0, \boldsymbol{y}_0, a), \boldsymbol{0})$$

is the desired singular limit. It follows from [17, Proposition 3.1] that the maps

$$(s_0, \boldsymbol{y}_0, a) \mapsto G_{a,m^{-1}}(s_0, \boldsymbol{y}_0)$$

converge to the map

$$(s_0, \boldsymbol{y}_0, a) \mapsto G_{a,0}(s_0, \boldsymbol{y}_0)$$

in $C^3(M \times \mathbb{S})$ as $m \to \infty$.

3.3. **A computable form of the singular limit.** From this point forward, we assume the setting of Theorem 1. We now extract the primary terms in the right side of (3.11). Recall that the shear integral $\boldsymbol{\Sigma}$ is defined by

$$\boldsymbol{\Sigma} = (\Sigma_1, \ldots, \Sigma_{n-1}) = \int_0^{2L} \boldsymbol{b}_1(\tau)^\mathsf{T} \mathsf{P}(\tau)\,d\tau.$$

and that the shear factor is given by $\sigma = \|\boldsymbol{\Sigma}\|$. We assume that the operator A is diagonalizable and that the $\boldsymbol{z}$-coordinate has been chosen such that $\mathsf{A} = \mathrm{diag}(\lambda_1, \ldots, \lambda_{n-1})$, where $0 > \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_{n-1}$ are the eigenvalues of A. Fix the normalized shear vector $\frac{\boldsymbol{\Sigma}}{\sigma}$ and the eigenvalue ratios $\mu_i = \frac{\lambda_1}{\lambda_i}$ for $1 \leqslant i \leqslant n-1$. Set $\rho = 1$ for notational simplicity. We regard σ, ε, and λ_1 as the parameters associated with the singular limit.

Expanding the second term on the right side of (3.11), we have

$$\int_{\hat{s}}^{\infty} \boldsymbol{b}_1(\tau)^\mathsf{T} \mathsf{P}(\tau) e^{(\tau-\hat{s})\mathsf{A}}\,d\tau = \int_{\hat{s}}^{\infty} \boldsymbol{\Sigma} e^{(\tau-\hat{s})\mathsf{A}}\,d\tau + \int_{\hat{s}}^{\infty} (\boldsymbol{b}_1(\tau)^\mathsf{T} \mathsf{P}(\tau) - \boldsymbol{\Sigma}) e^{(\tau-\hat{s})\mathsf{A}}\,d\tau$$

(3.12)

$$= \bar{\boldsymbol{d}} + \int_{\hat{s}}^{\infty} (\boldsymbol{b}_1(\tau)^\mathsf{T} \mathsf{P}(\tau) - \boldsymbol{\Sigma}) e^{(\tau-\hat{s})\mathsf{A}}\,d\tau,$$

where

$$\bar{\boldsymbol{d}} = \int_{\hat{s}}^{\infty} \boldsymbol{\Sigma} e^{(\tau-\hat{s})\mathsf{A}}\,d\tau = \left(-\frac{\Sigma_i}{\lambda_i}\right)_{i=1}^{n-1}.$$

8

Let $\tilde{H}_k : M \to \tilde{M}$ be the time-1 map generated by the system

(3.13a)
$$\frac{dt}{ds} = b_0(s)$$

(3.13b)
$$\frac{d\boldsymbol{z}}{ds} = \frac{\varepsilon \mathsf{P}^{-1}(s)\boldsymbol{\phi}(s, \boldsymbol{0})}{\psi_n^{(0)}(s)}$$

obtained from (3.8a)–(3.8b) by retaining only the terms of leading order. For $(s_0, \boldsymbol{y}_0) \in M$, write $\tilde{H}_k(s_0, \boldsymbol{y}_0) = (\tilde{s}, \tilde{\boldsymbol{z}})$. Integrating (3.13a) and (3.13b) gives

(3.14)
$$1 = \int_{s_0}^{\tilde{s}} b_0(\tau)\, d\tau,$$
$$\tilde{\boldsymbol{z}} = \boldsymbol{z}_0 + \varepsilon \int_{s_0}^{\tilde{s}} \frac{\mathsf{P}^{-1}(\tau)\boldsymbol{\phi}(\tau, \boldsymbol{0})}{\psi_n^{(0)}(\tau)}\, d\tau.$$

Proposition 3.1. *There exists a system constant $K_0 > 0$ such that*

(3.15)
$$\hat{s} = \tilde{s} + \xi_1(s_0, \boldsymbol{y}_0), \quad \hat{\boldsymbol{z}} = \tilde{\boldsymbol{z}} + \boldsymbol{\xi}_2(s_0, \boldsymbol{y}_0)$$

where
$$\|\xi_1|\{\boldsymbol{y}_0 = \boldsymbol{0}\}\|_{C^3(\mathbb{S})} \leqslant K_0 \varepsilon, \quad \|\boldsymbol{\xi}_2|\{\boldsymbol{y}_0 = \boldsymbol{0}\}\|_{C^3(\mathbb{S})} \leqslant K_0 \varepsilon |\lambda_1|.$$

Setting $\boldsymbol{y}_0 = \boldsymbol{0}$, define $g(s_0, a) = s_\infty(s_0, \boldsymbol{0}, a)$. Substituting (3.12), (3.14), and (3.15) into (3.11), the value $g(s_0, a)$ is defined implicitly by

(3.16)
$$\hat{t}(a) + 1 = \int_{s_0}^{g(s_0, a)} b_0(\tau)\, d\tau + \langle (\tilde{\boldsymbol{z}} + \boldsymbol{\xi}_2(s_0, \boldsymbol{0})), \bar{\boldsymbol{d}} \rangle$$
$$- \int_{\tilde{s}}^{\hat{s}} b_0(\tau)\, d\tau + \hat{\boldsymbol{z}} \cdot \int_{\hat{s}}^{\infty} (\boldsymbol{b}_1(\tau)^{\mathsf{T}} \mathsf{P}(\tau) - \boldsymbol{\Sigma}) e^{(\tau - \hat{s})\mathsf{A}}\, d\tau + \sum_{k=1}^{2} E_k(\infty).$$

Rescaling $\bar{\boldsymbol{d}}$, we define
$$\boldsymbol{d} = \left(\frac{\Sigma_i \mu_i}{\sigma} \right)_{i=1}^{n-1}, \qquad \Phi(s_0) = \left\langle \boldsymbol{d}, \int_{s_0}^{\tilde{s}} \frac{\mathsf{P}^{-1}(\tau)\boldsymbol{\phi}(\tau, \boldsymbol{0})}{\psi_n^{(0)}(\tau)}\, d\tau \right\rangle,$$

giving
$$\langle \tilde{\boldsymbol{z}}, \bar{\boldsymbol{d}} \rangle = \frac{\varepsilon \sigma}{|\lambda_1|} \Phi(s_0).$$

The higher-order terms are given by $\mathcal{E}_1 = E_1(\infty)$, $\mathcal{E}_2 = E_2(\infty)$,
$$\mathcal{E}_3 = \left\langle \hat{\boldsymbol{z}}, \int_{\hat{s}}^{\infty} (\boldsymbol{b}_1(\tau)^{\mathsf{T}} \mathsf{P}(\tau) - \boldsymbol{\Sigma}) e^{(\tau - \hat{s})\mathsf{A}}\, d\tau \right\rangle, \qquad \mathcal{E}_4 = -\int_{\tilde{s}}^{\hat{s}} b_0(\tau)\, d\tau, \qquad \mathcal{E}_5 = \langle \boldsymbol{\xi}_2(s_0, \boldsymbol{0}), \bar{\boldsymbol{d}} \rangle.$$

Setting $\mathcal{E} = \sum_{k=1}^{5} \mathcal{E}_k$ and substituting into (3.16), we obtain the final form of the singular limit:

(3.17)
$$\hat{t}(a) + 1 = \int_{s_0}^{g(s_0, a)} b_0(\tau)\, d\tau + \frac{\varepsilon \sigma}{|\lambda_1|} \Phi(s_0) + \mathcal{E}.$$

Proposition 3.2. *There exists a system constant $K_1 > 0$ such that the following hold.*
$$\|\mathcal{E}_1\|_{C^3(\mathbb{S})} \leqslant K_1 \frac{\sigma \varepsilon}{|\lambda_1|} \left(\frac{\varepsilon}{|\lambda_1|} \right)$$
$$\|\mathcal{E}_2\|_{C^3(\mathbb{S})} \leqslant K_1 \frac{\sigma \varepsilon}{|\lambda_1|} \left(\frac{\varepsilon}{\sigma} \right)$$
$$\|\mathcal{E}_3\|_{C^3(\mathbb{S})} \leqslant K_1 \frac{\sigma \varepsilon}{|\lambda_1|} (|\lambda_1|)$$
$$\|\mathcal{E}_4\|_{C^3(\mathbb{S})} \leqslant K_1 \frac{\sigma \varepsilon}{|\lambda_1|} \left(\frac{|\lambda_1|}{\sigma} \right)$$
$$\|\mathcal{E}_5\|_{C^3(\mathbb{S})} \leqslant K_1 \frac{\sigma \varepsilon}{|\lambda_1|} (|\lambda_1|)$$

9

4. Theory of rank one attractors

Let D denote the closed unit disk in $\mathbb{R}^{n-1}$ and let $M = \mathbb{S}^1 \times D$. We consider a family of maps $G_{\boldsymbol{a},b} : M \to M$, where $\boldsymbol{a} = (a_1, \ldots, a_k) \in \mathcal{V}$ is a vector of parameters and $b \in B_0$ is a scalar parameter. Here $\mathcal{V} = \mathcal{V}_1 \times \cdots \times \mathcal{V}_k \subset \mathbb{R}^k$ is a product of intervals and $B_0 \subset \mathbb{R} \setminus \{0\}$ is a subset of $\mathbb{R}$ with an accumulation point at 0. Points in M are denoted by (x, y) with $x \in \mathbb{S}^1$ and $y \in D$. Rank one theory postulates the following.

(H1) Regularity conditions.

    (a) For each $b \in B_0$, the function $(x, y, \boldsymbol{a}) \mapsto G_{\boldsymbol{a},b}(x, y)$ is C^3.

    (b) Each map $G_{\boldsymbol{a},b}$ is an embedding of M into itself.

    (c) There exists $K_D > 0$ independent of $\boldsymbol{a}$ and b such that for all $\boldsymbol{a} \in \mathcal{V}$, $b \in B_0$, and $z, z' \in M$, we have
$$\frac{|\det DG_{\boldsymbol{a},b}(z)|}{|\det DG_{\boldsymbol{a},b}(z')|} \leqslant K_D.$$

(H2) Existence of a singular limit. For $\boldsymbol{a} \in \mathcal{V}$, there exists a map $G_{\boldsymbol{a},0} : M \to \mathbb{S}^1 \times \{0\}$ such that the following holds. For every $(x, y) \in M$ and $\boldsymbol{a} \in \mathcal{V}$, we have
$$\lim_{b \to 0} G_{\boldsymbol{a},b}(x, y) = G_{\boldsymbol{a},0}(x, y)$$

Identifying $\mathbb{S}^1 \times \{0\}$ with $\mathbb{S}^1$, we refer to $G_{\boldsymbol{a},0}$ and the restriction $f_{\boldsymbol{a}} : \mathbb{S}^1 \to \mathbb{S}^1$ defined by $f_{\boldsymbol{a}}(x) = G_{\boldsymbol{a},0}(x, 0)$ as the *singular limit* of $G_{\boldsymbol{a},b}$.

(H3) C^3 convergence to the singular limit. We select a special index $j \in \{1, \ldots, k\}$. Fix $a_i \in \mathcal{V}_i$ for $i \neq j$. For every such choice of parameters a_i, the maps $(x, y, a_j) \mapsto G_{\boldsymbol{a},b}(x, y)$ converge in the C^3 topology to $(x, y, a_j) \mapsto G_{\boldsymbol{a},0}(x, y)$ on $M \times \mathcal{V}_j$ as $b \to 0$.

(H4) Existence of a sufficiently expanding map within the singular limit. There exists $\boldsymbol{a}^* = (a_1^*, \ldots, a_k^*) \in \mathcal{V}$ such that $f_{\boldsymbol{a}^*} \in \mathcal{M}$, where $\mathcal{M}$ is the set of Misiurewicz-type maps defined in Definition 4.1 below.

(H5) Parameter transversality. Let $C_{\boldsymbol{a}^*}$ denote the critical set of $f_{\boldsymbol{a}^*}$. For $a_j \in \mathcal{V}_j$, define the vector $\tilde{\boldsymbol{a}}_j \in \mathcal{V}$ by $\tilde{\boldsymbol{a}}_j = (a_1^*, \ldots, a_{j-1}^*, a_j, a_{j+1}^*, \ldots, a_k^*)$. We say that the family $\{f_{\boldsymbol{a}}\}$ satisfies the *parameter transversality* condition with respect to parameter a_j if the following holds. For each $x \in C_{\boldsymbol{a}^*}$, let $p = f_{\boldsymbol{a}^*}(x)$ and let $x(\tilde{\boldsymbol{a}}_j)$ and $p(\tilde{\boldsymbol{a}}_j)$ denote the continuations of x and p, respectively, as the parameter a_j varies around a_j^*. The point $p(\tilde{\boldsymbol{a}}_j)$ is the unique point such that $p(\tilde{\boldsymbol{a}}_j)$ and p have identical symbolic itineraries under $f_{\tilde{\boldsymbol{a}}_j}$ and $f_{\boldsymbol{a}^*}$, respectively. We have
$$\left. \frac{d}{da_j} f_{\tilde{\boldsymbol{a}}_j}(x(\tilde{\boldsymbol{a}}_j)) \right|_{a_j = a_j^*} \neq \left. \frac{d}{da_j} p(\tilde{\boldsymbol{a}}_j) \right|_{a_j = a_j^*}.$$

(H6) Nondegeneracy at 'turns'. For each $x \in C_{\boldsymbol{a}^*}$, there exists $1 \leqslant m \leqslant n - 1$ such that
$$\left. \frac{\partial}{\partial y_m} G_{\boldsymbol{a}^*,0}(x, y) \right|_{y=0} \neq 0.$$

(H7) Conditions for mixing.

    (a) We have $e^{\frac{1}{3}\lambda_0} > 2$, where λ_0 is defined within Definition 4.1.

    (b) Let $J_1, \ldots, J_r$ be the intervals of monotonicity of $f_{\boldsymbol{a}^*}$. Let $Q = (q_{im})$ be the matrix of 'allowed transitions' defined by
$$q_{im} = \begin{cases} 1, & \text{if } f_{\boldsymbol{a}^*}(J_i) \supset J_m, \\ 0, & \text{otherwise.} \end{cases}$$
    There exists $N > 0$ such that $Q^N > 0$.

We now define the family $\mathcal{M}$.

Definition 4.1. We say that $f \in C^2(\mathbb{S}^1, \mathbb{R})$ is a Misiurewicz map and we write $f \in \mathcal{M}$ if the following hold for some neighborhood U of the critical set $C = C(f) = \{x \in \mathbb{S}^1 : f'(x) = 0\}$.

(A) **(Outside of U)** There exist $\lambda_0 > 0$, $M_0 \in \mathbb{Z}^+$, and $0 < d_0 \leqslant 1$ such that
    (1) for all $m \geqslant M_0$, if $f^i(x) \notin U$ for $0 \leqslant i \leqslant m-1$, then $|(f^m)'(x)| \geqslant e^{\lambda_0 m}$,
    (2) for any $m \in \mathbb{Z}^+$, if $f^i(x) \notin U$ for $0 \leqslant i \leqslant m-1$ and $f^m(x) \in U$, then $|(f^m)'(x)| \geqslant d_0 e^{\lambda_0 m}$.
(B) **(Critical orbits)** For all $c \in C$ and $i > 0$, $f^i(c) \notin U$.
(C) **(Inside U)**
    (1) We have $f''(x) \neq 0$ for all $x \in U$, and
    (2) for all $x \in U \setminus C$, there exists $p_0(x) > 0$ such that $f^i(x) \notin U$ for all $i < p_0(x)$ and $|(f^{p_0(x)})'(x)| \geqslant d_0^{-1} e^{\frac{1}{3}\lambda_0 p_0(x)}$.

Rank one theory states that given a family $\{G_{\boldsymbol{a},b}\}$ satisfying **(H1)**–**(H6)**, a measure-theoretically significant subset of this family consists of maps admitting attractors with strong chaotic and stochastic properties. We formulate the precise results and we then describe the properties that the attractors possess.

Theorem 4.2 ([15, 18])**.** *Suppose the family $\{G_{\boldsymbol{a},b}\}$ satisfies **(H1)**, **(H2)**, **(H4)**, and **(H6)**. The following holds for all $1 \leqslant j \leqslant k$ such that the parameter a_j satisfies **(H3)** and **(H5)**. For all sufficiently small $b \in B_0$, there exists a subset $\Delta_j \subset \mathcal{V}_j$ of positive Lebesgue measure such that for $a_j \in \Delta_j$, $G_{\tilde{\boldsymbol{a}}_j,b}$ admits a strange attractor Ω with properties **(SA1)**, **(SA2)**, and **(SA3)**.*

Theorem 4.3 ([15, 16, 18])**.** *In the sense of Theorem 4.2,*

$$\textbf{(H1)–(H7)} \Longrightarrow \textbf{(SA1)–(SA4)}.$$

Remark 4.4. The proof of Theorem 4.2 for the special case $n = 2$ appears in [15]. The additional component **(H7)** $\Rightarrow$ **(SA4)** in Theorem 4.3 is proved in [16]. For general n, Wang and Young [18] prove the existence of an SRB measure for $G_{\tilde{\boldsymbol{a}}_j,b}$ if $a_j \in \Delta_j$. The complete proofs of **(SA1)**–**(SA3)** (and **(SA4)** assuming **(H7)**) for $G_{\tilde{\boldsymbol{a}}_j,b}$ with $a_j \in \Delta_j$ will appear in [14] for general n.

We now describe **(SA1)**–**(SA4)** precisely. Write $G = G_{\tilde{\boldsymbol{a}}_j,b}$.

(SA1) Positive Lyapunov exponent. Let U denote the basin of attraction of the attractor Ω. This means that U is an open set satisfying $G(\overline{U}) \subset U$ and

$$\Omega = \bigcap_{m=0}^{\infty} T^m(\overline{U}).$$

For almost every $z \in U$ with respect to Lebesgue measure, the orbit of z has a positive Lyapunov exponent. That is,

$$\lim_{m \to \infty} \frac{1}{m} \log \|DG^m(z)\| > 0.$$

(SA2) Existence of SRB measures and basin property.
    (a) The map G admits at least one and at most finitely many ergodic SRB measures each one of which has no zero Lyapunov exponents. Let $\nu_1, \cdots, \nu_r$ denote these measures.
    (b) For Lebesgue-a.e. $z \in U$, there exists $j(z) \in \{1, \ldots, r\}$ such that for every continuous function $\varphi : U \to \mathbb{R}$,

$$\frac{1}{m} \sum_{i=0}^{m-1} \varphi(G^i(x,y)) \to \int \varphi \, d\nu_{j(z)}.$$

(SA3) Statistical properties of dynamical observations.
    (a) For every ergodic SRB measure ν and every Hölder continuous function $\varphi : \Omega \to \mathbb{R}$, the sequence $\{\varphi \circ G^i : i \in \mathbb{Z}^+\}$ obeys a central limit theorem. That is, if $\int \varphi \, d\nu = 0$, then the sequence

$$\frac{1}{\sqrt{m}} \sum_{i=0}^{m-1} \varphi \circ G^i$$

converges in distribution (with respect to ν) to the normal distribution. The variance of the limiting normal distribution is strictly positive unless $\varphi = \psi \circ G - \psi$ for some $\psi \in L^2(\nu)$.

(b) Suppose that for some $N \geqslant 1$, G^N has an SRB measure ν that is mixing. Then given a Hölder exponent η, there exists $\tau = \tau(\eta) < 1$ such that for all Hölder $\varphi, \psi : \Omega \to \mathbb{R}$ with Hölder exponent η, there exists $K = K(\varphi, \psi)$ such that for all $m \in \mathbb{N}$,

$$\left| \int (\varphi \circ G^{mN}) \psi \, d\nu - \int \varphi \, d\nu \int \psi \, d\nu \right| \leqslant K(\varphi, \psi) \tau^m.$$

(SA4) Uniqueness of SRB measures and ergodic properties.

    (a) The map G admits a unique (and therefore ergodic) SRB measure ν, and

    (b) the dynamical system (G, ν) is mixing, or, equivalently, isomorphic to a Bernoulli shift.

5. VERIFICATION OF THE RANK ONE HYPOTHESES

We view the singular limit $\{G_{a,0} : a \in \mathbb{S}\}$ as a function of 3 parameters: ε, σ, and λ_1. We show that the family $\{G_{a,m^{-1}} : a \in \mathbb{S}, \ m \in \mathbb{Z}^+\}$ satisfies **(H1)**–**(H7)** if the parameters ε, σ, and λ_1 satisfy certain scaling assumptions.

5.1. 1D analysis: verification of (H4), (H5), and (H7). Recall that $g(s, a)$ is defined implicitly by

$$\hat{t}(a) + 1 = \int_s^{g(s,a)} b_0(\tau) \, d\tau + \frac{\varepsilon \sigma}{|\lambda_1|} \Phi(s) + \mathcal{E}.$$

Defining $f_a(s) = g(s, a)$, $\Lambda = \frac{\varepsilon \sigma}{|\lambda_1|}$, and $\Psi(s) = \Phi(s) + \Lambda^{-1} \mathcal{E}$, the singular limit becomes

$$(5.1) \qquad \hat{t}(a) + 1 = \int_s^{f_a(s)} b_0(\tau) \, d\tau + \Lambda \Psi(s).$$

For a map $f : \mathbb{S} \to \mathbb{S}$ and $\delta > 0$, let $C(f) = \{s : f'(s) = 0\}$ and let $C_\delta(f) = \{s : |s - \hat{s}| < \delta \text{ for some } \hat{s} \in C(f)\}$. We assume the following about Ψ: there exist positive constants K_2, d_0, d_1, and d_2, and a constant δ_0 satisfying $0 < \delta_0 < \frac{1}{2} d_1$, such that the following hold.

(A1) $\|\Psi\|_{C^3(\mathbb{S})} < K_2$

(A2) $|\Psi''(s)| > d_0$ for $s \in C_{\delta_0}(\Psi)$

(A3) If $\Psi'(s_1) = \Psi'(s_2) = 0$ and $s_1 \neq s_2$, then $|s_1 - s_2| > d_1$.

(A4) $|\Psi'(s)| > d_2$ for $s \in \mathbb{S} \setminus C_{\delta_0}(\Psi)$

Because Φ is a Morse function, Proposition 3.2 implies that assumptions (A1)–(A4) are satisfied if $\sigma \geqslant 1$, $|\lambda_1|$ is sufficiently small, and $\frac{\varepsilon}{|\lambda_1|}$ is sufficiently small.

We now compare the map f_a to the map Ψ. Let $\{\bar{v}_1, \ldots, \bar{v}_{q_0}\}$ be the set of critical points of Ψ. Set $\xi = \Lambda^{-\frac{3}{4}}$.

Lemma 5.1. *There exists $\Lambda_0 > 0$ and positive constants K_3, K_4, and K_5 such that the following hold for fixed $\Lambda > \Lambda_0$.*

(a) $C(f_a) = \{v_1, \ldots, v_{q_0}\}$ with $|v_i - \bar{v}_i| < K_3 \Lambda^{-1}$ for $1 \leqslant i \leqslant q_0$

(b) $|f_a''(s)| > K_4 \Lambda$ for all $s \in C_\xi(f_a)$

(c) $|f_a'(s)| > K_5 \Lambda^{\frac{1}{4}}$ for all $s \in \mathbb{S} \setminus C_{\frac{1}{2}\xi}(f_a)$

Proof of Lemma 5.1. Differentiating (5.1) with respect to s, we obtain

$$(5.2) \qquad b_0(s) - \Lambda \Psi'(s) = b_0(f_a(s)) f_a'(s).$$

Setting $f_a'(s) = 0$ gives $b_0(s) = \Lambda \Psi'(s)$. Since b_0 is bounded above and bounded away from 0, (A2)–(A4) imply (a). Solving for $f_a'(s)$, we have

$$(5.3) \qquad f_a'(s) = \frac{b_0(s) - \Lambda \Psi'(s)}{b_0(f_a(s))}.$$

On $\mathbb{S} \setminus C_{\frac{1}{2}\xi}(f_a)$ we have $|\Psi'(s)| > K\xi$ using (a), (A2), and (A4). Estimate (c) now follows from (5.3). Differentiating (5.2) with respect to s, we obtain

$$(5.4) \qquad b_0'(s) - \Lambda \Psi''(s) - b_0'(f_a(s))[f_a'(s)]^2 = b_0(f_a(s)) f_a''(s).$$

12

For all $s \in C_\xi(f_a)$, we have $|\Psi'(s)| < K\xi$ by (A1) and (a). This implies that $|f_a'(s)| < K\Lambda^{\frac{1}{4}}$ on C_ξ using (5.3). Therefore the second term on the left side of (5.4) dominates and (b) holds. ∎

5.1.1. *Critical curves.* Assume $\Lambda > \Lambda_0$ and let $\Delta \subset \mathbb{S}$ be a parameter interval. For $a \in \Delta$, we have $C(f_a) = \{v_1(a), \ldots, v_{q_0}(a)\}$ by Lemma 5.1. Write $\gamma^{(i)}(a) = v_i(a)$ for $1 \leqslant i \leqslant q_0$. For $1 \leqslant k \leqslant q_0$ and $i \in \mathbb{N}$, define $\gamma_i^{(k)}(a) := f_a^i(\gamma^{(k)}(a))$.

Differentiating $\gamma_1^{(k)}(a) = f_a(\gamma^{(k)}(a)) = f(\gamma^{(k)}(a), a)$ with respect to a, we have

$$
\begin{aligned}
\frac{d}{da}\gamma_1^{(k)}(a) &= \frac{\partial f}{\partial s}(\gamma^{(k)}(a), a) \cdot \frac{d}{da}\gamma^{(k)}(a) + \frac{\partial f}{\partial a}(\gamma^{(k)}(a), a) \\
&= \frac{\partial f}{\partial a}(\gamma^{(k)}(a), a).
\end{aligned}
$$
(5.5)

Differentiating (5.1) with respect to a and using the fact that $\frac{d}{da}\hat{t}(a) = b_0(a)$, we obtain

$$
\frac{\partial}{\partial a}f(s, a) = \frac{b_0(a)}{b_0(f(s, a))}.
$$
(5.6)

Thus

$$
\frac{d}{da}\gamma_1^{(k)}(a) \geqslant \frac{\min_{s \in \mathbb{S}} b_0(s)}{\max_{s \in \mathbb{S}} b_0(s)} > 0.
$$

More generally, an estimate on $\frac{d}{da}\gamma_{i+1}^{(k)}(a)$ for $i \in \mathbb{N}$ follows from the recursive formula

$$
\frac{d}{da}\gamma_{i+1}^{(k)}(a) = \frac{\partial f}{\partial s}(\gamma_i^{(k)}(a), a) \cdot \frac{d}{da}\gamma_i^{(k)}(a) + \frac{\partial f}{\partial a}(\gamma_i^{(k)}(a), a).
$$
(5.7)

Lemma 5.2 (Growth estimate for derivatives of critical curves). *There exists $\Lambda_1 \geqslant \Lambda_0$ such that the following holds for all $\Lambda > \Lambda_1$. For any $k \in \{1, \ldots, q_0\}$ and $i \in \mathbb{N}$ such that $\gamma_j^{(k)}(a) \in \mathbb{S} \setminus C_\xi(\Psi)$ for all $1 \leqslant j \leqslant i$, then*

$$
\left| \frac{d}{da}\gamma_{i+1}^{(k)}(a) \right| > \left(\frac{K_5}{2}\Lambda^{\frac{1}{4}} \right)^i > \Lambda^{\frac{i}{5}}.
$$
(5.8)

Proof of Lemma 5.2. Estimate (5.8) follows from (5.7), estimate (c) from Lemma 5.1, and the fact that for all $s \in \mathbb{S}$ and $a \in \Delta$ we have

$$
\frac{\partial}{\partial a}f(s, a) \leqslant \frac{\max_{s \in \mathbb{S}} b_0(s)}{\min_{s \in \mathbb{S}} b_0(s)} =: K_6.
$$

∎

Lemma 5.3 (Distortion estimate for critical curves). *There exists $\Lambda_2 \geqslant \Lambda_1$ and $D_1 > 0$ such that the following holds for all $\Lambda > \Lambda_2$. For any $k \in \{1, \ldots, q_0\}$ and any $n \geqslant 2$, let Δ be a parameter interval such that*

(a) $\gamma_i^{(k)}(\Delta) \subset \mathbb{S} \setminus C_\xi(\Psi)$ for $1 \leqslant i \leqslant n - 1$, and
(b) $\ell(\gamma_{n-1}^{(k)}(\Delta)) < \xi$ (ℓ denotes Lebesgue measure on $\mathbb{S}$).
Then for all $a, \hat{a} \in \Delta$, we have

$$
\left| \frac{\frac{d}{da}\gamma_n^{(k)}(a)}{\frac{d}{da}\gamma_n^{(k)}(\hat{a})} \right| < D_1.
$$
(5.9)

If $n = 1$, then (5.9) holds for all $k \in \{1, \ldots, q_0\}$ and for all $a, \hat{a} \in \mathbb{S}$.

Proof of Lemma 5.3. For $n = 1$ and $a, \hat{a} \in \mathbb{S}$, the estimate

$$
\left| \frac{\frac{d}{da}\gamma_1^{(k)}(a)}{\frac{d}{da}\gamma_1^{(k)}(\hat{a})} \right| < K_6^2
$$

follows from (5.5) and (5.6). For $n \geqslant 2$ and $a, \hat{a} \in \Delta$, let $s_i = \gamma_i^{(k)}(a)$ and $\hat{s}_i = \gamma_i^{(k)}(\hat{a})$. We have

$$
\left| \frac{\frac{d}{da}s_i}{\frac{d}{da}\hat{s}_i} \right| = \left| \frac{f_a'(s_{i-1}) \cdot \frac{d}{da}s_{i-1} + \frac{\partial}{\partial a}f_a(s_{i-1})}{f_{\hat{a}}'(\hat{s}_{i-1}) \cdot \frac{d}{da}\hat{s}_{i-1} + \frac{\partial}{\partial a}f_{\hat{a}}(\hat{s}_{i-1})} \right| = \left| \frac{f_a'(s_{i-1}) \cdot \frac{d}{da}s_{i-1}}{f_{\hat{a}}'(\hat{s}_{i-1}) \cdot \frac{d}{da}\hat{s}_{i-1}} \right| \left(1 + \mathcal{O}(\Lambda^{-\frac{1}{5}(i-1)}) \right).
$$

13

This implies the estimate

$$\log\left|\frac{\frac{d}{da}s_n}{\frac{d}{da}\hat{s}_n}\right| = \log\left|\frac{\frac{d}{da}s_1}{\frac{d}{da}\hat{s}_1}\right| + \sum_{i=1}^{n-1}\log\left|\frac{f_a'(s_i)}{f_{\hat{a}}'(\hat{s}_i)}\right| + \sum_{i=1}^{n-1}\log\left(1 + \mathcal{O}(\Lambda^{-\frac{i}{5}})\right)$$

$$\leqslant \sum_{i=1}^{n-1}\frac{|f_a'(s_i) - f_{\hat{a}}'(\hat{s}_i)|}{|f_{\hat{a}}'(\hat{s}_i)|} + \mathcal{O}(1).$$

The equality

$$|f_a'(s_i) - f_{\hat{a}}'(\hat{s}_i)| = \left|\frac{1}{b_0(f_a(s_i))}\left((b_0(s_{i+1}) - b_0(\hat{s}_{i+1}))f_{\hat{a}}'(\hat{s}_i) + \Lambda(\Psi'(s_i) - \Psi'(\hat{s}_i)) + (b_0(\hat{s}_i) - b_0(s_i))\right)\right|$$

implies the estimate

$$\log\left|\frac{\frac{d}{da}s_n}{\frac{d}{da}\hat{s}_n}\right| \leqslant K\sum_{i=1}^{n-1}|s_{i+1} - \hat{s}_{i+1}| + K\Lambda^{\frac{3}{4}}\sum_{i=1}^{n-1}|s_i - \hat{s}_i| + K\Lambda^{-\frac{1}{4}}\sum_{i=1}^{n-1}|\hat{s}_i - s_i| + \mathcal{O}(1)$$

$$= K\left(|s_n - \hat{s}_n| + \sum_{i=2}^{n-1}|s_i - \hat{s}_i|\right) + K\Lambda^{\frac{3}{4}}\sum_{i=1}^{n-1}|s_i - \hat{s}_i| + K\Lambda^{-\frac{1}{4}}\sum_{i=1}^{n-1}|\hat{s}_i - s_i| + \mathcal{O}(1)$$

$$\leqslant K|s_n - \hat{s}_n| + K|s_{n-1} - \hat{s}_{n-1}|\sum_{i=0}^{n-3}\Lambda^{-\frac{1}{5}i} + K|s_{n-1} - \hat{s}_{n-1}|(\Lambda^{\frac{3}{4}} + \Lambda^{-\frac{1}{4}})\sum_{i=0}^{n-2}\Lambda^{-\frac{1}{5}i} + \mathcal{O}(1)$$

$$= \mathcal{O}(1).$$

∎

5.1.2. *Verification of* **(H4)**: *Definition 4.1(B).* We prove the existence of a parameter a^* such that f_{a^*} satisfies Definition 4.1(B). We will then show that if Λ is sufficiently large, then for any parameter a, if f_a satisfies Definition 4.1(B), then $f_a \in \mathcal{M}$.

Proposition 5.4. *There exists $\Lambda_3 \geqslant \Lambda_2$ such that if $\Lambda \geqslant \Lambda_3$ and $\Delta \subset \mathbb{S}$ is a parameter interval satisfying $\ell(\Delta) = 3D_1K_6q_0\xi$, then there exists $a^* \in \Delta$ such that for all $c \in C(f_{a^*})$, $f_{a^*}^n(c) \in \mathbb{S}\setminus C_\xi(\Psi)$ for all $n \in \mathbb{N}$.*

Proof of Proposition 5.4. We inductively construct a nested sequence of parameter intervals $\Delta = \Delta_0 \supset \Delta_1 \supset \Delta_2 \supset \cdots$ such that $a^* \in \bigcap_{i=0}^\infty \Delta_i$ has the desired property.

Definition 5.5. The $(q_0 + 1)$-tuple $(\Delta_n; i_{1,n}, \ldots, i_{q_0,n})$ is called an ***admissible configuration*** if Δ_n is a subinterval of Δ_0 and if for every $k \in \{1, \ldots, q_0\}$, $i_{k,n} \leqslant n$ and the following conditions are satisfied.

(M1) $\gamma_i^{(k)}(\Delta_n) \cap C_\xi(\Psi) = \emptyset$ for all $i \leqslant i_{k,n}$

(M2) For all $a, \hat{a} \in \Delta_n$, we have the distortion estimate

$$\left|\frac{\frac{d}{da}\gamma_{i_{k,n}}^{(k)}(a)}{\frac{d}{da}\gamma_{i_{k,n}}^{(k)}(\hat{a})}\right| < D_1.$$

(M3) $\ell\left(\gamma_{i_{k,n}+1}^{(k)}(\Delta_n)\right) \geqslant 3D_1q_0\xi$

We inductively construct admissible configurations for all $n \in \mathbb{N}$ such that $i_{k,n} \to \infty$ as $n \to \infty$ for every k. We begin with $n = 1$. Let

$$\tilde{d} := \min_{\substack{s,t \in C(\Psi) \\ s \neq t}}|s - t|.$$

We assume that $3D_1K_6^2q_0\xi < \frac{1}{2}\tilde{d}$. Let $i_{k,1} = 1$ for all k. We choose Δ_1 as follows. We have

$$\frac{d}{da}\gamma_1^{(k)}(a) = \frac{b_0(a)}{b_0(\gamma_1^{(k)}(a))},$$

so

$$3D_1q_0\xi \leqslant \ell(\gamma_1^{(k)}(\Delta_0)) \leqslant 3D_1K_6^2q_0\xi < \frac{1}{2}\tilde{d}.$$

14

Consequently, $\gamma_1^{(k)}(\Delta_0)$ meets at most one component of $C_\xi(\Psi)$ and we have

$$\frac{\ell((\gamma_1^{(k)}|\Delta_0)^{-1}(C_\xi(\Psi)))}{\ell(\Delta_0)} \leqslant \frac{2}{3q_0}.$$

Even in the worst-case scenario in which the q_0 intervals $\{(\gamma_1^{(k)})^{-1}(C_\xi(\Psi)) : 1 \leqslant k \leqslant q_0\}$ are evenly spaced in Δ_0, there exists a subinterval Δ_1 of Δ_0 with $\ell(\Delta_1) \geqslant \frac{D_1 K_6 q_0 \xi}{q_0+1}$ such that $\gamma_1^{(k)}(\Delta_1) \cap C_\xi(\Psi) = \emptyset$ for all k. Property (M1) holds by design and (M2) follows from Lemma 5.3. Property (M3) holds if Λ is such that

$$\ell(\gamma_2^{(k)}(\Delta_1)) \geqslant \Lambda^{\frac{1}{5}} \ell(\Delta_1) \geqslant 3D_1 q_0 \xi.$$

Now assume that for $n \in \mathbb{N}$ we are given an admissible configuration $(\Delta_n; i_{1,n}, \ldots, i_{q_0,n})$. We construct an admissible configuration at step $n+1$ as follows. Partition the set $\{1, \ldots, q_0\}$ into 2 sets: A, the set of indices that are 'ready to advance', and $\{1, \ldots, q_0\} \setminus A$, the set of indices that are not ready to advance. The index k is in A if (I1) and (I2) hold:

(I1) $\ell(\gamma_{i_{k,n}}^{(k)}(\Delta_n)) < \xi$ (distortion estimate holds for the next iterate)

(I2) $\ell(\gamma_{i_{k,n}+1}^{(k)}(\Delta_n)) < \frac{1}{2}\tilde{d}$ (image of the next iterate meets at most one component of $C_\xi(\Psi)$)

Suppose that $A \neq \emptyset$. In this case, set

$$i_{k,n+1} = \begin{cases} i_{k,n} + 1, & \text{if } k \in A; \\ i_{k,n}, & \text{if } k \in \{1, \ldots, q_0\} \setminus A. \end{cases}$$

We now find Δ_{n+1} so that $(\Delta_{n+1}; i_{1,n+1}, \ldots, i_{q_0,n+1})$ is an admissible configuration. Let $k \in A$. Using (M3) and (I2), we have

$$3D_1 q_0 \xi \leqslant \ell(\gamma_{i_{k,n}+1}^{(k)}(\Delta_n)) < \frac{1}{2}\tilde{d}.$$

This implies that the fraction of $\gamma_{i_{k,n}+1}^{(k)}(\Delta_n)$ in $C_\xi(\Psi)$ is bounded above by $\frac{2\xi}{3D_1 q_0 \xi} = \frac{2}{3D_1 q_0}$. Using (I1) and Lemma 5.3, we have

$$\frac{\ell((\gamma_{i_{k,n}+1}^{(k)}|\Delta_n)^{-1}(C_\xi(\Psi)))}{\ell(\Delta_n)} \leqslant \frac{2}{3q_0}.$$

Arguing as in the $n=1$ case, there exists a subinterval Δ_{n+1} of Δ_n such that $\ell(\Delta_{n+1}) \geqslant \frac{1}{3(q_0+1)}\ell(\Delta_n)$ and for all $k \in A$, $\gamma_{i_{k,n}+1}^{(k)}(\Delta_{n+1}) \cap C_\xi(\Psi) = \emptyset$. For $k \in A$, (M1) holds by design and (M2) follows from (I1). The inequality

$$\ell(\gamma_{i_{k,n}+1}^{(k)}(\Delta_{n+1})) \geqslant \frac{D_1^{-1}}{3(q_0+1)}(3D_1 q_0 \xi) = \frac{q_0}{q_0+1}\xi$$

implies that (M3) holds if Λ is such that

$$\ell(\gamma_{i_{k,n}+2}^{(k)}(\Delta_{n+1})) \geqslant \left(\frac{q_0 \xi}{q_0+1}\right)\left(\frac{K_5}{2}\Lambda^{\frac{1}{4}}\right) \geqslant 3D_1 q_0 \xi.$$

Now let $k \in \{1, \ldots, q_0\} \setminus A$. Properties (M1) and (M2) are inherited from step n. If (I1) fails for index k, then (M2) gives

$$\ell(\gamma_{i_{k,n}}^{(k)}(\Delta_{n+1})) \geqslant \frac{\xi}{3D_1(q_0+1)},$$

so index k satisfies (M3) if Λ is such that

$$\ell(\gamma_{i_{k,n}+1}^{(k)}(\Delta_{n+1})) \geqslant \left(\frac{\xi}{3D_1(q_0+1)}\right)\left(\frac{K_5}{2}\Lambda^{\frac{1}{4}}\right) \geqslant 3D_1 q_0 \xi.$$

If (I1) holds but (I2) fails, then (M3) holds for index k if Λ is such that

$$\ell(\gamma_{i_{k,n}+1}^{(k)}(\Delta_{n+1})) \geqslant \left(\frac{1}{3D_1(q_0+1)}\right)\left(\frac{1}{2}\tilde{d}\right) \geqslant 3D_1 q_0 \xi.$$

If $A = \emptyset$, then let Δ_n' be the left half of Δ_n. We claim that $(\Delta_n'; i_{1,n}, \ldots, i_{q_0,n})$ is an admissible configuration. For each index k, properties (M1) and (M2) trivially hold. Property (M3) is established (for Λ sufficiently large) by arguing as above in the 2 cases

    (1) (I1) does not hold, and
    (2) (I1) holds but (I2) fails.

Repeat the halving process until $A \neq \emptyset$. ∎

5.1.3. *Verification of **(H4)**:* f_{a^*} *satisfies Definition 4.1(B)* $\Rightarrow$ $f_{a^*} \in \mathcal{M}$. We show that if Λ is sufficiently large and $a^* \in \mathbb{S}$ is as in Proposition 5.4, then $f_{a^*} \in \mathcal{M}$. This implication is a consequence of Lemma 5.1 and the following binding estimate.

Proposition 5.6. *There exists $K_7 > 0$ such that for Λ sufficiently large and a^* as in Proposition 5.4, we have the following. For $c \in C(f_{a^*})$ and $s \in \mathbb{S}$ satisfying $|s - c| \leqslant \Lambda^{-\frac{11}{12}}$, let $m(s)$ be the smallest value of $m \in \mathbb{Z}^+$ such that $|f_{a^*}^m(s) - f_{a^*}^m(c)| > \frac{1}{2}\xi$. Then $m(s) > 1$ and*

$$|(f_{a^*}^{m(s)})'(s)| \geqslant (K_7\Lambda)^{\frac{m(s)}{16}}.$$

Proof of Proposition 5.6. We begin with a spatial distortion lemma.

Lemma 5.7 (Spatial distortion estimate). *There exists $D_2 \geqslant 1$ such that the following holds for all $a \in \mathbb{S}$. For $s, \hat{s} \in \mathbb{S}$, let $m \in \mathbb{Z}^+$ be such that π_i, the segment between $f_a^i(s)$ and $f_a^i(\hat{s})$, satisfies $\ell(\pi_i) < \frac{1}{2}\xi$ and $\pi_i \cap C_{\frac{1}{2}\xi}(f_a) = \emptyset$ for all $0 \leqslant i < m$. Then*

$$\left| \frac{(f_a^m)'(s)}{(f_a^m)'(\hat{s})} \right| \leqslant D_2.$$

Proof of Lemma 5.7. Writing $s_i = f_a^i(s)$ and $\hat{s}_i = f_a^i(\hat{s})$ and using Lemma 5.1 and its proof, we have

$$
\begin{aligned}
\log \left| \frac{(f_a^m)'(s)}{(f_a^m)'(\hat{s})} \right| &= \sum_{i=0}^{m-1} \log \left| \frac{f_a'(s_i)}{f_a'(\hat{s}_i)} \right| \\
&\leqslant \sum_{i=0}^{m-1} \frac{|f_a'(s_i) - f_a'(\hat{s}_i)|}{|f_a'(\hat{s}_i)|} \\
&\leqslant K_5^{-1} \Lambda^{-\frac{1}{4}} \left(K_6 + \frac{\Lambda \|\Psi'\|_{C^0(\mathbb{S})}}{\min_{w \in \mathbb{S}} b_0(w)} \right) \sum_{i=0}^{m-1} |s_i - \hat{s}_i| \\
&\leqslant (K\Lambda^{-\frac{1}{4}} + K\Lambda^{\frac{3}{4}})|s_{m-1} - \hat{s}_{m-1}| \sum_{i=0}^{m-1} (K_5\Lambda^{\frac{1}{4}})^{-i} \\
&= \mathcal{O}(1).
\end{aligned}
$$

∎

Returning to the proof of Proposition 5.6, write $f = f_{a^*}$. We first show that $m(s) > 1$. We have

$$|f(s) - f(c)| = \frac{1}{2}|f''(\zeta)|(s - c)^2$$

for some ζ satisfying $|\zeta - c| \leqslant \Lambda^{-\frac{11}{12}}$. Arguing as in the proof of Lemma 5.1, $|f''(\zeta)| \leqslant K\Lambda$. Therefore

$$|f(s) - f(c)| \leqslant K\Lambda^{-\frac{5}{6}} \leqslant \frac{\xi}{2}$$

for Λ sufficiently large.

Now assume $m(s) > 1$. Using Lemma 5.7, we have

$$
\begin{aligned}
\frac{\xi}{2} &< |f^{m(s)}(s) - f^{m(s)}(c)| \\
&= |(f^{m(s)-1})'(\zeta_1)| \cdot |f(s) - f(c)| \qquad \text{(for some ζ_1 between $f(s)$ and $f(c)$)} \\
&\leqslant D_2|(f^{m(s)-1})'(f(c))| \cdot |f(s) - f(c)|
\end{aligned}
$$

16

and therefore

$$(5.10) \qquad \xi < D_2|(f^{m(s)-1})'(f(c))| \cdot |f''(\zeta)| \cdot (s-c)^2.$$

Reversing inequality (5.10) at time $m(s) - 1$, we have

$$(5.11) \qquad \xi \geqslant D_2^{-1}|(f^{m(s)-2})'(f(c))| \cdot |f''(\zeta)| \cdot (s-c)^2.$$

Estimating $|(f^{m(s)-1})'(f(c))|$ from below using (5.10) gives

$$\begin{aligned}
|(f^{m(s)})'(s)| &= |f'(s) - f'(c)| \cdot |(f^{m(s)-1})'(f(s))| \\
&\geqslant D_2^{-1}|(f^{m(s)-1})'(f(c))| \cdot |f''(\zeta_4)| \cdot |s-c| \qquad \text{(for some } \zeta_4 \text{ between } s \text{ and } c) \\
&\geqslant \frac{\xi}{D_2^2|s-c|}\left(\frac{|f''(\zeta_4)|}{|f''(\zeta)|}\right).
\end{aligned}$$

(5.12)

Arguing as in the proof of Lemma 5.1, $\frac{|f''(\zeta_4)|}{|f''(\zeta)|} \geqslant K > 0$ since ζ_4 and ζ are between s and c. Using this fact and estimating $|s-c|^{-1}$ from below using (5.11), (5.12) implies

$$\begin{aligned}
|(f^{m(s)})'(s)| &\geqslant \frac{K\xi}{D_2^2}\left(\frac{D_2^{-1}|(f^{m(s)-2})'(f(c))| \cdot |f''(\zeta)|}{\xi}\right)^{\frac{1}{2}} \\
&\geqslant (K\Lambda)^{\frac{m(s)}{8} - \frac{1}{8}} \\
&\geqslant (K\Lambda)^{\frac{m(s)}{16}}.
\end{aligned}$$

∎

5.1.4. *Verification of (H5) and (H7).* The following lemma facilitates the verification of **(H5)**.

Lemma 5.8 ([12, 11]). *Let $f = f_{a^*}$. Suppose that for all $x \in C(f_{a^*})$, we have*

$$\sum_{k=0}^{\infty} \frac{1}{|(f^k)'(f(x))|} < \infty.$$

Then for each $x \in C(f_{a^*})$,

$$\sum_{k=0}^{\infty} \frac{[(\partial_a f_a)(f^k(x))]_{a=a^*}}{(f^k)'(f(x))} = \left[\frac{d}{da}f_a(x(a)) - \frac{d}{da}p(a)\right]_{a=a^*}.$$

Property **(H5)** follows from Lemma 5.8 for Λ sufficiently large. To see this, suppose $f_{a^*} \in \mathcal{M}$ and let $c \in C(f_{a^*})$. For $k \in \mathbb{Z}^+$, we have

$$|(f_{a^*}^k)'(f(c))| \geqslant \left(K_5\Lambda^{\frac{1}{4}}\right)^k$$

by Lemma 5.1(c). Since $K_6^{-1} \leqslant \frac{\partial}{\partial a}f(s,a) \leqslant K_6$, we conclude that if Λ is sufficiently large, then

$$\sum_{k=0}^{\infty} \frac{[\partial_a f_a(f_{a^*}^k(c))]_{a=a^*}}{(f_{a^*}^k)'(f_{a^*}(c))} \geqslant K_6^{-1} - \sum_{k=1}^{\infty} \frac{K_6}{\left(K_5\Lambda^{\frac{1}{4}}\right)^k} > 0.$$

Property **(H7)** follows from Lemma 5.1 and Proposition 5.6 provided Λ is sufficiently large.

APPENDIX A. SOME PROOFS

We assume throughout Section A that $L = 1$. Notice that if $\boldsymbol{V}$ denotes a vector field, then

$$(A.1) \qquad \frac{d\boldsymbol{z}}{ds} = \boldsymbol{V} \implies \frac{d\|\boldsymbol{z}\|}{ds} = \frac{1}{2\|\boldsymbol{z}\|}\frac{d\|\boldsymbol{z}\|^2}{ds} = \frac{\boldsymbol{z}}{\|\boldsymbol{z}\|} \cdot \frac{d\boldsymbol{z}}{ds} = \frac{\boldsymbol{z}}{\|\boldsymbol{z}\|} \cdot \boldsymbol{V}.$$

We will use this fact together with the following Grönwall-type inequality:

Lemma A.1. *Assume that β is a constant, the function φ is continuous on the interval $[\hat{s}, \check{s}]$, and that the function u is differentiable and satisfies $\frac{du}{ds} \leqslant \beta u + \varphi$ on $(\hat{s}, \check{s})$. Then, for all $s \in (\hat{s}, \check{s})$,*

$$u(s) \leqslant u(\hat{s})e^{\beta(s-\hat{s})} + \int_{\hat{s}}^{s} e^{\beta(s-\tau)}\varphi(\tau)\,d\tau.$$

Proof. Suppose $v(s) = u(\hat{s})e^{\beta(s-\hat{s})} + \int_{\hat{s}}^{s} e^{\beta(s-\tau)}\varphi(\tau)\,d\tau$. Then v satisfies the equation $\frac{dv}{ds}(s) = \beta v(s) + \varphi(s)$ with $v(\hat{s}) = u(\hat{s})$. Since $u - v$ is differentiable, $\frac{d}{ds}(u-v) \leqslant \beta(u-v)$, and $(u-v)(\hat{s}) = 0$, a standard Grönwall argument shows that $u \leqslant v$. ∎

We get immediately

Corollary A.2. *Suppose that in Lemma A.1 $\frac{du}{ds}(s) \leqslant \frac{\lambda_1}{2}u + C_0 e^{\lambda_1(s-\hat{s})}$. Then*

$$u(s) \leqslant \left(u(\hat{s}) + \frac{2C_0}{|\lambda_1|}\right)e^{\frac{\lambda_1}{2}(s-\hat{s})}.$$

Our first application of a Grönwall inequality is

Lemma A.3. *Assume $\boldsymbol{z}$ solves the forced equation* (3.8b) *with $\boldsymbol{z}(s_0) = \boldsymbol{z}_0$ and fix a constant $K > 0$. If $\varepsilon/|\lambda_1|$ is sufficiently small,*

$$(A.2) \qquad \|\partial_s^m \partial_{s_0}^l \boldsymbol{z}(s)\| \leqslant C\varepsilon \qquad (0 \leqslant l + m \leqslant 3)$$

as long as $s - s_0 \leqslant K$. Moreover,

$$(A.3) \qquad \|\partial_{z_0}\boldsymbol{z} - 1\| \leqslant C|\lambda_1|.$$

Proof. Equation (3.8b) reads

$$(A.4) \qquad \frac{d\boldsymbol{z}}{ds} = \mathsf{A}\boldsymbol{z} + \boldsymbol{h}_3(s, \boldsymbol{z}) \qquad \text{with} \qquad \boldsymbol{h}_3(s, \boldsymbol{z}) = \mathcal{O}_s(\varepsilon) + \mathcal{O}_{s,z}(\varepsilon\boldsymbol{z}) + \mathcal{O}_{s,z}(\varepsilon^2) + \mathcal{O}_{s,z}(\|\boldsymbol{z}\|^2).$$

Assuming $\|\boldsymbol{z}\|/|\lambda_1|$ and $\varepsilon/|\lambda_1|$ are sufficiently small, (A.1) implies

$$\frac{d\|\boldsymbol{z}\|}{ds} \leqslant \frac{\lambda_1}{2}\|\boldsymbol{z}\| + C_0\varepsilon.$$

By Lemma A.1,

$$\|\boldsymbol{z}(s)\| \leqslant \|\boldsymbol{z}_0\|e^{\frac{\lambda_1}{2}(s-s_0)} + \frac{2C_0\varepsilon}{|\lambda_1|}\left(1 - e^{\frac{\lambda_1}{2}(s-s_0)}\right) \leqslant \|\boldsymbol{z}_0\|e^{\frac{\lambda_1}{2}(s-s_0)} + C_0\varepsilon(s - s_0).$$

For $s - s_0 \leqslant K$ we get $\|\boldsymbol{z}(s)\|/|\lambda_1| \leqslant \|\boldsymbol{z}_0\|/|\lambda_1| + C_0 K\varepsilon/|\lambda_1|$, which proves the assumption legitimate.

Differentiating (A.4) with respect to s up to two times yields an expression for $\partial_s^m \boldsymbol{z}(s)$. One immediately obtains

$$\|\partial_s^m \boldsymbol{z}(s)\| \leqslant C\varepsilon$$

for $0 \leqslant m \leqslant 3$ and $s - s_0 \leqslant K$.

Equation (A.4) implies

$$(A.5) \qquad \boldsymbol{z}(s) = e^{(s-s_0)\mathsf{A}}\boldsymbol{z}_0 + \int_{s_0}^{s} e^{(s-\tau)\mathsf{A}}\boldsymbol{h}_3(\tau, \boldsymbol{z}(\tau))\,d\tau.$$

Differentiating this with respect to s_0 up to three times and evaluating at $s = s_0$ yields

$$\partial_{s_0}\boldsymbol{z}(s_0) = -\mathsf{A}\boldsymbol{z}_0 - \boldsymbol{h}_3(s_0, \boldsymbol{z}_0)$$

$$\partial_{s_0}^2\boldsymbol{z}(s_0) = \mathsf{A}^2\boldsymbol{z}_0 + \mathsf{A}\boldsymbol{h}_3(s_0, \boldsymbol{z}_0) - \partial_s\boldsymbol{h}_3(s_0, \boldsymbol{z}_0) - D\boldsymbol{h}_3(s_0, \boldsymbol{z}_0)\partial_{s_0}\boldsymbol{z}(s_0)$$

$$\partial_{s_0}^3\boldsymbol{z}(s_0) = -\mathsf{A}^3\boldsymbol{z}_0 - \mathsf{A}^2\boldsymbol{h}_3(s_0, \boldsymbol{z}_0) + 2\mathsf{A}\partial_s\boldsymbol{h}_3(s_0, \boldsymbol{z}_0) - \partial_s^2\boldsymbol{h}_3(s_0, \boldsymbol{z}_0) + \mathsf{A}D\boldsymbol{h}_3(s_0, \boldsymbol{z}_0)\partial_{s_0}\boldsymbol{z}(s_0)$$

$$- D(\partial_s\boldsymbol{h}_3)(s_0, \boldsymbol{z}_0)\partial_{s_0}\boldsymbol{z}(s_0) - D\boldsymbol{h}_3(s_0, \boldsymbol{z}_0)\frac{d}{ds_0}\partial_{s_0}\boldsymbol{z}(s_0) - D^2\boldsymbol{h}_3(s_0, \boldsymbol{z}_0)(\partial_{s_0}\boldsymbol{z}(s_0), \partial_{s_0}\boldsymbol{z}(s_0))$$

$$- D\boldsymbol{h}_3(s_0, \boldsymbol{z}_0)\partial_{s_0}^2\boldsymbol{z}(s_0).$$

Clearly, for $1 \leqslant l \leqslant 3$,

$$\|\partial_{s_0}^l\boldsymbol{z}(s_0)\| \leqslant C\varepsilon.$$

18

Such initial conditions are needed for analyzing the variational equations

(A.6) $\qquad \dfrac{d}{ds}\partial_{s_0}\boldsymbol{z} = (\mathsf{A} + D\boldsymbol{h}_3(s,\boldsymbol{z}))\,\partial_{s_0}\boldsymbol{z}$

(A.7) $\qquad \dfrac{d}{ds}\partial_{s_0}^2\boldsymbol{z} = (\mathsf{A} + D\boldsymbol{h}_3(s,\boldsymbol{z}))\,\partial_{s_0}^2\boldsymbol{z} + D^2\boldsymbol{h}_3(s,\boldsymbol{z})(\partial_{s_0}\boldsymbol{z},\partial_{s_0}\boldsymbol{z})$

$\qquad \dfrac{d}{ds}\partial_{s_0}^3\boldsymbol{z} = (\mathsf{A} + D\boldsymbol{h}_3(s,\boldsymbol{z}))\,\partial_{s_0}^3\boldsymbol{z} + 3D^2\boldsymbol{h}_3(s,\boldsymbol{z})(\partial_{s_0}^2\boldsymbol{z},\partial_{s_0}\boldsymbol{z}) + D^3\boldsymbol{h}_3(s,\boldsymbol{z})(\partial_{s_0}\boldsymbol{z},\partial_{s_0}\boldsymbol{z},\partial_{s_0}\boldsymbol{z}).$

One then checks recursively, using (A.1) and Corollary A.2, that

$$\|\partial_{s_0}^l\boldsymbol{z}(s)\| \leqslant C\varepsilon e^{\frac{\lambda_1}{2}(s-s_0)} \leqslant C\varepsilon$$

hold for $1 \leqslant l \leqslant 3$ and $s - s_0 \leqslant K$.

Equations (A.6) and (A.7) provide us with an expression for $\partial_s\partial_{s_0}^l\boldsymbol{z}(s)$ with $l = 1$ and $l = 2$. Moreover, (A.6) can be differentiated with respect to s to yield an expression for $\partial_s^2\partial_{s_0}\boldsymbol{z}(s)$. The bounds in (A.2) are then readily obtained.

Finally, we will prove (A.3). To this end, notice that

(A.8) $$\dfrac{d}{ds}\partial_{z_0}\boldsymbol{z} = (\mathsf{A} + D\boldsymbol{h}_3(s,\boldsymbol{z}))\,\partial_{z_0}\boldsymbol{z}.$$

In particular, each row, $\partial_{z_{0,i}}\boldsymbol{z}$, of $\partial_{z_0}\boldsymbol{z}$ satisfies this equation. Hence, by principle (A.1), $\frac{d}{ds}\|\partial_{z_{0,i}}\boldsymbol{z}\| \leqslant \frac{\lambda_1}{2}\|\partial_{z_{0,i}}\boldsymbol{z}\|$, so that the matrix $\partial_{z_0}\boldsymbol{z}$ remains perpetually bounded. Integrating both sides of (A.8) from s_0 to s and recalling $\partial_{z_0}\boldsymbol{z}(s_0) = \mathbb{1}$ gives

$$\|\partial_{z_0}\boldsymbol{z}(s) - \mathbb{1}\| \leqslant |s - s_0|\left(|\lambda_{n-1}| + \sup_{s_0 \leqslant s' \leqslant s}\|D\boldsymbol{h}_3(s',\boldsymbol{z})\| \sup_{s_0 \leqslant s' \leqslant s}\|\partial_{z_0}\boldsymbol{z}(s')\|\right),$$

where $\|\cdot\|$ now denotes the matrix norm induced by the Euclidean norm. This estimate implies (A.3). $\blacksquare$

Proof of Proposition 3.1. Throughout the proof, $\|\cdot\|_{C^3}$ will stand for the C^3-norm with respect to s_0.

By (3.8a) and (3.14), $\tilde{s}$ and $\hat{s}$ have to satisfy

(A.9) $$\int_{s_0}^{\tilde{s}} b_0(\tau)\,d\tau = \rho = \int_{s_0}^{\hat{s}} b_0(\tau) + v(\tau)\,d\tau,$$

where $v(s) = \boldsymbol{b}_1^\mathsf{T}(s)\mathsf{P}(s)\boldsymbol{z}(s) + \mathcal{O}_{s,\boldsymbol{z}}(\varepsilon) + \mathcal{O}_{s,\boldsymbol{z}}(\varepsilon\|\boldsymbol{z}(s)\|) + \mathcal{O}_{s,\boldsymbol{z}}(\|\boldsymbol{z}(s)\|^2)$ and $\boldsymbol{z} = \boldsymbol{z}(s)$ solves (3.8b) with $\boldsymbol{z}(s_0) = \boldsymbol{z}_0$.

We use the implicit function theorem to find $\tilde{s}$. Clearly, $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R} : (s_0, s) \mapsto \int_{s_0}^s b_0(\tau)\,d\tau - \rho$ is C^3. Observe that $F(s_0, s_0) = -\rho$ and $\lim_{s\to\infty} F(s_0, s) = \infty$ as $\min b_0 = m > 0$. By the intermediate value theorem, there exists a number $\tilde{s}$ such that $F(s_0, \tilde{s}) = 0$. Because $\frac{\partial}{\partial s}F(s_0, s) = b_0(s) \geqslant m$, the implicit function theorem implies that $\tilde{s}$ is a C^3-function of s_0. Notice that $F(s_0 + 1, s + 1) \equiv F(s_0, s)$, so that $\tilde{s}(s_0 + 1) = \tilde{s}(s_0) + 1$ which implies that $s_0 \mapsto \tilde{s}(s_0) - s_0$ is periodic.

Now that we have $\tilde{s}$, let us define the function

$$g(\xi) := -\rho + \int_{s_0}^{\tilde{s}+\xi} b_0(\tau) + v(\tau)\,d\tau.$$

Notice that, denoting $\xi_1 = \hat{s} - \tilde{s}$, the right side of (A.9) is equivalent to $g(\xi_1) = 0$. The Taylor expansion $g(\xi) = g(0) + g'(0)\xi + \delta_2 g(\xi)$ yields

$$G(\xi) := -\frac{1}{g'(0)}\left(g(0) + \delta_2 g(\xi)\right) = \xi,$$

which we regard, for all fixed $\boldsymbol{z}_0$, as a fixed point equation on the space of C^3 functions $\xi = \xi(s_0)$. Assuming G is a contraction in a closed, origin-centered, ball $\bar{B}_r \subset C^3$ of radius r, there exists a unique solution, ξ_1, to $G(\xi) = \xi$ inside the ball. Next, we prove that for a suitably small value of r, G is indeed a contraction.

First, notice that

$$g(0) = \int_{s_0}^{\tilde{s}} v(\tau)\, d\tau = (\tilde{s} - s_0) \int_0^1 v((1-\tau)s_0 + \tau\tilde{s})\, d\tau$$

$$g'(0) = b_0(\tilde{s}) + v(\tilde{s})$$

$$\delta_2 g(\xi) = \xi^2 \int_0^1 (1-\tau)\, g''(\xi\tau)\, d\tau = \xi^2 \int_0^1 (1-\tau)\, (b_0' + v')(\tilde{s} + \xi\tau)\, d\tau$$

are smooth functions of s_0. Because $\tilde{s}$ is C^3 in s_0 and $\inf_{s_0} g'(0) > 0$, the bounds (A.2) yield

$$\left\| \frac{1}{g'(0)} \right\|_{C^3} \leqslant C \quad \text{and} \quad \left\| \frac{g(0)}{g'(0)} \right\|_{C^3} \leqslant C\varepsilon.$$

Moreover,

$$\|\delta_2 g(\xi)\|_{C^3} \leqslant C\|\xi\|_{C^3}^2 \sup_{\zeta \in \bar{B}_r} \|(b_0' + v')(\tilde{s} + \zeta)\|_{C^3}.$$

Hence, $\|G(\xi)\|_{C^3} \leqslant C_0(\varepsilon + r^2)$ for some C_0. Choosing $r = 2C_0\varepsilon$, we have $G(\bar{B}_r) \subset \bar{B}_r$ for ε small enough.

Second, let ξ^1 and ξ^2 be elements of $\bar{B}_r$. Since the map $\xi \mapsto G(\xi)$ is differentiable and the operator norm of the derivative obeys the bound $\sup_{\xi \in \bar{B}_r} \|DG(\xi)\|_{\mathcal{L}(C^3)} \leqslant C \sup_{\xi \in \bar{B}_r} \|D\delta_2 g(\xi)\|_{\mathcal{L}(C^3)} \leqslant Cr$, the mean value theorem yields $\|G(\xi^1) - G(\xi^2)\|_{C^3} \leqslant Cr\|\xi^1 - \xi^2\|_{C^3}$. Hence, G is a contraction on $\bar{B}_r$ if ε is sufficiently small.

We will now prove that the fixed point, ξ_1, of G is a periodic function of s_0. Let us denote $\boldsymbol{z}(s, s_0, \boldsymbol{z}_0)$ the solution and $v(s)|_{s_0}$ the function v defined above, when the initial condition $\boldsymbol{z}(s_0) = \boldsymbol{z}_0$ is being used. Because $\boldsymbol{h}_3(s+2, \boldsymbol{z}) = \boldsymbol{h}_3(s, \boldsymbol{z})$ in (A.4), we have $\boldsymbol{z}(s+2, s_0+2, \boldsymbol{z}_0) = \boldsymbol{z}(s, s_0, \boldsymbol{z}_0)$ and $v(s+2)|_{s_0+2} = v(s)|_{s_0}$. Since $g(\xi_1) = 0$ for all values of s_0 and $\tilde{s}(s_0 + 2) = \tilde{s}(s_0) + 2$, the computation

$$g(\xi_1)(s_0 + 2) = -\rho + \int_{s_0+2}^{\tilde{s}(s_0+2)+\xi_1(s_0+2)} b_0(\tau) + v(\tau)|_{s_0+2}\, d\tau = -\rho + \int_{s_0+2}^{\tilde{s}(s_0)+2+\xi_1(s_0+2)} b_0(\tau) + v(\tau)|_{s_0+2}\, d\tau$$

$$= -\rho + \int_{s_0}^{\tilde{s}(s_0)+\xi_1(s_0+2)} b_0(\tau+2) + v(\tau+2)|_{s_0+2}\, d\tau = -\rho + \int_{s_0}^{\tilde{s}(s_0)+\xi_1(s_0+2)} b_0(\tau) + v(\tau)|_{s_0}\, d\tau$$

$$= g(\xi_1)(s_0) + \int_{\tilde{s}(s_0)+\xi_1(s_0)}^{\tilde{s}(s_0)+\xi_1(s_0+2)} b_0(\tau) + v(\tau)|_{s_0}\, d\tau,$$

implies that the last integral vanishes despite the fact that the integrand is positive, so we must have $\xi_1(s_0 + 2) = \xi_1(s_0)$.

As the last step, we will bound the difference $\hat{\boldsymbol{z}} - \tilde{\boldsymbol{z}}$. Let $\boldsymbol{z}^{(1)}$ and $\boldsymbol{z}^{(2)}$ solve (3.8b) and (3.13b), respectively, with the initial condition $\boldsymbol{z}^{(1)}(s_0) = \boldsymbol{z}^{(2)}(s_0) = \boldsymbol{z}_0$. Both of these are C^3 functions of $(s_0, \boldsymbol{z}_0)$ by the smoothness of the vector fields. By definition, $\hat{\boldsymbol{z}} = \boldsymbol{z}^{(1)}(\hat{s})$ and $\tilde{\boldsymbol{z}} = \boldsymbol{z}^{(2)}(\tilde{s})$. We need a bound on the C^3 norm of the difference $\boldsymbol{\xi}_2(s_0) = \hat{\boldsymbol{z}} - \tilde{\boldsymbol{z}}$ for fixed $\boldsymbol{z}_0$. Notice that $\boldsymbol{\xi}_2(s_0) = (\boldsymbol{z}^{(1)} - \boldsymbol{z}^{(2)})(\hat{s}) + (\boldsymbol{z}^{(2)}(\hat{s}) - \boldsymbol{z}^{(2)}(\tilde{s}))$.

Observe that the difference $\boldsymbol{\delta} = \boldsymbol{z}^{(1)} - \boldsymbol{z}^{(2)}$ satisfies the differential equation

$$\frac{d\boldsymbol{\delta}}{ds} = \mathsf{A}\boldsymbol{z}^{(1)} + \mathcal{O}_{s,\boldsymbol{z}^{(1)}}(\varepsilon\boldsymbol{z}^{(1)}) + \mathcal{O}_{s,\boldsymbol{z}^{(1)}}(\varepsilon^2) + \mathcal{O}_{s,\boldsymbol{z}^{(1)}}(\|\boldsymbol{z}^{(1)}\|^2) = \boldsymbol{w}.$$

Here $\boldsymbol{z}^{(1)}$, and hence $\boldsymbol{w}$, is to be regarded as a predetermined function for which we already have good bounds. Indeed, let $\mathbb{S} = \{(s_0, s) : 0 \leqslant s_0 < 2,\ s_0 \leqslant s \leqslant K\}$ and $\|\cdot\|_{C_{\mathbb{S}}^3}$ stand for the C^3 norm on this set. According to (A.2), $\|\boldsymbol{w} - \mathsf{A}\boldsymbol{z}^{(1)}\|_{C_{\mathbb{S}}^3} \leqslant C\varepsilon^2$ whereas, recalling that all eigenvalues of A are proportional to λ_1, $\|\mathsf{A}\boldsymbol{z}^{(1)}\|_{C_{\mathbb{S}}^3} \leqslant C|\lambda_1|\varepsilon$. In other words, $\|\boldsymbol{w}\|_{C_{\mathbb{S}}^3} \leqslant C|\lambda_1|\varepsilon$. As $\boldsymbol{\delta}(s_0) = 0$, we have

$$(\boldsymbol{z}^{(1)} - \boldsymbol{z}^{(2)})(\hat{s}) = \boldsymbol{\delta}(\hat{s}) = \int_{s_0}^{\hat{s}} \boldsymbol{w}(\tau)\, d\tau.$$

Because $\hat{s}$ is C^3 in s_0, it follows that $\|(\boldsymbol{z}^{(1)} - \boldsymbol{z}^{(2)})(\hat{s})\|_{C^3} \leqslant C|\lambda_1|\varepsilon$. By (3.13b), the remaining contribution reads

$$\boldsymbol{z}^{(2)}(\hat{s}) - \boldsymbol{z}^{(2)}(\tilde{s}) = \int_{\tilde{s}}^{\hat{s}} \frac{\varepsilon \mathsf{P}^{-1}(\tau)\boldsymbol{\phi}(\boldsymbol{0}, \tau)}{\psi_n^{(0)}(\tau)}\, d\tau.$$

We have seen above that $\|\hat{s} - \tilde{s}\|_{C^3} \leqslant C\varepsilon$, which implies $\|\boldsymbol{z}^{(2)}(\hat{s}) - \boldsymbol{z}^{(2)}(\tilde{s})\|_{C^3} \leqslant C\varepsilon^2$ and finally $\|\boldsymbol{\xi}_2(s_0)\|_{C^3} \leqslant C|\lambda_1|\varepsilon$. ∎

Remark A.4. It follows from the previous proof that, under the conditions of Proposition 3.1,

$$(A.10) \qquad \|\partial_{\boldsymbol{z}_0}\hat{s}\| \leqslant C.$$

Indeed, $\partial_{\boldsymbol{z}_0}\hat{s} = \partial_{\boldsymbol{z}_0}\xi_1$, as $\partial_{\boldsymbol{z}_0}\tilde{s} = 0$. From the fixed point equation $\xi_1 = G(\xi_1)$ we get $\partial_{\boldsymbol{z}_0}\xi_1 = (1 - DG(\xi_1))^{-1}(\partial_{\boldsymbol{z}_0}G)(\xi_1)$ and then the claimed bound. Moreover,

$$(A.11) \qquad \|\partial_{\boldsymbol{z}_0}\hat{\boldsymbol{z}} - 1\| \leqslant C|\lambda_1|.$$

For let $\boldsymbol{z}(s) = \boldsymbol{z}(s; s_0, \boldsymbol{z}_0)$ be the solution to (3.8b) with $\boldsymbol{z}(s_0; s_0, \boldsymbol{z}_0) = \boldsymbol{z}_0$ and recall that $\hat{s}$ depends on $(s_0, \boldsymbol{z}_0)$. By definition, $\hat{\boldsymbol{z}} = \boldsymbol{z}(\hat{s}; s_0, \boldsymbol{z}_0)$ so that $\partial_{\boldsymbol{z}_0}\hat{\boldsymbol{z}} = \partial_s \boldsymbol{z}(\hat{s})\partial_{\boldsymbol{z}_0}\hat{s} + \partial_{\boldsymbol{z}_0}\boldsymbol{z}(\hat{s}) = 1 + \mathcal{O}(\lambda_1)$ by the bounds in Lemma A.3.

Let us view the solution

$$(A.12) \qquad \boldsymbol{z}(s) = \boldsymbol{z}(s, \hat{s}, \hat{\boldsymbol{z}}), \qquad \boldsymbol{z}(\hat{s}) \equiv \hat{\boldsymbol{z}}$$

to equation (3.7b) as a function of three variables and abbreviate $\partial_s = \partial/\partial_s$, $\hat{\partial}_i = \partial/\partial\hat{z}_i$, $\hat{\partial}_{i_1 \cdots i_k} = \hat{\partial}_{i_1} \cdots \hat{\partial}_{i_k}$, and $\partial_{\hat{s}} = \partial/\partial\hat{s}$.

Proposition A.5. *Assuming $\|\hat{\boldsymbol{z}}\|/|\lambda_1|$ is small enough, we have, for $0 \leqslant k + l + m \leqslant 3$ and $s \geqslant \hat{s}$, the following bounds:*

$$\left\|\partial_s^m \partial_{\hat{s}}^l \boldsymbol{z}(s)\right\| \leqslant C\|\hat{\boldsymbol{z}}\|e^{\frac{\lambda_1}{2}(s-\hat{s})}$$

$$\left\|\partial_s^m \partial_{\hat{s}}^l \hat{\partial}_{i_1 \cdots i_k}\boldsymbol{z}(s)\right\| \leqslant \frac{C}{|\lambda_1|^{k-1}}e^{\frac{\lambda_1}{2}(s-\hat{s})} \qquad (k > 0).$$

Proof. The initial conditions $(\partial\boldsymbol{z}/\partial\hat{\boldsymbol{z}})(\hat{s}) = 1$, $\hat{\partial}_{ij}\boldsymbol{z}(\hat{s}) = \boldsymbol{0}$, and $\hat{\partial}_{ijk}\boldsymbol{z}(\hat{s}) = \boldsymbol{0}$ follow from (A.12), as the $\hat{\boldsymbol{z}}$-derivatives can be computed *after* evaluating $\boldsymbol{z}$ at $s = \hat{s}$. Similarly, taking $\hat{s}$-derivatives of

$$\boldsymbol{z}(s) = e^{(s-\hat{s})\mathsf{A}}\left(\hat{\boldsymbol{z}} + \int_{\hat{s}}^s e^{-(\tau-\hat{s})\mathsf{A}}\boldsymbol{h}_1(\tau, \boldsymbol{z}(\tau))\,d\tau\right)$$

yields first, analogously to how the identities below (A.5) were obtained,

$$\partial_{\hat{s}}\boldsymbol{z}(\hat{s}) = -\mathsf{A}\hat{\boldsymbol{z}} - \boldsymbol{h}_1(\hat{s}, \hat{\boldsymbol{z}})$$

$$\partial_{\hat{s}}^2\boldsymbol{z}(\hat{s}) = \mathsf{A}^2\hat{\boldsymbol{z}} + \mathsf{A}\boldsymbol{h}_1(\hat{s}, \hat{\boldsymbol{z}}) - \partial_s\boldsymbol{h}_1(\hat{s}, \hat{\boldsymbol{z}}) - D\boldsymbol{h}_1(\hat{s}, \hat{\boldsymbol{z}})\partial_{\hat{s}}\boldsymbol{z}(\hat{s})$$

$$\partial_{\hat{s}}^3\boldsymbol{z}(\hat{s}) = -\mathsf{A}^3\hat{\boldsymbol{z}} - \mathsf{A}^2\boldsymbol{h}_1(\hat{s}, \hat{\boldsymbol{z}}) + 2\mathsf{A}\partial_s\boldsymbol{h}_1(\hat{s}, \hat{\boldsymbol{z}}) - \partial_s^2\boldsymbol{h}_1(\hat{s}, \hat{\boldsymbol{z}}) + \mathsf{A}D\boldsymbol{h}_1(\hat{s}, \hat{\boldsymbol{z}})\partial_{\hat{s}}\boldsymbol{z}(\hat{s})$$

$$\qquad - D(\partial_s\boldsymbol{h}_1)(\hat{s}, \hat{\boldsymbol{z}})\partial_{\hat{s}}\boldsymbol{z}(\hat{s}) - D\boldsymbol{h}_1(\hat{s}, \hat{\boldsymbol{z}})\frac{d}{d\hat{s}}\partial_{\hat{s}}\boldsymbol{z}(\hat{s}) - D^2\boldsymbol{h}_1(\hat{s}, \hat{\boldsymbol{z}})(\partial_{\hat{s}}\boldsymbol{z}(\hat{s}), \partial_{\hat{s}}\boldsymbol{z}(\hat{s}))$$

$$\qquad - D\boldsymbol{h}_1(\hat{s}, \hat{\boldsymbol{z}})\partial_{\hat{s}}^2\boldsymbol{z}(\hat{s}).$$

These formulas can then be differentiated with respect to $\hat{\boldsymbol{z}}$ in order to find higher-order initial conditions. As $\boldsymbol{h}_1(s, \boldsymbol{z}) = \mathcal{O}(\|\boldsymbol{z}\|^2)$, we obtain the following estimates:

$$\|\partial_{\hat{s}}^l\boldsymbol{z}(\hat{s})\| \leqslant C\|\hat{\boldsymbol{z}}\| \qquad l = 1, 2, 3$$

$$\|\partial_{\hat{s}}^l\hat{\partial}_i\boldsymbol{z}(\hat{s})\| \leqslant C \qquad l = 1, 2$$

$$\|\partial_{\hat{s}}\hat{\partial}_{ij}\boldsymbol{z}(\hat{s})\| \leqslant C.$$

Combining (3.7b) and (A.1),

$$\frac{d}{ds}\|\boldsymbol{z}\| = \frac{\boldsymbol{z} \cdot \mathsf{A}\boldsymbol{z} + \boldsymbol{z} \cdot \boldsymbol{h}_1(s, \boldsymbol{z})}{\|\boldsymbol{z}\|} \leqslant \lambda_1\|\boldsymbol{z}\| + \|\boldsymbol{h}_1(s, \boldsymbol{z})\| \leqslant \frac{\lambda_1}{2}\|\boldsymbol{z}\|$$

if $\|\boldsymbol{z}\|/|\lambda_1|$ is small enough. Below, $\|\hat{\boldsymbol{z}}\|/|\lambda_1|$ will always be assumed small enough. Thus, for all $s > \hat{s}$,

$$(A.13) \qquad \|\boldsymbol{z}(s)\| \leqslant \|\hat{\boldsymbol{z}}\|e^{\frac{\lambda_1}{2}(s-\hat{s})}.$$

Differentiating (3.7b) with respect to various components of $\hat{\boldsymbol{z}}$, we obtain the variational equations

(A.14) $$\frac{d}{ds}\hat{\partial}_i\boldsymbol{z} = (\mathsf{A} + D\boldsymbol{h}_1(s,\boldsymbol{z}))\,\hat{\partial}_i\boldsymbol{z}$$

(A.15) $$\frac{d}{ds}\hat{\partial}_{ij}\boldsymbol{z} = (\mathsf{A} + D\boldsymbol{h}_1(s,\boldsymbol{z}))\,\hat{\partial}_{ij}\boldsymbol{z} + D^2\boldsymbol{h}_1(s,\boldsymbol{z})(\hat{\partial}_i\boldsymbol{z},\hat{\partial}_j\boldsymbol{z})$$

(A.16) $$\frac{d}{ds}\hat{\partial}_{ijk}\boldsymbol{z} = (\mathsf{A} + D\boldsymbol{h}_1(s,\boldsymbol{z}))\,\hat{\partial}_{ijk}\boldsymbol{z} + D^2\boldsymbol{h}_1(s,\boldsymbol{z})(\hat{\partial}_i\boldsymbol{z},\hat{\partial}_{jk}\boldsymbol{z})$$
$$+ D^2\boldsymbol{h}_1(s,\boldsymbol{z})(\hat{\partial}_k\boldsymbol{z},\hat{\partial}_{ij}\boldsymbol{z}) + D^2\boldsymbol{h}_1(s,\boldsymbol{z})(\hat{\partial}_j\boldsymbol{z},\hat{\partial}_{ki}\boldsymbol{z}) + D^3\boldsymbol{h}_1(s,\boldsymbol{z})(\hat{\partial}_i\boldsymbol{z},\hat{\partial}_k\boldsymbol{z},\hat{\partial}_{ij}\boldsymbol{z}).$$

Combining (A.14), (A.1), and (A.13), we have

(A.17) $$\|\hat{\partial}_i\boldsymbol{z}(s)\| \leqslant e^{\frac{\lambda_1}{2}(s-\hat{s})}$$

in analogy with (A.13). Combining (A.15), (A.1), (A.13), and (A.17),

$$\frac{d}{ds}\|\hat{\partial}_{ij}\boldsymbol{z}\| \leqslant \frac{\lambda_1}{2}\|\hat{\partial}_{ij}\boldsymbol{z}\| + C\|\hat{\partial}_i\boldsymbol{z}\|\|\hat{\partial}_j\boldsymbol{z}\| \leqslant \frac{\lambda_1}{2}\|\hat{\partial}_{ij}\boldsymbol{z}\| + Ce^{\lambda_1(s-\hat{s})}.$$

Applying Corollary A.2,

(A.18) $$\|\hat{\partial}_{ij}\boldsymbol{z}(s)\| \leqslant \frac{C}{|\lambda_1|}e^{\frac{\lambda_1}{2}(s-\hat{s})}.$$

Similarly, combining (A.16), (A.1), (A.13), (A.17), and (A.18),

(A.19) $$\|\hat{\partial}_{ijk}\boldsymbol{z}(s)\| \leqslant \frac{C}{|\lambda_1|^2}e^{\frac{\lambda_1}{2}(s-\hat{s})}.$$

Differentiating equations (3.7b), (A.14), and (A.15) with respect to $\hat{s}$ produces equations for $\partial_{\hat{s}}^l\boldsymbol{z}$, $\partial_{\hat{s}}^l\hat{\partial}_i\boldsymbol{z}$, and $\partial_{\hat{s}}\hat{\partial}_{ij}\boldsymbol{z}$. For example,

$$\frac{d}{ds}\partial_{\hat{s}}\boldsymbol{z} = (\mathsf{A} + D\boldsymbol{h}_1(s,\boldsymbol{z}))\,\partial_{\hat{s}}\boldsymbol{z}$$
$$\frac{d}{ds}\partial_{\hat{s}}^2\boldsymbol{z} = (\mathsf{A} + D\boldsymbol{h}_1(s,\boldsymbol{z}))\,\partial_{\hat{s}}^2\boldsymbol{z} + D^2\boldsymbol{h}_1(s,\boldsymbol{z})(\partial_{\hat{s}}\boldsymbol{z},\partial_{\hat{s}}\boldsymbol{z}).$$

Such equations can be handled in a similar fashion and it is easy to verify that the additional $\hat{s}$-derivatives do not change the bounds by more than a constant prefactor. The bounds with $m = 0$ in the proposition are now clear. The bounds with $m \neq 0$ follow immediately from the appropriate differential equation; for instance, bounding the right-hand side of (A.14) yields the bound on $\frac{d}{ds}\hat{\partial}_i\boldsymbol{z}$. ∎

Proof of Proposition 3.2. We first view the error terms $\mathcal{E}_k$ with $1 \leqslant k \leqslant 3$ as smooth functions of $(\hat{s}, \hat{\boldsymbol{z}})$ and bound their derivatives with respect to these variables. The bounds on the C^3-norms with respect to s_0 follow by the chain rule. Bounding the C^3-norms of $\mathcal{E}_4$ and $\mathcal{E}_5$ is trivial and is done at the end of the proof. Throughout the proof, $\|\cdot\|_{C^3}$ will stand for the C^3-norm with respect to s_0.

Terms $\mathcal{E}_1$ and $\mathcal{E}_2$. It is convenient to express $\mathcal{E}_1$ and $\mathcal{E}_2$ in the form

$$\mathcal{E}_1 = \int_0^\infty \boldsymbol{b}_1(\hat{s}+\tau)^\mathsf{T}\mathsf{P}(\hat{s}+\tau)\int_0^\tau e^{(\tau-\xi)\mathsf{A}}\boldsymbol{h}_1(\hat{s}+\xi,\boldsymbol{z}(\hat{s}+\xi))\,d\xi\,d\tau$$
$$\mathcal{E}_2 = \int_0^\infty h_2(\hat{s}+\tau,\boldsymbol{z}(\hat{s}+\tau))\,d\tau.$$

First of all, $\|(\boldsymbol{b}_1^\mathsf{T}\mathsf{P})(\hat{s}+\tau)\|_{C^3} \leqslant C\|\boldsymbol{b}_1^\mathsf{T}\mathsf{P}\|_{C^3}$ for every τ, because $\|\hat{s} - s_0\|_{C^3} \leqslant C$, so that

$$\|\mathcal{E}_1\|_{C^3} \leqslant C\|\boldsymbol{b}_1^\mathsf{T}\mathsf{P}\|_{C^3}\int_0^\infty\int_0^\tau e^{\lambda_1(\tau-\xi)}\|\boldsymbol{h}_1(\hat{s}+\xi,\boldsymbol{z}(\hat{s}+\xi))\|_{C^3}\,d\xi\,d\tau$$

$$\|\mathcal{E}_2\|_{C^3} \leqslant \int_0^\infty \|h_2(\hat{s}+\tau,\boldsymbol{z}(\hat{s}+\tau))\|_{C^3}\,d\tau.$$

Since $\boldsymbol{h}_1(s,\boldsymbol{z})$ and $h_2(s,\boldsymbol{z})$ are periodic in the variable s, their partial derivatives of any order (less than four) with respect to s are periodic functions of s and can be bounded exactly as $\boldsymbol{h}_1(s,\boldsymbol{z})$ and $h_2(s,\boldsymbol{z})$. To save a considerable amount of space, we write $\eta = \hat{s} + \xi$ and $\boldsymbol{\zeta} = (\hat{s}+\xi,\boldsymbol{z}(\hat{s}+\xi))$ below.

Notice that the first three total $\hat{s}$-derivatives of $\boldsymbol{z}(\eta)$ are

$$\frac{d}{d\hat{s}}\boldsymbol{z}(\eta) = \partial_s\boldsymbol{z}(\eta) + \partial_{\hat{s}}\boldsymbol{z}(\eta)$$

$$\frac{d^2}{d\hat{s}^2}\boldsymbol{z}(\eta) = \partial_s^2\boldsymbol{z}(\eta) + 2\partial_s\partial_{\hat{s}}\boldsymbol{z}(\eta) + \partial_{\hat{s}}^2\boldsymbol{z}(\eta)$$

$$\frac{d^3}{d\hat{s}^3}\boldsymbol{z}(\eta) = \partial_s^3\boldsymbol{z}(\eta) + 3\partial_s^2\partial_{\hat{s}}\boldsymbol{z}(\eta) + 3\partial_s\partial_{\hat{s}}^2\boldsymbol{z}(\eta) + \partial_{\hat{s}}^3\boldsymbol{z}(\eta).$$

Taking $\hat{\boldsymbol{z}}$-derivatives of the first two formulas above,

$$\frac{d}{d\hat{s}}\hat{\partial}_i\boldsymbol{z}(\eta) = \partial_s\hat{\partial}_i\boldsymbol{z}(\eta) + \partial_{\hat{s}}\hat{\partial}_i\boldsymbol{z}(\eta)$$

$$\frac{d}{d\hat{s}}\hat{\partial}_{ij}\boldsymbol{z}(\eta) = \partial_s\hat{\partial}_{ij}\boldsymbol{z}(\eta) + \partial_{\hat{s}}\hat{\partial}_{ij}\boldsymbol{z}(\eta)$$

$$\frac{d^2}{d\hat{s}^2}\hat{\partial}_i\boldsymbol{z}(\eta) = \partial_s^2\hat{\partial}_i\boldsymbol{z}(\eta) + \partial_s\partial_{\hat{s}}\hat{\partial}_i\boldsymbol{z}(\eta) + \partial_{\hat{s}}^2\hat{\partial}_i\boldsymbol{z}(\eta).$$

Proposition A.5 then implies the bounds for $0 \leqslant k + l \leqslant 3$:

$$\left\|\frac{d^l}{d\hat{s}^l}\boldsymbol{z}(\eta)\right\| \leqslant C\|\hat{\boldsymbol{z}}\|e^{\frac{\lambda_1}{2}\xi}$$

$$\left\|\frac{d^l}{d\hat{s}^l}\hat{\partial}_{i_1\ldots i_k}\boldsymbol{z}(\eta)\right\| \leqslant \frac{C}{|\lambda_1|^{k-1}}e^{\frac{\lambda_1}{2}\xi} \qquad (k > 0).$$

These will be used to bound the C^3-norm of $\boldsymbol{h}_1(\boldsymbol{\zeta})$. To this end, we compute

$$\frac{d}{d\hat{s}}\boldsymbol{h}_1(\boldsymbol{\zeta}) = \partial_s\boldsymbol{h}_1(\boldsymbol{\zeta}) + D\boldsymbol{h}_1(\boldsymbol{\zeta})\frac{d}{d\hat{s}}\boldsymbol{z}(\eta)$$

$$= \mathcal{O}\big(\|\hat{\boldsymbol{z}}\|^2 e^{\lambda_1\xi}\big)$$

$$\frac{d^2}{d\hat{s}^2}\boldsymbol{h}_1(\boldsymbol{\zeta}) = \partial_s^2\boldsymbol{h}_1(\boldsymbol{\zeta}) + 2D(\partial_s\boldsymbol{h}_1)(\boldsymbol{\zeta})\frac{d}{d\hat{s}}\boldsymbol{z}(\eta) + D\boldsymbol{h}_1(\boldsymbol{\zeta})\frac{d^2}{d\hat{s}^2}\boldsymbol{z}(\eta) + D^2\boldsymbol{h}_1(\boldsymbol{\zeta})\left(\frac{d}{d\hat{s}}\boldsymbol{z}(\eta), \frac{d}{d\hat{s}}\boldsymbol{z}(\eta)\right)$$

$$= \mathcal{O}\big(\|\hat{\boldsymbol{z}}\|^2 e^{\lambda_1\xi}\big)$$

$$\frac{d^3}{d\hat{s}^3}\boldsymbol{h}_1(\boldsymbol{\zeta}) = \partial_s^3\boldsymbol{h}_1(\boldsymbol{\zeta}) + D(\partial_s^2\boldsymbol{h}_1)(\boldsymbol{\zeta})\frac{d}{d\hat{s}}\boldsymbol{z}(\eta) + 2D(\partial_s^2\boldsymbol{h}_1)(\boldsymbol{\zeta})\frac{d}{d\hat{s}}\boldsymbol{z}(\eta) + 3D(\partial_s\boldsymbol{h}_1)(\boldsymbol{\zeta})\frac{d^2}{d\hat{s}^2}\boldsymbol{z}(\eta)$$

$$+ 3D^2(\partial_s\boldsymbol{h}_1)(\boldsymbol{\zeta})\left(\frac{d}{d\hat{s}}\boldsymbol{z}(\eta), \frac{d}{d\hat{s}}\boldsymbol{z}(\eta)\right) + D\boldsymbol{h}_1(\boldsymbol{\zeta})\frac{d^3}{d\hat{s}^3}\boldsymbol{z}(\eta) + 3D^2\boldsymbol{h}_1(\boldsymbol{\zeta})\left(\frac{d}{d\hat{s}}\boldsymbol{z}(\eta), \frac{d^2}{d\hat{s}^2}\boldsymbol{z}(\eta)\right)$$

$$+ D^3\boldsymbol{h}_1(\boldsymbol{\zeta})\left(\frac{d}{d\hat{s}}\boldsymbol{z}(\eta), \frac{d}{d\hat{s}}\boldsymbol{z}(\eta), \frac{d}{d\hat{s}}\boldsymbol{z}(\eta)\right).$$

$$= \mathcal{O}\big(\|\hat{\boldsymbol{z}}\|^2 e^{\lambda_1\xi}\big)$$

$$\frac{d}{d\hat{z}_i}\boldsymbol{h}_1(\boldsymbol{\zeta}) = D\boldsymbol{h}_1(\boldsymbol{\zeta})\hat{\partial}_i\boldsymbol{z}(\eta)$$

$$= \mathcal{O}\big(\|\hat{\boldsymbol{z}}\|e^{\lambda_1\xi}\big)$$

$$\frac{d^2}{d\hat{z}_i d\hat{z}_j}\boldsymbol{h}_1(\boldsymbol{\zeta}) = D\boldsymbol{h}_1(\boldsymbol{\zeta})\hat{\partial}_{ij}\boldsymbol{z}(\eta) + D^2\boldsymbol{h}_1(\boldsymbol{\zeta})(\hat{\partial}_i\boldsymbol{z}(\eta),\hat{\partial}_j\boldsymbol{z}(\eta))$$

$$= \mathcal{O}\left(\left(\frac{\|\hat{\boldsymbol{z}}\|}{|\lambda_1|}+1\right)e^{\lambda_1\xi}\right) = \mathcal{O}\big(e^{\lambda_1\xi}\big)$$

$$\frac{d^3}{d\hat{z}_i d\hat{z}_j d\hat{z}_k}\boldsymbol{h}_1(\boldsymbol{\zeta}) = D\boldsymbol{h}_1(\boldsymbol{\zeta})\hat{\partial}_{ijk}\boldsymbol{z}(\eta) + D^2\boldsymbol{h}_1(\boldsymbol{\zeta})(\hat{\partial}_i\boldsymbol{z}(\eta),\hat{\partial}_{jk}\boldsymbol{z}(\eta)) + D^2\boldsymbol{h}_1(\boldsymbol{\zeta})(\hat{\partial}_k\boldsymbol{z}(\eta),\hat{\partial}_{ij}\boldsymbol{z}(\eta))$$

$$+ D^2\boldsymbol{h}_1(\boldsymbol{\zeta})(\hat{\partial}_j\boldsymbol{z}(\eta),\hat{\partial}_{ki}\boldsymbol{z}(\eta)) + D^3\boldsymbol{h}_1(\boldsymbol{\zeta})(\hat{\partial}_i\boldsymbol{z}(\eta),\hat{\partial}_k\boldsymbol{z}(\eta),\hat{\partial}_{ij}\boldsymbol{z}(\eta)).$$

$$= \mathcal{O}\left(\left(\frac{\|\hat{\boldsymbol{z}}\|}{|\lambda_1|^2}+\frac{1}{|\lambda_1|}\right)e^{\lambda_1\xi}\right) = \mathcal{O}\left(\frac{1}{|\lambda_1|}e^{\lambda_1\xi}\right)$$

Taking $\hat{z}$-derivatives of $\frac{d}{d\hat{s}}\boldsymbol{h}_1(\boldsymbol{\zeta})$, $\frac{d^2}{d\hat{s}^2}\boldsymbol{h}_1(\boldsymbol{\zeta})$, and the resulting expression for $\frac{d^2}{d\hat{z}_i d\hat{s}}\boldsymbol{h}_1(\boldsymbol{\zeta})$, we get

$$\frac{d^2}{d\hat{z}_i d\hat{s}}\boldsymbol{h}_1(\boldsymbol{\zeta}) = D(\partial_s\boldsymbol{h}_1)(\boldsymbol{\zeta})\hat{\partial}_i\boldsymbol{z}(\eta) + D\boldsymbol{h}_1(\boldsymbol{\zeta})\frac{d}{d\hat{s}}\hat{\partial}_i\boldsymbol{z}(\eta) + D^2\boldsymbol{h}_1(\boldsymbol{\zeta})\left(\frac{d}{d\hat{s}}\boldsymbol{z}(\eta),\hat{\partial}_i\boldsymbol{z}(\eta)\right)$$

$$= \mathcal{O}\big(\|\hat{\boldsymbol{z}}\|e^{\lambda_1\xi}\big)$$

$$\frac{d^3}{d\hat{z}_i d\hat{s}^2}\boldsymbol{h}_1(\boldsymbol{\zeta}) = D(\partial_s^2\boldsymbol{h}_1)(\boldsymbol{\zeta})\hat{\partial}_i\boldsymbol{z}(\eta) + 2D(\partial_s\boldsymbol{h}_1)(\boldsymbol{\zeta})\frac{d}{d\hat{s}}\hat{\partial}_i\boldsymbol{z}(\eta) + 2D^2(\partial_s\boldsymbol{h}_1)(\boldsymbol{\zeta})\left(\frac{d}{d\hat{s}}\boldsymbol{z}(\eta),\hat{\partial}_i\boldsymbol{z}(\eta)\right)$$

$$+ D\boldsymbol{h}_1(\boldsymbol{\zeta})\frac{d^2}{d\hat{s}^2}\hat{\partial}_i\boldsymbol{z}(\eta) + D^2\boldsymbol{h}_1(\boldsymbol{\zeta})\left(\frac{d^2}{d\hat{s}^2}\boldsymbol{z}(\eta),\hat{\partial}_i\boldsymbol{z}(\eta)\right)$$

$$+ 2D^2\boldsymbol{h}_1(\boldsymbol{\zeta})\left(\frac{d}{d\hat{s}}\hat{\partial}_i\boldsymbol{z}(\eta),\frac{d}{d\hat{s}}\boldsymbol{z}(\eta)\right) + D^3\boldsymbol{h}_1(\boldsymbol{\zeta})\left(\hat{\partial}_i\boldsymbol{z}(\eta),\frac{d}{d\hat{s}}\boldsymbol{z}(\eta),\frac{d}{d\hat{s}}\boldsymbol{z}(\eta)\right)$$

$$= \mathcal{O}\big(\|\hat{\boldsymbol{z}}\|e^{\lambda_1\xi}\big)$$

$$\frac{d^3}{d\hat{z}_i d\hat{z}_j d\hat{s}}\boldsymbol{h}_1(\boldsymbol{\zeta}) = D(\partial_s\boldsymbol{h}_1)(\boldsymbol{\zeta})\hat{\partial}_{ij}\boldsymbol{z}(\eta) + D^2(\partial_s\boldsymbol{h}_1)(\boldsymbol{\zeta})\left(\hat{\partial}_i\boldsymbol{z}(\eta),\hat{\partial}_j\boldsymbol{z}(\eta)\right) + D\boldsymbol{h}_1(\boldsymbol{\zeta})\frac{d}{d\hat{s}}\hat{\partial}_{ij}\boldsymbol{z}(\eta)$$

$$+ D^2\boldsymbol{h}_1(\boldsymbol{\zeta})\left(\frac{d}{d\hat{s}}\hat{\partial}_i\boldsymbol{z}(\eta),\hat{\partial}_j\boldsymbol{z}(\eta)\right) + D^2\boldsymbol{h}_1(\boldsymbol{\zeta})\left(\frac{d}{d\hat{s}}\hat{\partial}_j\boldsymbol{z}(\eta),\hat{\partial}_i\boldsymbol{z}(\eta)\right)$$

$$+ D^2\boldsymbol{h}_1(\boldsymbol{\zeta})\left(\frac{d}{d\hat{s}}\boldsymbol{z}(\eta),\hat{\partial}_{ij}\boldsymbol{z}(\eta)\right) + D^3\boldsymbol{h}_1(\boldsymbol{\zeta})\left(\frac{d}{d\hat{s}}\boldsymbol{z}(\eta),\hat{\partial}_i\boldsymbol{z}(\eta),\hat{\partial}_j\boldsymbol{z}(\eta)\right).$$

$$= \mathcal{O}\left(\left(\frac{\|\hat{\boldsymbol{z}}\|}{|\lambda_1|}+1\right)e^{\lambda_1\xi}\right) = \mathcal{O}\big(e^{\lambda_1\xi}\big)$$

We bound the derivatives of $h_2(\hat{s}+\tau,\boldsymbol{z}(\hat{s}+\tau))$ in exactly the same way.

Term $\mathcal{E}_3$. Setting $\boldsymbol{v}(\tau) = \boldsymbol{b}_1(\tau)^\mathsf{T}\mathsf{P}(\tau) - \boldsymbol{\Sigma}$, we have $\mathcal{E}_3 = \hat{\boldsymbol{z}}\cdot\int_{\hat{s}}^{\infty}\boldsymbol{v}(\tau)e^{(\tau-\hat{s})\mathsf{A}}\,d\tau$. Using the facts that $\boldsymbol{v}$ is 2-periodic, that its integral vanishes, and that A is negative definite,

$$\int_{\hat{s}}^{\infty}\boldsymbol{v}(\tau)e^{(\tau-\hat{s})\mathsf{A}}\,d\tau = \int_0^{\infty}\boldsymbol{v}(\hat{s}+\tau)e^{\tau\mathsf{A}}\,d\tau = \sum_{k=0}^{\infty}\left(\int_0^2\boldsymbol{v}(\hat{s}+\tau)e^{\tau\mathsf{A}}\,d\tau\right)e^{2k\mathsf{A}}$$

$$= \left(\int_0^2\boldsymbol{v}(\hat{s}+\tau)e^{\tau\mathsf{A}}\,d\tau\right)\left(1-e^{2\mathsf{A}}\right)^{-1}$$

$$= \left(\int_0^2\boldsymbol{v}(\hat{s}+\tau)\left(e^{\tau\mathsf{A}}-1\right)\,d\tau\right)\left(1-e^{2\mathsf{A}}\right)^{-1}.$$

Hence,
$$\frac{d^k}{d\hat{s}^k}\int_{\hat{s}}^{\infty}\boldsymbol{v}(\tau)e^{(\tau-\hat{s})\mathsf{A}}\,d\tau=\left(\int_0^2\boldsymbol{v}^{(k)}(\hat{s}+\tau)\left(e^{\tau\mathsf{A}}-1\right)d\tau\right)\left(1-e^{2\mathsf{A}}\right)^{-1}.$$

Recalling that A is diagonal, we obtain for each value of k the upper bound

$$(\text{A.20})\qquad\left|\frac{d^k}{d\hat{s}^k}\int_{\hat{s}}^{\infty}\left(\boldsymbol{v}(\tau)e^{(\tau-\hat{s})\mathsf{A}}\right)_i\,d\tau\right|=\left\|v_i^{(k)}\right\|_{\infty}\left(\int_0^2\left|e^{\tau\lambda_i}-1\right|d\tau\right)(1-e^{2\lambda_i})^{-1}\leqslant C\left\|\boldsymbol{v}^{(k)}\right\|_{\infty}.$$

Incorporating $(s_0,\boldsymbol{z}_0=\boldsymbol{0})\mapsto(\hat{s},\hat{\boldsymbol{z}})$. We set $\boldsymbol{z}_0=\boldsymbol{0}$ and denote $(\hat{s},\hat{\boldsymbol{z}})=H_k(s_0,\boldsymbol{0})$.

As $\hat{\boldsymbol{z}}=\boldsymbol{z}(\hat{s})=\boldsymbol{z}(\hat{s}(s_0),s_0,\boldsymbol{0})$, we have $\frac{d^k\hat{\boldsymbol{z}}}{ds_0^k}=\frac{d^k}{ds_0^k}\boldsymbol{z}(\hat{s}(s_0),s_0,\boldsymbol{0})$. The bounds

$$\left\|\frac{d^k\hat{\boldsymbol{z}}}{ds_0^k}\right\|\leqslant C\varepsilon\qquad(1\leqslant k\leqslant 3)$$

follow from the fact that $\hat{s}$ is a C^3-function of s_0 and the bounds in (A.2).

Since $\hat{s}$ and $\hat{\boldsymbol{z}}$ are functions of s_0, for any function $u=u(\hat{s},\hat{\boldsymbol{z}})$,

$$\frac{d}{ds_0}u=\frac{d\hat{s}}{ds_0}\partial_{\hat{s}}u+\frac{d\hat{z}_i}{ds_0}\partial_{\hat{z}_i}u$$

$$\frac{d^2}{ds_0^2}u=\frac{d^2\hat{s}}{ds_0^2}\partial_{\hat{s}}u+\frac{d^2\hat{z}_i}{ds_0^2}\partial_{\hat{z}_i}u+\left(\frac{d\hat{s}}{ds_0}\right)^2\partial_{\hat{s}\hat{s}}u+2\frac{d\hat{s}}{ds_0}\frac{d\hat{z}_i}{ds_0}\partial_{\hat{s}\hat{z}_i}u+\frac{d\hat{z}_i}{ds_0}\frac{d\hat{z}_j}{ds_0}\partial_{\hat{z}_i\hat{z}_j}u$$

$$\frac{d^3}{ds_0^3}u=\frac{d^3\hat{s}}{ds_0^3}\partial_{\hat{s}}u+\frac{d^3\hat{z}_i}{ds_0^3}\partial_{\hat{z}_i}u+2\frac{d\hat{s}}{ds_0}\frac{d^2\hat{s}}{ds_0^2}\partial_{\hat{s}\hat{s}}u+2\left(\frac{d^2\hat{s}}{ds_0^2}\frac{d\hat{z}_i}{ds_0}+\frac{d\hat{s}}{ds_0}\frac{d^2\hat{z}_i}{ds_0^2}\right)\partial_{\hat{s}\hat{z}_i}u+2\frac{d^2\hat{z}_i}{ds_0^2}\frac{d\hat{z}_j}{ds_0}\partial_{\hat{z}_i\hat{z}_j}u$$

$$+\frac{d^2\hat{s}}{ds_0^2}\frac{d}{ds_0}\partial_{\hat{s}}u+\frac{d^2\hat{z}_i}{ds_0^2}\frac{d}{ds_0}\partial_{\hat{z}_i}u+\left(\frac{d\hat{s}}{ds_0}\right)^2\frac{d}{ds_0}\partial_{\hat{s}\hat{s}}u+2\frac{d\hat{s}}{ds_0}\frac{d\hat{z}_i}{ds_0}\frac{d}{ds_0}\partial_{\hat{s}\hat{z}_i}u+\frac{d\hat{z}_i}{ds_0}\frac{d\hat{z}_j}{ds_0}\frac{d}{ds_0}\partial_{\hat{z}_i\hat{z}_j}u.$$

Here summation over repeated indices is understood and we leave it to the reader to expand the remaining s_0-derivatives on the last line. Using the bounds derived earlier, we then get

$$\|\boldsymbol{h}_1(\boldsymbol{\zeta})\|\leqslant C\|\hat{\boldsymbol{z}}\|^2e^{\lambda_1\xi}$$

$$\left\|\frac{d}{ds_0}\boldsymbol{h}_1(\boldsymbol{\zeta})\right\|\leqslant C\left(\|\hat{\boldsymbol{z}}\|^2+\varepsilon\|\hat{\boldsymbol{z}}\|\right)e^{\lambda_1\xi}$$

$$\left\|\frac{d^2}{ds_0^2}\boldsymbol{h}_1(\boldsymbol{\zeta})\right\|\leqslant C\left(\|\hat{\boldsymbol{z}}\|^2+\varepsilon\|\hat{\boldsymbol{z}}\|+\varepsilon^2\right)e^{\lambda_1\xi}$$

$$\left\|\frac{d^3}{ds_0^3}\boldsymbol{h}_1(\boldsymbol{\zeta})\right\|\leqslant C\left(\|\hat{\boldsymbol{z}}\|^2+\varepsilon\|\hat{\boldsymbol{z}}\|+\varepsilon^2+\frac{\varepsilon^3}{|\lambda_1|}\right)e^{\lambda_1\xi}.$$

Similar bounds are obtained for $h_2(\boldsymbol{\zeta})$. We conclude that

$$\|\mathcal{E}_1\|_{C^3}\leqslant C\|\boldsymbol{b}_1^{\mathsf{T}}\mathsf{P}\|_{C^3}\frac{\varepsilon^2}{|\lambda_1|^2}\leqslant C\sigma\frac{\varepsilon^2}{|\lambda_1|^2}$$

$$\|\mathcal{E}_2\|_{C^3}\leqslant C\frac{\varepsilon^2}{|\lambda_1|}$$

$$\|\mathcal{E}_3\|_{C^3}\leqslant C\varepsilon\sigma.$$

The final inequality involving $\mathcal{E}_1$ holds because $\frac{\boldsymbol{b}_1^{\mathsf{T}}\mathsf{P}}{\sigma}$ is independent of σ.

Terms $\mathcal{E}_4$ and $\mathcal{E}_5$. Writing $\mathcal{E}_4$ in the form

$$\mathcal{E}_4=(\tilde{s}-\hat{s})\int_0^1 b_0((1-\tau)\tilde{s}+\tau\hat{s})\,d\tau$$

and recalling that $\tilde{s}$ and $\hat{s}$ are both C^3 functions of s_0 allows us to estimate

$$\|\mathcal{E}_4\|_{C^3}\leqslant C\|\tilde{s}-\hat{s}\|_{C^3}\leqslant C\varepsilon.$$

Proposition 3.1 was used here. Finally, by the same proposition,

$$\|\mathcal{E}_5\|_{C^3}=\|\langle\boldsymbol{\xi}_2(s_0,\boldsymbol{0}),\bar{\boldsymbol{d}}\rangle\|_{C^3}\leqslant C\varepsilon\|\boldsymbol{\Sigma}\|,$$

which finishes the proof. ∎

Lemma A.6. *For all s_0 and a, we have $\|\partial_{\boldsymbol{z}_0} s_\infty(s_0, \boldsymbol{0}, a)\| > 0$.*

Proof. Differentiating both sides of (3.11) with respect to $\boldsymbol{z}_0$, we get

$$(A.21) \qquad 0 = b_0(s_\infty)\partial_{\boldsymbol{z}_0} s_\infty + \left\langle \boldsymbol{\Sigma}\left(\int_0^\infty e^{\tau\mathsf{A}}\, d\tau\right), \partial_{\boldsymbol{z}_0}\hat{\boldsymbol{z}}\right\rangle + \boldsymbol{R},$$

where

$$\boldsymbol{R} = -b_0(\hat{s})\partial_{\boldsymbol{z}_0}\hat{s} + \left\langle\left(\int_0^\infty (\boldsymbol{b}_1^\mathsf{T}\mathsf{P} - \boldsymbol{\Sigma})(\hat{s} + \tau)e^{\tau\mathsf{A}}\, d\tau\right), \partial_{\boldsymbol{z}_0}\hat{\boldsymbol{z}}\right\rangle$$
$$+ \left\langle \hat{\boldsymbol{z}}, \left(\int_0^\infty (\boldsymbol{b}_1^\mathsf{T}\mathsf{P})'(\hat{s} + \tau)e^{\tau\mathsf{A}}\, d\tau\right)(\partial_{\boldsymbol{z}_0}\hat{s})\right\rangle + \sum_{k=1}^2 \partial_{\boldsymbol{z}_0}\mathcal{E}_k.$$

Because (A.20) holds for any periodic, zero-integral function, the two integrals appearing in $\boldsymbol{R}$ are $\mathcal{O}(\sigma)$ in the limit $\lambda_1 \to 0$. Terms $\partial_{\boldsymbol{z}_0}\hat{s}$ and $\partial_{\boldsymbol{z}_0}\hat{\boldsymbol{z}}$ are bounded by (A.10) and (A.11), respectively. Estimating $\partial_{\boldsymbol{z}_0}\mathcal{E}_1$ and $\partial_{\boldsymbol{z}_0}\mathcal{E}_2$, we conclude that

$$\|\boldsymbol{R}\| = \mathcal{O}(\sigma) + \mathcal{O}\left(\frac{\sigma\varepsilon}{|\lambda_1|^2}\right).$$

From (A.21) and (A.11), we have

$$(A.22) \qquad \partial_{\boldsymbol{z}_0} s_\infty = \frac{1}{b_0(s_\infty)}\left[\boldsymbol{\Sigma}\mathsf{A}^{-1}(1 + \mathcal{O}(\lambda_1)) + \mathcal{O}(\sigma) + \mathcal{O}\left(\frac{\sigma\varepsilon}{|\lambda_1|^2}\right)\right]$$

as $\lambda_1 \to 0$. Since $\boldsymbol{\Sigma}\mathsf{A}^{-1} = (-\Sigma_i\lambda_i^{-1})_{i=1}^{n-1}$, if $\frac{\varepsilon}{|\lambda_1|}$ is sufficiently small, then the first term on the right side of (A.22) dominates and thus $\|\partial_{\boldsymbol{z}_0} s_\infty\| > 0$. ∎

REFERENCES

1. Michael Benedicks and Lennart Carleson, *On iterations of $1 - ax^2$ on $(-1, 1)$*, Ann. of Math. (2) **122** (1985), no. 1, 1–25. MR MR799250 (87c:58058)
2. _____, *The dynamics of the Hénon map*, Ann. of Math. (2) **133** (1991), no. 1, 73–169. MR MR1087346 (92d:58116)
3. R. E. Lee DeVille, Eric Vanden-Eijnden, and Cyrill B. Muratov, *Two distinct mechanisms of coherence in randomly perturbed dynamical systems*, Phys. Rev. E (3) **72** (2005), no. 3, 031105, 10. MR MR2179903 (2006f:37074)
4. Ian Falconer, Georg A. Gottwald, Ian Melbourne, and Kjetil Wormnes, *Application of the 0-1 test for chaos to experimental data*, SIAM J. Appl. Dyn. Syst. **6** (2007), no. 2, 395–402 (electronic). MR MR2318660 (2008d:37151)
5. Georg A. Gottwald and Ian Melbourne, *A new test for chaos in deterministic systems*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **460** (2004), no. 2042, 603–611. MR MR2034658 (2005d:37168)
6. Brian R. Hunt, Tim Sauer, and James A. Yorke, *Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces*, Bull. Amer. Math. Soc. (N.S.) **27** (1992), no. 2, 217–238. MR MR1161274 (93k:28018)
7. _____, *Prevalence. An addendum to: "Prevalence: a translation-invariant 'almost every' on infinite-dimensional spaces" [Bull. Amer. Math. Soc. (N.S.) **27** (1992), no. 2, 217–238; MR1161274 (93k:28018)]*, Bull. Amer. Math. Soc. (N.S.) **28** (1993), no. 2, 306–307. MR MR1191479 (93k:28019)
8. M. V. Jakobson, *Absolutely continuous invariant measures for one-parameter families of one-dimensional maps*, Comm. Math. Phys. **81** (1981), no. 1, 39–88. MR MR630331 (83j:58070)
9. Kevin K. Lin and Lai-Sang Young, *Shear-induced chaos*, Nonlinearity **21** (2008), no. 5, 899–922. MR MR2412320 (2009f:37034)
10. William Ott and James A. Yorke, *Prevalence*, Bull. Amer. Math. Soc. (N.S.) **42** (2005), no. 3, 263–290 (electronic). MR MR2149086 (2006d:28012)
11. Ph. Thieullen, C. Tresser, and L.-S. Young, *Positive Lyapunov exponent for generic one-parameter families of unimodal maps*, J. Anal. Math. **64** (1994), 121–172. MR MR1303510 (95g:58134)
12. Philippe Thieullen, Charles Tresser, and Lai-Sang Young, *Exposant de Lyapunov positif dans des familles à un paramètre d'applications unimodales*, C. R. Acad. Sci. Paris Sér. I Math. **315** (1992), no. 1, 69–72. MR MR1172409 (93m:58067)
13. Warwick Tucker, *A rigorous ODE solver and Smale's 14th problem*, Found. Comput. Math. **2** (2002), no. 1, 53–117. MR MR1870856 (2003b:37055)
14. Qiudong Wang and Lai-Sang Young, in preparation.
15. _____, *Strange attractors with one direction of instability*, Comm. Math. Phys. **218** (2001), no. 1, 1–97. MR MR1824198 (2002m:37050)
16. _____, *From invariant curves to strange attractors*, Comm. Math. Phys. **225** (2002), no. 2, 275–304. MR MR1889226 (2003e:37045)

17. _____, *Strange attractors in periodically-kicked limit cycles and Hopf bifurcations*, Comm. Math. Phys. **240** (2003), no. 3, 509–529. MR MR2005855 (2004i:37070)
18. _____, *Toward a theory of rank one attractors*, Ann. of Math. (2) **167** (2008), no. 2, 349–480. MR MR2415378
19. Lai-Sang Young, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. (2) **147** (1998), no. 3, 585–650. MR MR1637655 (99h:58140)
20. _____, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188. MR MR1750438 (2001j:37062)
21. G. M. Zaslavsky, *The simplest case of a strange attractor*, Phys. Lett. A **69** (1978/79), no. 3, 145–147. MR MR587772 (82g:58062)

COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK, NY 10012
URL, William Ott: `http://www.cims.nyu.edu/~ott`

URL, Mikko Stenlund: `http://www.math.helsinki.fi/mathphys/mikko.html`