ON THE SPECTRUM OF RANDOMLY PERTURBED EXPANDING MAPS

V. Baladi and L.-S. Young

August 1992

Abstract. We consider small random perturbations of expanding and piecewise expanding maps and prove the robustness of their invariant densities and rates of mixing. We do this by proving some simple lemmas about the robustness of the spectra of certain operators. These abstract results are then applied to the Perron-Frobenius operators of the models in question.

Introduction

Let $f : M \to M$ be a dynamical system preserving some natural probability measure μ_0 with density ρ_0. This paper is motivated by the following question: *does exponential mixing imply stochastic stability*? Roughly speaking, *exponential mixing* of (f, μ_0) means that, for two observables φ and ψ on M, the correlation between $\varphi \circ f^n$ and ψ decays exponentially fast with n. *Stochastic stability* means that, if we add a small amount of random noise to f, obtaining at noise level ϵ a Markov process with invariant density ρ_ϵ, then ρ_ϵ tends to ρ_0 as ϵ tends to zero.

The following heuristic argument suggests an affirmative answer to this question. Consider the Perron-Frobenius operator $\mathcal{L}$ associated with f, acting on a suitable class of functions. The exponential mixing property is equivalent to the presence of a gap in the spectrum of $\mathcal{L}$ between the eigenvalue equal to unity and the "next largest eigenvalue." Corresponding to the noisy situation is a noisy Perron-Frobenius operator $\mathcal{L}_\epsilon$, which should not be too different from $\mathcal{L}$, for small ϵ. By the usual geometric arguments for

---

1991 *Mathematics Subject Classification.* 58F11 58F30; 58C25 58F03 58F15 58F19 58G32 60J10.

V. Baladi started the present work during a postdoctoral fellowship at IBM, T.J. Watson Center.
L.-S. Young is partially supported by the National Science Foundation.

hyperbolic operators, the eigenfunction corresponding to the eigenvalue 1 for $\mathcal{L}_\epsilon$ should be near that for $\mathcal{L}$, proving stochastic stability.

Also, since the "second largest" eigenvalue of $\mathcal{L}$ determines the rate of decay of correlations, if there is a gap between the "second largest" and the "third largest" eigenvalue, then a similar reasoning will show that the presence of small amounts of noise should not affect significantly the rate of mixing of the system. When further gaps exist, this reasoning can be extended to other eigenvalues of $\mathcal{L}$ (the "resonances" of Ruelle [1986]).

In this paper, we examine three models against the ideas sketched in the last two paragraphs. These models are (1) expanding maps of the circle perturbed by convolution and (2) expanding maps of Riemannian manifolds followed by diffusions, both with the space of $\mathcal{C}^r$ test functions; and (3) piecewise expanding maps of the interval, with test functions of bounded variation. All of these models, when unperturbed, have the exponential mixing property. That is known and will be taken for granted here. In all three cases we successfully carry out the steps sketched above, proving stochastic stability and robustness of the rates of mixing. (For (3), some additional conditions on the gaps of the spectrum are needed.) We shall see, however, that $\mathcal{L}_\epsilon$ does not converge to $\mathcal{L}$ in the operator norm topology. Indeed, the manner in which $\mathcal{L}_\epsilon$ approaches $\mathcal{L}$ in each case is delicate and depends on the dynamics as well as the function spaces in question.

Some of our results are new; others are not. We will state them precisely and give references when appropriate in subsequent sections. We wish to emphasize here our relatively unified method of proof: we first prove some simple perturbation lemmas for abstract operators that apply simultaneously to all three models. Once that is done, we prove some dynamical lemmas relating $\mathcal{L}_\epsilon$ to $\mathcal{L}$ for each model. We hope that this approach goes beyond the situations considered in the present article.

We will keep the setting of this paper simple and the amount of machinery to a minimum. In a forthcoming paper by the first named author some of the results here will be brought to fuller generality. Transfer operators with more general weights will be considered, and the Fredholm determinants of the perturbed operators will be shown to converge to that of $\mathcal{L}$ on certain regions of the complex plane. (The study of this last problem was suggested by D. Ruelle.)

We express thanks to Pierre Collet, François Ledrappier, and Bruno Sevennec for very useful conversations. V. Baladi acknowledges the hospitality and financial support of the Mathematics Department at U.C.L.A., the I.H.E.S., and the Niels Bohr Institute. L.-S. Young is grateful to the Mittag-Leffler Institute for its hospitality and support.

1. Background, definitions and notations

Let $f : M \to M$ be a differentiable or piecewise differentiable transformation of a compact Riemannian manifold. Assume that f preserves a Borel probability measure μ_0 of the form $\mu_0 = \rho_0 \, dm$, where m denotes Riemannian volume. Our aim in this work is to study the invariant density and rate of mixing of (f, μ_0) under small random perturbations, and we do that by studying the spectral properties of the perturbed Perron-Frobenius operators associated with f. The purpose of this section is to give precise definitions for all of these terms.

Let $\mathcal{B}$ denote the σ-algebra of Borel sets of M and $\mathcal{P}$ the space of Borel probability measures on M. Recall that a random perturbation of f is a family of Markov chains $\mathcal{X}^\epsilon$ (with small $\epsilon > 0$) defined on the measure space $(M, \mathcal{B})$, with transition probabilities $\{P^\epsilon(x, \cdot)\}$ in $\mathcal{P}$ (i.e., $P\{\mathcal{X}^\epsilon_{n+1} \in E \, : \, \mathcal{X}^\epsilon_n = x\} = P^\epsilon(x, E)$). We assume that the following conditions are satisfied:

(1) The map $x \mapsto P^\epsilon(x, \cdot)$ is continuous for each ϵ.

(2) Each $P^\epsilon(x, \cdot)$ is absolutely continuous with respect to Lebesgue measure m.

(3) For any continuous test function $g : M \to \mathbb{R}$

$$\lim_{\epsilon \to 0} \left(\sup_{x \in M} \left| \int_M g(y) P^\epsilon(x, dy) - g(fx) \right| \right) = 0 \, .$$

If M is compact, it follows from (1) and (2) that each Markov chain $\mathcal{X}^\epsilon$ admits an absolutely continuous invariant probability measure μ_ϵ, i.e., a probability measure $\mu_\epsilon = \rho_\epsilon dm$ such that

$$\mu_\epsilon(E) = \int P^\epsilon(x, E) d\mu_\epsilon(x)\,, \forall E \in \mathcal{B}\,.$$

Moreover, one can often show that μ_ϵ is unique if f has some transitivity properties. (For more details, see e.g. Kifer [1988a]. See also Benedicks–Young [1992])

We say that (f, μ_0) is *stochastically stable* under the perturbation $\mathcal{X}^\epsilon$ if μ_ϵ tends to μ_0 weakly as $\epsilon \to 0$. Various dynamical systems have been shown to be stochastically stable in this sense (see e.g. Kifer [1974] and the results and references in [1988a], Benedicks-Young [1992] etc.). Sometimes, one has a stronger notion of stochastic stability. If $(\mathcal{F}, \|\cdot\|)$ is a Banach space of functions $\rho : M \to \mathbb{R}$ containing ρ_0 and ρ_ϵ, then we say that (f, μ_0) is *stochastically stable in* $(\mathcal{F}, \|\cdot\|)$ if $\|\rho_\epsilon - \rho_0\|$ tends to zero as $\epsilon \to 0$. (See e.g. Keller [1982] and Collet [1984] for certain interval maps, with $\mathcal{F} = L^1(dm)$.)

We are also going to consider the convergence of the rate of mixing. Recall that one says that τ_0 is the *rate of decay of correlations of* (f, μ_0) *for functions in* $(\mathcal{F}, \|\cdot\|)$ if τ_0 is the smallest number such that the following holds: for each $\tau > \tau_0$ and each pair $\varphi, \psi \in \mathcal{F}$, there exists $C = C(\tau, \|\varphi\|, \|\psi\|)$ such that

$$|\int (\varphi \circ f^n) \cdot \psi d\mu_0 - \int \varphi d\mu_0 \int \psi d\mu_0| \le C\tau^n\,, \quad \forall n \ge 1\,.$$

We are mostly interested in the case where $\tau_0 < 1$.

Consider now the Markov chain $(\mathcal{X}^\epsilon, \mu_\epsilon)$, and let $P_n^\epsilon(x, \cdot)$ be the n-step transition probability. We say that τ_ϵ is the *rate of decay of correlations of* $(\mathcal{X}^\epsilon, \mu_\epsilon)$ *for functions in* $(\mathcal{F}, \|\cdot\|)$ if τ_ϵ is the smallest number such that the following holds: for each $\tau > \tau_\epsilon$ and each pair $\varphi, \psi \in \mathcal{F}$, there exists $C = C(\tau, \|\varphi\|, \|\psi\|)$ such that

$$|\int (\int \varphi(y) P_n^\epsilon(x, dy)) \cdot \psi(x) d\mu_\epsilon(x) - \int \varphi d\mu_\epsilon \int \psi d\mu_\epsilon| \le C\tau^n\,, \quad \forall n \ge 1\,.$$

We say that the rate of mixing of (f, μ_0) in $\mathcal{F}$ is *robust* if τ_ϵ tends to τ_0 as ϵ goes to zero. (The relation between τ_ϵ and τ_0 has been considered in e.g. Ruelle [1986], for mixing Anosov flows.)

Next we define the Perron-Frobenius operator associated with f. For this, we fix a suitable Banach space of functions $(\mathcal{F}, \|\cdot\|)$ as above, and for $\varphi \in \mathcal{F}$, we define

$$\mathcal{L}\varphi(x) = \sum_{f(y)=x} \frac{\varphi(y)}{|\det Df_y|}\,.$$

Or, equivalently, if $\varphi \in \mathcal{F}$ is the density of a signed measure μ on M, then $\mathcal{L}\varphi$ is the density of $f_*\mu$ where $f_*\mu$ is the push-forward of μ by f, i.e., $(f_*\mu)(E) = \mu(f^{-1}E)$, for all $E \in \mathcal{B}$. We assume that $\mathcal{L} : \mathcal{F} \to \mathcal{F}$ is a well-defined bounded operator, and that $\rho_0 \in \mathcal{F}$. Then 1 is an eigenvalue of $\mathcal{L}$, and our invariant density ρ_0 is an eigenfunction for the eigenvalue 1.

In our models, as in virtually all situations where the spectrum of the Perron-Frobenius operator is understood, the operator $\mathcal{L}$ is quasi-compact, i.e., its essential spectral radius $\text{ ess sp }(\mathcal{L})$ is strictly less than its spectral radius. In particular, for every $\tau > \text{ ess sp }(\mathcal{L})$, the set $\sigma(\mathcal{L}) \cap \{z \,:\, |z| \geq \tau\}$ consists of a finite number of eigenvalues with finite dimensional eigenspaces. If we further assume that (f, μ_0) is exact — which is the case for the models considered in this paper — then the spectrum of $\mathcal{L}$ can be written as $\sigma(\mathcal{L}) = \sigma_0 \cup \{1\}$, where 1 is a simple eigenvalue (this means that its algebraic multiplicity is equal to one, see e.g. Kato [1976] for terminology) and $|\sigma_0| := \sup\{|z| \,:\, z \in \sigma_0\} < 1$ (see Hofbauer–Keller [1982], Ruelle [1989]).

The relationship between τ_0 and σ_0 is as follows: since

$$\int (\varphi \circ f^n) \cdot \psi \, d\mu_0 = \int \varphi \cdot \mathcal{L}^n(\psi \, \rho_0) \, dm\,,$$

we have

$$|\int (\varphi \circ f^n)\, \psi \, d\mu_0 - \int \varphi \, d\mu_0 \int \psi \, d\mu_0| = |\int \varphi \big[\mathcal{L}^n(\psi \, \rho_0) - \big(\int \psi \rho_0 \, dm\big)\, \rho_0\big] \, dm\,|\,.$$

If $\int |\varphi|\, dm \leq \text{const} \cdot \|\varphi\|$ — and this is certainly true in our models — the last expression above is

$$
\begin{aligned}
&\leq C \cdot \|\mathcal{L}^n(\psi\rho_0) - \pi(\psi\rho_0)\| \\
&\leq C' \cdot \tau^n ,
\end{aligned}
$$

where τ is any number strictly larger than $|\sigma_0|$, the constants C and C' depend only on $\|\varphi\|$, $\|\psi\|$ and τ, and π is the projection onto the one-dimensional eigenspace of 1. Thus we have $\tau_0 = |\sigma_0|$.

If $|\sigma_0| >$ ess sp $(\mathcal{L})$, then $\tau_0 = |\sigma_0|$ will be referred to as an *isolated* rate of decay.

Corresponding to the perturbation $\mathcal{X}^\epsilon$ of f, we define the Perron-Frobenius operator $\mathcal{L}_\epsilon$ as follows: if $\varphi \in \mathcal{F}$ is the density of μ, then $\mathcal{L}_\epsilon\varphi$ is the density of $\mathcal{X}^\epsilon_*\mu$ where $\mathcal{X}^\epsilon_*\mu(E) = \int P^\epsilon(x, E) d\mu(x)$. Moreover, if $\rho_\epsilon \in \mathcal{F}$ and if 1 is the only point of $\sigma(\mathcal{L}_\epsilon)$ on the unit circle, it is a simple eigenvalue; we then write $\sigma(\mathcal{L}_\epsilon) = \{1\} \cup \sigma_0(\mathcal{L}_\epsilon)$ and the interpretation of τ_ϵ as $|\sigma_0(\mathcal{L}_\epsilon)|$ carries over as well.

In the next three sections, we will consider for each of our models the following questions:

(1) does $\|\rho_\epsilon - \rho_0\| \to 0$?

(2) does $\tau_\epsilon \to \tau_0$ (assuming that τ_0 is an isolated rate of decay)?

If the answers to (1) and (2) are affirmative then we may also ask

(3) how does $\|\rho_\epsilon - \rho_0\|$ or $|\tau_\epsilon - \tau_0|$ scale with ϵ as $\epsilon \to 0$?

2. Perturbation lemmas for abstract operators

It will become clear in the next three sections that the setting we have to deal with is the following: let $(X, \|\cdot\|)$ be a complex Banach space, and let $\{T_\epsilon, \epsilon \geq 0\}$ be a family of bounded linear operators on X. We make the following assumption about T_0:

There exist two real numbers $0 < \kappa_1 < \kappa_0 \leq 1$ such that the spectrum of T_0 decomposes as $\Sigma_0 \cup \Sigma_1$ where

$$
\begin{aligned}
\kappa_0 &= \inf\{|z| \,:\, z \in \Sigma_0\} \\
\kappa_1 &= \sup\{|z| \,:\, z \in \Sigma_1\} .
\end{aligned}
\tag{A.1}
$$

Let X_i be the eigenspace corresponding to Σ_i, and let $\pi_i : X_0 \oplus X_1 \to X_i$ be the associated projection. Let $\sigma(\cdot)$ denote the spectrum of an operator. Our first result is

Lemma 1. *Assume that there exists $\kappa_1 < \kappa < \kappa_0$ such that for each sufficiently large $n \in \mathbb{Z}^+$, there exists $\epsilon(n)$ such that for all $0 < \epsilon < \epsilon(n)$*

$$\|T_\epsilon^n - T_0^n\| \le \kappa^n \,. \tag{A.2}$$

Then, for each sufficiently small $0 < \epsilon$, there exists a decomposition of $\sigma(T_\epsilon)$ into

$$\sigma(T_\epsilon) = \Sigma_0^\epsilon \cup \Sigma_1^\epsilon$$

such that

$$\kappa' := \sup\{|z| \,:\, z \in \Sigma_1^\epsilon\} < \kappa_0' := \inf\{|z| \,:\, z \in \Sigma_0^\epsilon\} \,,$$

where κ' and κ_0' can be made arbitrarily close to κ and κ_0 by choosing ϵ small enough.

Note that we do not assume that T_ϵ^n converges to T_0^n as $\epsilon \to 0$ for fixed n, not even pointwise.

Proof of Lemma 1. Fix κ_1', κ', κ_0' near κ_1, κ, κ_0 such that

$$\kappa_1 < \kappa_1' < \kappa < \kappa' < \kappa_0' < \kappa_0 \,.$$

Let N be large enough for all the purposes below, in particular, we require that

$$\begin{aligned} x \in X_0 &\Longrightarrow \|T_0^N x\| \ge (\kappa_0')^N \|x\| \\ x \in X_1 &\Longrightarrow \|T_0^N x\| \le (\kappa_1')^N \|x\| \,. \end{aligned}$$

Let $\epsilon < \epsilon(N)$, and let λ satisfy $\kappa' < |\lambda| < \kappa_0'$. We will show that $\lambda \notin \sigma(T_\epsilon)$.

It suffices to prove that the resolvent $R(T_\epsilon^N, \lambda^N)$ exists as a bounded operator. We write down what it must be if it exists:

$$\begin{aligned} R(T_\epsilon^N, \lambda^N) &= \left[(\lambda^N I - T_0^N) - (T_\epsilon^N - T_0^N)\right]^{-1} \\ &= \left[(\lambda^N I - T_0^N) \cdot \big(I - R(T_0^N, \lambda^N)(T_\epsilon^N - T_0^N)\big)\right]^{-1} \\ &= \sum_{n=0}^{\infty} \big(R(T_0^N, \lambda^N)(T_\epsilon^N - T_0^N)\big)^n \cdot R(T_0^N, \lambda^N) \,. \end{aligned} \tag{2.1}$$

Assuming $\|T_\epsilon^N - T_0^N\| < \kappa^N$, it is enough to show $\|R(T_0^N, \lambda^N)\| < (1/\kappa)^N$. Since $R(T_0^N, \lambda^N)X_i = X_i$ for $i = 0, 1$, we have for $x \in X$, $\|x\| = 1$

$$\begin{aligned}\|R(T_0^N, \lambda^N)\| &\le \|R(T_0^N, \lambda^N)\pi_0 x\| + \|R(T_0^N, \lambda^N)\pi_1 x\| \\ &\le \|R(T_0^N, \lambda^N)|_{X_0}\|\|\pi_0\| + \|R(T_0^N, \lambda^N)|_{X_1}\|\|\pi_1\|\,.\end{aligned}$$

so that it suffices to bound $\|R(T_0^N, \lambda^N)|_{X_i}\|$, $i = 0, 1$. If $x \in X_0$, for $\kappa_0' < \hat{\kappa}_0 < \kappa_0$,

$$\begin{aligned}\|T_0^N X - \lambda^N x\| &\ge \|T_0^N x\| - |\lambda|^N \|x\| \\ &\ge \big((\hat{\kappa}_0)^N - (\kappa_0')^N\big)\|x\| \\ &\ge \text{const} \cdot (\hat{\kappa}_0)^N \|x\|\,,\end{aligned}$$

and if $x \in X_1$,

$$\begin{aligned}\|T_0^N X - \lambda^N x\| &\ge -\|T_0^N x\| + |\lambda|^N \|x\| \\ &\ge \big(-(\kappa_1')^N - (\kappa')^N\big)\|x\| \\ &\ge \text{const} \cdot (\kappa')^N \|x\|\,.\end{aligned}$$

Hence, for large enough N,

$$\|R(T_0^N, \lambda^N)\| \le \frac{\text{const} \cdot (\|\pi_0\| + \|\pi_1\|)}{(\kappa')^N} \le \frac{1}{\kappa^N}\,. \tag{2.2}$$

Define

$$\Sigma_0^\epsilon := \{z \in \sigma(T_\epsilon) \,:\, |z| \ge \kappa_0'\} \qquad \Sigma_1^\epsilon := \{z \in \sigma(T_\epsilon) \,:\, |z| \le \kappa'\}\,. \quad \square$$

Let $\pi_0^\epsilon : X_0^\epsilon \oplus X_1^\epsilon \to X_0^\epsilon$ be the projection associated with the spectral decomposition of T_ϵ. For $\Gamma \subset \mathbb{C}$ write $\Gamma^N := \{z^N \,:\, z \in \Gamma\}$. We also use the notation $B_r := \{|z| = r\}$.

Lemma 2. *If Assumptions* (A.1) *and* (A.2) *hold then* $\|\pi_0 - \pi_0^\epsilon\| \to 0$ *as* $\epsilon \to 0$.

Proof of Lemma 2. Note that π_0 can be regarded as the projection associated with $(T^N, (\Sigma_0)^N)$ for any N, and similarly for π_0^ϵ. We will again consider N large and $\epsilon < \epsilon(N)$.

Let $C := B_{\hat\kappa^N} \cup B_{r_0^N}$ for some $\kappa' < \hat\kappa < \kappa'_0$ with $\hat\kappa < (\kappa')^2/\kappa$, and $r_0 > |\sigma(T_0)|$. Then Σ_0^N and $(\Sigma_0^\epsilon)^N$ are contained in the annular region bounded by C, and we have

$$\pi_0 = \frac{1}{2i\pi}\int_C R(T_0^N,\lambda)\,d\lambda \qquad\qquad \pi_0^\epsilon = \frac{1}{2i\pi}\int_C R(T_\epsilon^N,\lambda)\,d\lambda\,.$$

We will estimate $\|\pi_0 - \pi_0^\epsilon\|$ by

$$\begin{aligned}\|\pi_0 - \pi_0^\epsilon\| &\le \frac{1}{2\pi}\int_C \|R(T_0^N,\lambda) - R(T_\epsilon^N,\lambda)\|\,d\lambda \\ &\le \frac{1}{2\pi}\cdot \ell(B_{\hat\kappa^N})\cdot \max_{\lambda\in B_{\hat\kappa^N}} \|R(T_0^N,\lambda) - R(T_\epsilon^N,\lambda)\| \\ &\qquad + \text{ the corresponding term for } B_{r_0^N} \\ &=: (1) + (2)\,.\end{aligned} \tag{2.3}$$

Using (2.1) we have

$$\|R(T_0^N,\lambda) - R(T_\epsilon^N,\lambda)\| \le \sum_{n=1}^\infty \|R(T_0^N,\lambda)\|^{n+1}\cdot \|T_\epsilon^N - T_0^N\|^n\,.$$

Since $\ell(B_{\hat\kappa^N}) = 2\pi\hat\kappa^N$, and $\|R(T_0^N,\lambda)\| \le \text{const}/(\kappa')^N$ for $\lambda \in B_{\hat\kappa^N}$ (by (2.2)), we obtain

$$\begin{aligned}(1) &\le \hat\kappa^N \cdot \sum_{n=1}^\infty \left(\frac{\text{const}}{{\kappa'}^N}\right)^{n+1} (\kappa^N)^n \\ &\le \text{const}\cdot \hat\kappa^N \cdot \frac{\kappa^N}{({\kappa'}^N)^2} \to 0 \text{ as } N\to\infty\,.\end{aligned}$$

For (2), we use $\ell(B_{r_0{}^N}) = 2\pi {r_0}^N$, to get

$$(2) \le \text{const}\cdot r_0^N \cdot \frac{\kappa^N}{r_0^{2N}} \to 0 \text{ as } N\to\infty\,. \quad \square$$

For $n \ge 1$ define

$$C_n(\epsilon) := \sup_{\substack{x\in X_0\\ x\neq 0}} \frac{\|T_\epsilon^n x - T_0^n x\|}{\|x\|}\,.$$

(By (A.1), $C_n(\epsilon) < \kappa^n$ for large enough n and small enough ϵ.)

Lemma 3. *Assume that* (A.1)-(A.2) *hold, that* $\|T_\epsilon\|$ *is uniformly bounded, and that*

$$dim\ X_0 < \infty\,. \tag{A.3}$$

Let d denote the maximum algebraic multiplicity of $z \in \sigma(T_0|_{X_0})$ and let κ' and $\kappa_0' < \kappa_0$ be given from Lemma 1. Then for fixed large N and $\epsilon < \epsilon(N)$:

(1) *Hausdorff-distance*$(\sigma(T_0|_{X_0}), \sigma(T_\epsilon|_{X_0^\epsilon})) \leq$ *const* $\cdot\, (C_1(\epsilon) + \frac{C_N(\epsilon)}{{\kappa_0'}^N})^{1/d}$.

(2) *If $\hat{x}_0 \in X_0$ is an eigenvector for T_0 with $T_0\hat{x}_0 = \nu_0\hat{x}_0$, then T_ϵ has an eigenvector $\hat{x}_0^\epsilon \in X_0^\epsilon$ with eigenvalue ν_0^ϵ which is const $\cdot\, (C_1(\epsilon) + \frac{C_N(\epsilon)}{{\kappa_0'}^N})^{1/d}$-near ν_0 such that*

$$\|\hat{x}_0^\epsilon - \hat{x}_0\| \leq const \cdot (C_1(\epsilon) + \frac{C_N(\epsilon)}{{\kappa_0'}^N})^{1/d}\,.$$

The assumption that $\|T_\epsilon\|$ is uniformly bounded is not essential since for some large iterate $\|T_\epsilon^N\| \leq \|T_0^N\| + \kappa^N$ for all small enough ϵ.

Proof of Lemma 3. First we show that $X_0^\epsilon = \mathrm{graph}(S_\epsilon)$ for some linear $S_\epsilon : X_0 \to X_1$ with $\|S_\epsilon\| \to 0$ as $\epsilon \to 0$. To see this, consider ϵ small and let $x \in X_0^\epsilon$. Since $\|x - \pi_0 x\| \leq \|\pi_0^\epsilon - \pi_0\|\|x\|$, it follows that if $x = (x_0, x_1) \in X_0 \oplus X_1$, then $\|x_1\| \ll \|x_0\|$. This inequality implies in particular that if x, $x' \in X_0^\epsilon$ and $\pi_0 x = \pi_0 x'$ then $x = x'$.

Next, we estimate $\|S_\epsilon\|$. We know by (A.3) that there exists $x_0 \in X_0$, $\|x_0\| = 1$ such that

$$\|S_\epsilon\| \leq \frac{\|\pi_1 T_\epsilon^N(x_0, S_\epsilon x_0)\|}{\|\pi_0 T_\epsilon^N(x_0, S_\epsilon x_0)\|}\,.$$

This is

$$\leq \frac{\|\pi_1\|\Big(\big((\kappa_1')^N + \kappa^N\big)\|S_\epsilon\| + C_N(\epsilon)\Big)}{(\kappa_0')^N - \|\pi_0\|(1 + \|S_\epsilon\|)\cdot\kappa^N}\,, \tag{2.4}$$

from which we see that

$$\|S_\epsilon\| \leq \mathrm{const}\frac{C_N(\epsilon)}{(\kappa_0')^N}\,.$$

Define $\hat{T}_\epsilon : X_0 \to X_0$ by

$$\hat{T}_\epsilon(x) = \pi_0 \circ T_\epsilon(x, S_\epsilon x)\,.$$

Then for $x \in X_0$ with $\|x\| = 1$, we have

$$\begin{aligned}\|\hat{T}_\epsilon x - T_0 x\| &\le \|\pi_0\| \cdot (\|T_\epsilon x - T_0 x\| + \|T_\epsilon S_\epsilon x\|) \\ &\le \text{const} \cdot (C_1(\epsilon) + \|T_\epsilon\| \cdot \frac{C_N(\epsilon)}{{\kappa'_0}^N}) \,.\end{aligned}$$

There is a similar bound for $\|\pi_1 \circ T_\epsilon(x, S_\epsilon x) - \pi_1 T_0 x\|$ with $x \in X_0$. The assertions of Lemma 3 follow immediately. (See e.g. Wilkinson [1965].) $\square$

3. The simplest model:

expanding maps of the circle and perturbations by convolutions

A. The unperturbed model.

Assume first that our manifold M is equal to the circle S^1. Let f be a $\mathcal{C}^r$ transformation of S^1 ($2 \le r < \infty$) which is expanding, i.e., $|f'| \ge \lambda > 1$. The *expanding constant* of f is the largest λ such that this inequality holds. This implies the existence of a unique absolutely continuous invariant probability measure μ_0 with respect to which f is mixing (in fact, exact).

We set $\mathcal{F} = \mathcal{C}^{r-1}(S^1)$ and let $\|\cdot\|$ be the usual $\mathcal{C}^{r-1}$-norm. Let $\mathcal{L} : \mathcal{F} \to \mathcal{F}$ be the Perron–Frobenius operator associated with f:

$$\mathcal{L}\varphi(x) = \sum_{f(y)=x} \frac{\varphi(y)}{|f'(y)|} \,.$$

It is proved in Ruelle [1989] (see also Collet–Isola [1991]) that $\mathcal{L}$ is quasi-compact with essential spectral radius bounded above by $(1/\lambda)^{r-1}$.

We remark that if the map f is $\mathcal{C}^\infty$ or $\mathcal{C}^\omega$, we can let $\mathcal{L}$ act on the Fréchet space $\mathcal{C}^\infty(S^1)$ of $\mathcal{C}^\infty$ functions, respectively the Banach space $\mathcal{C}^\omega(S^1)$ of real analytic functions endowed with the supremum norm. Using the fact (Ruelle [1989]) that, for a $\mathcal{C}^r$ map, the eigenfunctions of $\mathcal{L}$ acting on $\mathcal{C}^{r'}$ for $1 \le r' < r-1$ are all elements of $\mathcal{C}^{r-1}(S^1)$, it makes sense to speak of the eigenvalues of $\mathcal{L}$ when acting on $\mathcal{C}^\infty(S^1)$, even though $\mathcal{C}^\infty(S^1)$ is not a Banach space. In particular, one can view $\mathcal{L} : \mathcal{C}^\infty(S^1) \to \mathcal{C}^\infty(S^1)$ as a "compact" operator. If $r = \omega$, the operator $\mathcal{L}$ is (truly) compact, and much is known about it (Ruelle [1976], Mayer [1976], etc.). We will not discuss further the cases $r = \infty, \omega$, but our results clearly hold there too.

We remark also that $\tau_0 = |\sigma_0|$ is not always an isolated rate of decay. Consider for instance the map $z \to z^2$ on S^1 and its the transfer operator acting on real analytic functions. By following the computation in Ruelle [1986], one checks that the relevant Fredholm determinant is equal to $(1 - z)$, so that the only eigenvalue is 1. This implies (Ruelle [1976,1989,1990]) that the transfer operator acting on $\mathcal{C}^r(S^1)$, with $1 \le r \le \infty$ has no eigenvalue besides 1 whose modulus is bigger than the essential spectral radius. The other "algebraic" maps $z \mapsto z^k$, for integers $k \ge 3$, have the same property. However, as pointed out to us by Mark Pollicott, the above examples do not seem to be generic: a necessary condition for the lack of nontrivial eigenvalues in the spectrum of the operator acting on analytic functions is the fact that the trace of the Fredholm operator is equal to 1. By considering analytic perturbations of the algebraic examples, one can arrange that the value of this trace changes. For example, the projection on the circle of the periodic map $x \mapsto 2x (\mathrm{mod}\, 1) + \delta \sin 2\pi x$ only has one fixed point (if $\delta > 0$ is not too large), and the trace of its Perron-Frobenius operator can easily be computed to be $1/(1 - \delta) > 1$, so that there is at least one eigenvalue besides 1 whose real part is strictly positive (Pollicott [1991]).

B. Type of perturbation: convolutions.

For $\epsilon > 0$, let $\theta_\epsilon : \mathbb{R} \to \mathbb{R}$ be a function in $L^1(dm)$ satisfying

$$\theta_\epsilon \ge 0\,,\ \mathrm{supp}\,\theta_\epsilon \subset [-\epsilon, \epsilon]\,,\ \text{and} \int \theta_\epsilon = 1\,.$$

Consider the random perturbation $\mathcal{X}^\epsilon$ where the transition probabilities $P^\epsilon(x, dy)$ have densities $\theta_\epsilon(y - fx)$. (I.e., the density only depends on the difference $y - fx$.) Equivalently (use Fubini), one can describe this process as given by f followed by a random translation by ω, where ω is distributed according to θ_ϵ. We call such a perturbation a *random perturbation by convolution* (see Kifer [1988a, Chapter IV]).

Since f is topologically transitive the invariant probability measure is unique and its density ρ_ϵ is positive on each open subset of S^1.

The perturbed Perron–Frobenius operator $\mathcal{L}_\epsilon : \mathcal{C}^{r-1}(S^1) \to \mathcal{C}^{r-1}(S^1)$ can be written as follows: for $\varphi \in \mathcal{C}^{r-1}(S^1)$,

$$\begin{aligned}(\mathcal{L}_\epsilon\varphi)(x) &= \int (\mathcal{L}\varphi)(x-\omega)\theta_\epsilon(\omega)\, d\omega \\ &= \int \varphi(y)\theta_\epsilon(x - fy) dm(y)\,.\end{aligned}$$

Analogous operators have been used by Keller [1982, §5] and Collet [1984] among others. It is clearly linear and bounded on $\mathcal{C}^{r-1}(S^1)$ and enters the class studied by Ruelle [1990]. In particular, it is quasi-compact and the density ρ_ϵ is hence in $\mathcal{C}^{r-1}$.

If we had made the additional asssumption that θ_ϵ is $\mathcal{C}^{r-1}$ then, by using a simple modification of the usual Ascoli argument to show that a kernel operator

$$\varphi(x) \to \int_{S^1} K(x,y)\varphi(y)\, dm(y)\,,\ \ \varphi \in \mathcal{C}^0(S^1)\,,$$

with $\mathcal{C}^0$ kernel $K(\cdot,\cdot)$, is compact (see e.g. Yosida [1980, p. 277]), we could easily show that $\mathcal{L}_\epsilon$ is compact on $\mathcal{C}^{r-1}(S^1)$.

C. Statement of our results.

We now state our main results, which give partial answers to the questions posed in Section 1 for this simplest model:

Theorem 1. *Let $f : S^1 \to S^1$ be a $\mathcal{C}^r$ expanding map ($r \geq 2$) of the circle as defined in Section 3.A, with expanding constant λ, and let $\mu_0 = \rho_0\, dm$ be its unique absolutely continuous invariant probability measure. Let $\mathcal{X}^\epsilon$ be a small random perturbation of f of the type described in Section 3.B, with invariant measure $\mu_\epsilon = \rho_\epsilon dm$. Then:*

(1) *The dynamical system (f, μ_0) is stochastically stable under $\mathcal{X}^\epsilon$ in the space of $\mathcal{C}^{r-1}$ functions, i.e., $\|\rho_\epsilon - \rho_0\|_{r-1}$ tends to 0 as $\epsilon \to 0$. Moreover, we have $\|\rho_\epsilon - \rho_0\|_{r-2} = O(\epsilon)$.*

Let τ_0 and τ_ϵ be the rates of decay of correlation functions for f and $\mathcal{X}^\epsilon$ respectively, in the space of $\mathcal{C}^{r-1}$ functions.

(2) *If $\tau_0 > \lambda^{-(r-1)}$, then the rate of mixing is robust, i.e., $\tau_\epsilon \to \tau_0$ as $\epsilon \to 0$. Furthermore, if $\tau_0 > \lambda^{-(r-2)}$ then $|\tau_\epsilon - \tau_0| = O(\epsilon^{1/d})$ for some integer $d \geq 1$.*

We show in fact that

(3) *For each $\delta > 0$ outside of $\{|z| \leq \lambda^{-(r-1)} + \delta\}$, the spectrum of $\mathcal{L}_\epsilon$ converges to that of $\mathcal{L}$ as $\epsilon \to 0$.*

The proofs below yield the same results for small *deterministic* perturbations by translations (i.e., maps $f^\epsilon = f + t$ with $|t| \leq \epsilon$), as well as for perturbations of expanding and $\mathcal{C}^r$ transformations of higher-dimensional tori.

D. Dynamical lemmas.

In this section we prove the dynamical lemmas which will allow us to reduce Theorem 1 to an abstract statement about linear operators acting on Banach spaces (see Section 2). The setting and notations are as in Sections 3.A and 3.B.

Lemma 4.

(1) *For a fixed $n \in \mathbb{Z}^+$ and $\varphi \in \mathcal{C}^{r-1}$*

$$\|\mathcal{L}_\epsilon^n \varphi - \mathcal{L}^n \varphi\| \to 0 \ \text{ as } \epsilon \to 0 \, .$$

(2) *For a fixed $n \in \mathbb{Z}^+$ and $\varphi \in \mathcal{C}^{r-1}$, we have in the $\mathcal{C}^{r-2}$ norm $\| \cdot \|_{r-2}$*

$$\|\mathcal{L}_\epsilon^n \varphi - \mathcal{L}^n \varphi\|_{r-2} = O(\epsilon) \, , \quad \epsilon \to 0 \, .$$

Proof of Lemma 4. It suffices to show the lemma for $n = 1$, the inductive step follows from the triangle inequality

$$\begin{aligned} \|\mathcal{L}_\epsilon^n \varphi - \mathcal{L}^n \varphi\| &= \|\mathcal{L}_\epsilon(\mathcal{L}_\epsilon^{n-1}\varphi) - \mathcal{L}(\mathcal{L}^{n-1}\varphi)\| \\ &\leq \|\mathcal{L}_\epsilon(\mathcal{L}_\epsilon^{n-1}\varphi - \mathcal{L}^{n-1}\varphi)\| + \|\mathcal{L}_\epsilon(\mathcal{L}^{n-1}\varphi) - \mathcal{L}(\mathcal{L}^{n-1}\varphi)\| \, . \end{aligned}$$

(The induction hypothesis need only be applied to φ and $\mathcal{L}^{n-1}\varphi$.)

(1) Since $\mathcal{L}_\epsilon\varphi = \theta_\epsilon * \mathcal{L}\varphi$, each derivative satisfies $D^k(\mathcal{L}_\epsilon\varphi) = \theta_\epsilon * D^k(\mathcal{L}\varphi)$. It hence suffices to consider $\mathcal{C}^0$-norms. But if ψ is continuous the convolution $\theta_\epsilon * \psi$ converges uniformly to ψ.

(2) To show the claimed asymptotic scaling in the $\mathcal{C}^{r-2}$ norm, it again suffices to consider the case $r = 2$. Observe that for any $\psi \in \mathcal{C}^1$ the Mean Value Theorem implies

$$\begin{aligned}|\theta_\epsilon * \psi(x) - \psi(x)| &\le \int \theta_\epsilon(t)|(\psi(x-t) - \psi(x))|\, dt \\ &\le \sup_\xi |\psi'(\xi)| \cdot \int \theta_\epsilon(t) \cdot t\, dt \\ &\le \sup_\xi |\psi'(\xi)| \cdot 2\epsilon\,. \quad \square\end{aligned}$$

We want to emphasize that in general $\mathcal{L}_\epsilon$ does *not* converge to $\mathcal{L}$ in the operator topology when $\epsilon \to 0$. (For example, if θ is $\mathcal{C}^{r-1}$, the operators $\mathcal{L}_\epsilon$ are all compact and convergence in norm would imply that $\mathcal{L}$ is compact too — but this is well-known to be false: see the explicit construction of essential spectral values in Collet–Isola [1991] for this model, and in Keller [1984] for the model considered in Section 5.)

The key lemma follows:

Lemma 5. *For $\Lambda > \lambda^{-(r-1)}$, there exists $N_0 \in \mathbb{Z}^+$ such that, for each $n \ge N_0$, there exists $\epsilon(n) > 0$ such that, for each $\epsilon < \epsilon(n)$, one has*

$$\|\mathcal{L}_\epsilon^n - \mathcal{L}^n\| < \Lambda^n\,.$$

Proof of Lemma 5. We use the following notations: C represents a constant independent of n and ϵ; $c_{n,\epsilon}$ represents a constant depending only on n and ϵ (and not on test functions), and tending to zero as $\epsilon \to 0$, for each fixed n. We also write g for $1/|f'|$. Recall that

$$\begin{aligned}(\mathcal{L}^n\varphi)(x) &= \sum_{y: f^n y = x} \varphi(y)(g(y)\cdot g(fy)\cdots g(f^{n-1}y)) \\ &= \sum_{y: f^n y = x} (\mathcal{L}^n\varphi_y)\,,\end{aligned}$$

where the second equality defines $(\mathcal{L}^n \varphi_y)$. Writing, for $\vec{t} = (t_1, \dots, t_n)$,

$$f^n_{\vec{t}}(z) = f(\dots f(f(z) + t_1) + t_2) \dots) + t_n \,,$$

we have

$$\begin{aligned}
(\mathcal{L}^n_\epsilon \varphi)(x) &= \int \cdots \int dt_1 \dots dt_n \, \theta_\epsilon(t_1) \dots \theta_\epsilon(t_n) \sum_{y_{\vec{t}} : f^n_{\vec{t}}(y_{\vec{t}}) = x} \varphi(y_{\vec{t}}) g(y_{\vec{t}}) \cdots g(f^{n-1}_{\vec{t}} y_{\vec{t}}) \\
&= \int \cdots \int dt_1 \dots dt_n \, \theta_\epsilon(t_1) \dots \theta_\epsilon(t_n) \sum_{y_{\vec{t}} : f^n_{\vec{t}}(y_{\vec{t}}) = x} (\mathcal{L}^n_{\vec{t}} \varphi)_{y_{\vec{t}}} \\
&= \int \cdots \int dt_1 \dots dt_n \, \theta_\epsilon(t_1) \dots \theta_\epsilon(t_n) \, (\mathcal{L}^n_{\vec{t}} \varphi)(x) \,,
\end{aligned}$$

where the last two equalities define $(\mathcal{L}^n_{\vec{t}} \varphi)$ and $(\mathcal{L}^n_{\vec{t}} \varphi)_{y_{\vec{t}}}$.

We have used the fact that all orbits are *strongly shadowable*: that is, if ϵ is small enough, then for a fixed x and a fixed n-tuple $(t_1, \dots, t_n)$, with $|t_i| \le \epsilon$, there is a natural bijection between the y such that $f^n(y) = x$ and the $y_{\vec{t}}$ such that $f^n_{\vec{t}}(y_{\vec{t}}) = x$. Moreover, for each pair $(y, y_{\vec{t}})$ corresponding to a choice of an inverse branch of f^n at x we have

$$g(y) \cdot g(fy) \cdots g(f^{n-1} y) = g(y_{\vec{t}}) \cdot g(f_{\vec{t}} y_{\vec{t}}) \cdots g(f^{n-1}_{\vec{t}} y_{\vec{t}}) \pm c_{n,\epsilon} \,. \tag{3.1}$$

We first show the lemma in the case $r = 2$. Let us compare $\mathcal{L}$ and $\mathcal{L}_\epsilon$ in the $\mathcal{C}^0$-norm, noting $|\varphi| = \sup |\varphi|$ and $|\varphi'| = \sup |\varphi'|$.

$$\begin{aligned}
(\mathcal{L}^n_{\vec{t}} \varphi)_{y_{\vec{t}}} &= (\varphi(y) \pm c_{n,\epsilon} |\varphi'|)(\prod_{j=0}^{n-1} g(f^j y) \pm c_{n,\epsilon}) \\
&= (\mathcal{L}^n \varphi)_y \pm c_{n,\epsilon}(|\varphi| + |\varphi'|) \,.
\end{aligned} \tag{3.2}$$

Hence, summing over inverse branches, and integrating over the t_i,

$$(\mathcal{L}^n_\epsilon \varphi)(x) = (\mathcal{L}^n \varphi)(x) \pm c_{n,\epsilon} \|\varphi\|_1 \,. \tag{3.3}$$

We now consider first derivatives, using the Leibnitz Theorem and decomposing

$$\frac{d}{dx} (\mathcal{L}^n_{\vec{t}} \varphi)_{y_{\vec{t}}}$$

into a first part A which is a sum of terms where some g factor is differentiated and a second part B where φ is differentiated. For the first part we have

$$\begin{aligned}
A &= \sum_{j=0}^{n-1} \varphi(y_{\vec{t}}) g(y_{\vec{t}}) \cdots [g'(f_{\vec{t}}^{j} y_{\vec{t}}) g(f_{\vec{t}}^{j} y_{\vec{t}}) \ldots g(y_{\vec{t}})] g(f_{\vec{t}}^{j+1} y_{\vec{t}}) \ldots g(f_{\vec{t}}^{n-1} y_{\vec{t}}) \\
&= \sum_{j} (\varphi(y) \pm c_{n,\epsilon} |\varphi'|) \big(g(y) \cdots [g'(f^j(y)) \cdots] \cdots g(f^{n-1} y) \pm c_{n,\epsilon} \big) \qquad (3.4) \\
&= (\text{ the corresponding part for } \frac{d}{dx} (\mathcal{L}^n \varphi)_y \ ) \pm c_{n,\epsilon} (|\varphi| + |\varphi'|) \,.
\end{aligned}$$

For the second part, we get

$$\begin{aligned}
B &= \varphi'(y_{\vec{t}}) \cdot \prod_{j=0}^{n-1} g(f_{\vec{t}}^{j} y_{\vec{t}}) \cdot \prod_{j=0}^{n-1} g(f_{\vec{t}}^{j} y_{\vec{t}}) \\
&= (\varphi'(y) \pm 2|\varphi'|) \cdot (\prod_{j=0}^{n-1} g(f^j y) \pm c_{n,\epsilon}) \cdot (\prod_{j=0}^{n-1} g(f^j y) \pm c_{n,\epsilon}) \qquad (3.5) \\
&= \varphi'(y) \big(\prod_{j=0}^{n-1} g(f^j y) \big)^2 \pm c_{n,\epsilon} |\varphi'| \pm 2|\varphi'| \lambda^{-n} \prod_{j=0}^{n-1} g(f^j y) \,.
\end{aligned}$$

Summing over inverse branches, and integrating over the t_i, we obtain

$$(\mathcal{L}_\epsilon^n \varphi)' = (\mathcal{L}^n \varphi)' \pm c_{n,\epsilon} \|\varphi\|_1 \pm 2\|\varphi\|_1 \lambda^{-n} \sum_{y : f^n(y) = x} \prod g(f^j(y)) \,. \qquad (3.6)$$

Since the sum in the last term of the right-hand-side is equal to $\mathcal{L}^n(1)(x)$, we know that it is uniformly bounded since $\mathcal{L}^n(1)$ converges.

For arbitrary differentiability r, note that for $k \le r-2$, the terms of the kth derivative $(\mathcal{L}^n \varphi)_y^{(k)}$ involve only the ℓth derivative of φ for $\ell \le k$ so that

$$|(\mathcal{L}_\epsilon^n \varphi)^{(k)} - (\mathcal{L}^n \varphi)^{(k)}| \le c_{n,\epsilon} \|\varphi\|_{k+1} \le c_{n,\epsilon} \|\varphi\|_{r-1} \,.$$

The only potentially troublesome term is part B of $(\mathcal{L}_\epsilon^n \varphi)^{(r-1)}(x)$, i.e.,

$$\int \ldots \int dt_1 \ldots dt_n \, \theta_\epsilon(t_1) \ldots \theta_\epsilon(t_n) \sum_{y_{\vec{t}}} \varphi^{(r-1)}(y_{\vec{t}}) \big(\prod_j g(f_{\vec{t}}^{j} y_{\vec{t}}) \big)^r ,$$

but the same argument as above yields an additional error term of the type

$$c_{n,\epsilon}\|\varphi\|_{r-1} + C \cdot \lambda^{-n(r-1)}\|\varphi\|_{r-1} \,. \quad (3.7) \quad \square$$

In fact, we have not used the expanding condition as stated but only a slightly weaker condition:

$$\exists \lambda > 1 \text{ such that } \lim_{n\to\infty}\big(\inf_x |f^{n\prime}(x)|^{1/n}\big) > \lambda \,.$$

Remark. If we go a little more carefully through the proof of Lemma 5, we can see that $\epsilon(n) = O(\Lambda^n/n) = O(\Lambda^n)$ as $n \to \infty$ *if we assume that* $f^{(r)}$ *is Lipschitz.* Indeed, by a simple distortion estimate, in (3.1),

$$|c_{n,\epsilon}| \le g(y)\cdots g(f^{n-1}y)\cdot(1 - e^{n\cdot L(f')\cdot\eta(\epsilon)}) \,,$$

where $L(f')$ is a Lipschitz constant for f', and $d(f^k y, f^k_{\vec{t}} y_{\vec{t}}) \le \eta(\epsilon) = O(\epsilon)$ is the bound from the shadowing lemma. Hence, in (3.2)

$$|c_{n,\epsilon}| \le g(y)\cdots g(f^{n-1}y)\cdot(\epsilon + (1+\epsilon)\cdot(1 - e^{n\cdot L(f')\cdot\eta(\epsilon)})) \,,$$

so that in (3.3)

$$|c_{n,\epsilon}| \le C(L(f'))\cdot n \cdot \epsilon \,,$$

where the constant C is independent of n, φ and ϵ. One obtains analogous bounds for the constants $c_{n,\epsilon}$ in (3.4) and (3.5) (note that the Lipschitz constant of f'' appears in (3.4)). Thus, in (3.6)

$$|c_{n,\epsilon}| \le C(L(f'), L(f''))\cdot n \cdot \epsilon \,.$$

Finally, in (3.7)

$$|c_{n,\epsilon}| \le C(L(f'), \ldots, L(f^{(r)}))\cdot n \cdot \epsilon \,.$$

The distortion argument is essentially the remark that

$$|\log(\frac{h_1(x_1)\cdot h_2(x_2)}{h_1(y_1)\cdot h_2(y_2)})| = \sum_{i=1,2} |\log h_i(x_i) - \log h_i(y_i)| \le 2\cdot \sup L(\log h_i)\cdot \sup|x_i - y_i| \,.$$

E. Proof of Theorem 1.

Unless otherwise stated we will use the results in Section 2 with X the space of $\mathcal{C}^{r-1}$ functions on S^1, $\|\cdot\|$ the $\mathcal{C}^{r-1}$ norm, $T_0 = \mathcal{L}$ and $T_\epsilon = \mathcal{L}_\epsilon$.

To prove (1), we let $\Sigma_0 = \{1\}$, Lemma 5 together with the fact that (f, μ_0) is exact tell us the conditions (A.1) to (A.3) in Section 2 are met. We also know that $\|\mathcal{L}_\epsilon\|$ is uniformly bounded, that 1 is always an eigenvalue of $\mathcal{L}_\epsilon$ and ρ_ϵ is an eigenfunction for 1. We conclude from Lemma 1 that X_0^ϵ must be the linear span of ρ_ϵ. Lemma 3 then tells us that for any $\kappa_0' < 1$, $\|\rho_\epsilon - \rho_0\| = O(C_1(\epsilon) + \frac{C_N(\epsilon)}{\kappa_0'^N})^{1/d})$ which tends to zero as $\epsilon \to 0$ by Lemma 4 (1), proving stochastic stability in $(\mathcal{C}^{r-1}(S^1), \|\cdot\|)$. Since $C_N(\epsilon) := \|\mathcal{L}_\epsilon^N \rho_0 - \rho_0\|$, the speed with which $C_N(\epsilon)$ tends to 0 depends on the modulus of continuity of $D^{r-1}\rho_0$. In particular, if we rewrite everything with $X = \mathcal{C}^{r-2}(S^1)$ and $\|\cdot\|$ the $\mathcal{C}^{r-2}$ norm, then $D^{r-2}\rho_0$ is Lipschitz and we have by Lemma 4 (2) $C_N(\epsilon) = O(\epsilon)$. This completes the proof of (1).

To prove (2), we let $\Sigma_0 = \sigma(\mathcal{L}) \cap \{|z| \geq \tau_0\}$. Note that conditions (A.1) and (A.2) in Section 2 are guaranteed by our assumption that $\tau_0 > \lambda^{-(r-1)} \geq \text{ess sp}(\mathcal{L})$. Since $\sigma(\mathcal{L}_\epsilon) \subset (\sigma(\mathcal{L}_\epsilon|_{X_0^\epsilon}) \cup \sigma(\mathcal{L}_\epsilon|_{X_1^\epsilon}))$, we know that $\tau_\epsilon = \sup\{|z| : z \in \sigma(\mathcal{L}_\epsilon|_{X_0^\epsilon}), z \neq 1\}$. Lemma 3 then tells us that for any $\tau_0' < \tau_0$, $|\tau_0 - \tau_\epsilon| = O((C_1(\epsilon) + \frac{C_N(\epsilon)}{\tau_0'^N})^{1/d})$, proving the robustness of τ_0.

To see how $|\tau_\epsilon - \tau_0|$ scales with ϵ, we let $\mathcal{L}$ act on $(\mathcal{C}^{r-2}(S^1), \|\cdot\|_{r-2})$ instead of $(\mathcal{C}^{r-1}(S^1), \|\cdot\|_{r-1})$. Since the eigenfunctions of $\mathcal{L}$ are always $\mathcal{C}^{r-1}$, the rates of decay of correlation are the same in both cases provided that $\tau_0 > \lambda^{-(r-2)}$ (note that this implies in particular $r > 2$). So even as we change the space on which $\mathcal{L}$ acts, the definition of Σ_0 remains unchanged. In fact, X_0 stays the same (Ruelle [1989]). In the definition of $C_N(\epsilon)$, we are now dealing with $\mathcal{C}^{r-2}$ norms for functions in X_0, a finite dimensional subspace of $\mathcal{C}^{r-1}(S^1)$. By Lemma 4 (2), we have $C_N(\epsilon) = O(\epsilon)$. Hence $|\tau_\epsilon - \tau_0| = O(\epsilon^{1/d})$.

To prove (3), let $\Sigma_0 = \sigma(\mathcal{L}) \cap \{|z| \geq \lambda^{-(r-1)} + \delta\}$. $\square$

Remark. By the remark following the proof of Lemma 5, we have $\epsilon(N) = O(\Lambda^N/N) = O(\Lambda^N)$ as $N \to \infty$ (if $f^{(r)}$ is Lipschitz) and hence $N = O(\log\epsilon/\log\Lambda)$ as $\epsilon \to 0$. This implies that there exists a constant $C > 0$ such that for any $1 \geq \beta > \Lambda$ we have $C_N(\epsilon)/\beta^N = O(\epsilon^{C\cdot(1-(\log\beta/\log\Lambda))})$.

If we could control analogously $\epsilon(N)$ *from Lemma 9 (9')*, this observation, combined with Lemma 8, could be used in Theorem 3 (3') to obtain a scaling result in the L^1-norm (assuming that $f^{(2)}$ in Section 5 is piecewise Lipschitz).

4. Expanding maps of manifolds followed by stochastic flows

This is a generalization of Section 3.

A. The unperturbed model.

Here, M is a $\mathcal{C}^\infty$ compact, connected Riemannian manifold without boundary, and $f : M \to M$ is a $\mathcal{C}^r$ map for some $2 \leq r < \infty$. We assume that f is *expanding*, i.e., there exists $\lambda > 1$ such that for all x in M and all v in T_xM, we have $|Df_xv| \geq \lambda|v|$. The largest such λ is called the *expanding constant* of f. It is well-known that an expanding map f admits a unique absolutely continuous invariant probability measure $\mu_0 = \rho_0 dm$ with respect to which f is exact (see e.g. Mañé [1987]).

Let $\mathcal{F} = \{\varphi : M \to \mathbb{R} : \varphi \text{ is } \mathcal{C}^{r-1}\}$. For $\varphi \in \mathcal{F}$, we define $\|\varphi\|$ to be the $\mathcal{C}^{r-1}$-norm of φ, defined using a set of charts that will remain fixed throughout. The Perron-Frobenius operator $\mathcal{L} : \mathcal{F} \to \mathcal{F}$ is defined as usual. Ruelle's results stated in the last section are in fact proved in this more general setting. In particular, we have the inequality

$$\text{ess sp}(\mathcal{L}) \leq \lambda^{-(r-1)} .$$

B. Type of perturbation: time-ϵ-maps of stochastic flows.

Let $X_0, X_1, \ldots, X_m$ be $\mathcal{C}^\infty$ vector fields on M, and consider the stochastic differential equation of Stratonovich type

$$d\xi_t = X_0\, dt + \sum_{i=1}^{m} X_i \circ d\beta_t^i \,, \tag{4.1}$$

where $\{\beta_t^i\}$ is the standard m-dimensional Brownian motion. We define $\mathcal{X}^\epsilon$, our ϵ-perturbation of f, to be $\xi_\epsilon \circ f$, i.e., $\mathcal{X}^\epsilon$ is the Markov chain whose transition probabilities are given by

$$P^\epsilon(x, E) = \text{Prob}\,\{(\xi_\epsilon \circ f)(x) \in E\}\,.$$

(Observe that ϵ now plays the role of a small *time*, in Section 3 the number ϵ was a small *displacement*.)

As in the last section, we wish to view $\mathcal{X}^\epsilon$ as the composition of random maps. To do that we realize the solution of (4.1) as a stochastic flow $\{\xi_t\}_{t\geq 0}$, i.e., we realize the solution of (4.1) as a Diff $^\infty(M)$-valued stochastic process $\{\xi_t\}$ satisfying

(i) $\xi_0 = \text{Id}$, the identity map,

(ii) for $t_0 < t_1 < \ldots < t_n$, the increments $\xi_{t_i} \circ \xi_{t_{i-1}}^{-1}$ are independent,

(iii) for $s < t$, the composition $\xi_t \circ \xi_s^{-1}$ depends only on $t - s$,

(iv) with probability 1, the stochastic flow ξ_t has continuous sample paths.

(See, e.g. Kunita [1990] for more information.) Let ν_ϵ denote the distribution of ξ_ϵ on Diff $^\infty(M)$. Then $\mathcal{X}^\epsilon$ is equivalent to the random map

$$\cdots \circ (\xi_\epsilon(\omega_2) \circ f) \circ (\xi_\epsilon(\omega_1) \circ f)\,,$$

where $\xi_\epsilon(\omega_1)$, $\xi_\epsilon(\omega_2)$, ... are i.i.d. with law ν_ϵ.

Using this representation of $\mathcal{X}^\epsilon$, we can write the perturbed Perron-Frobenius operator $\mathcal{L}_\epsilon : \mathcal{C}^{r-1}(M) \to \mathcal{C}^{r-1}(M)$ as follows. Let $f_\omega = \xi_\epsilon(\omega) \circ f$, then

$$(\mathcal{L}_\epsilon \varphi)(x) = \int \nu_\epsilon(d\omega)(\mathcal{L}_\omega \varphi)(x)\,,$$

where

$$(\mathcal{L}_\omega \varphi)(x) = \sum_{y : f_\omega y = x} \frac{\varphi(y)}{|\det Df_\omega(y)|}\,.$$

In fact, $\mathcal{L}_\epsilon$ is still in the framework studied by Ruelle [1990] and in particular is quasicompact. Again, $\mathcal{L}_\epsilon$ has 1 as an eigenvalue, with eigenfuction $\rho_\epsilon \in \mathcal{C}^{r-1}$ equal to the density of the invariant measure for $\mathcal{X}^\epsilon$.

In the remainder of this subsection we summarize a few technical estimates about the $\mathcal{C}^r$-norms of ξ_ϵ that will be needed later on. For $\xi \in \mathrm{Diff}^{\,r}(M)$, we define the $\mathcal{C}^r$-norm $\|\xi\|_r$ to be $\|\xi\|_r = \sum_{i=0}^r |D^i\xi|$, where $|D^i\xi|$ is computed using a fixed system of charts, and let $|||\xi||| := \max(\|\xi\|_r, \|\xi^{-1}\|_r)$. We assume that $|||\mathrm{Id}||| = 1$. For $\delta > 0$, we define the sets

$$\mathcal{U}_\delta := \{\xi \in \mathrm{Diff}^{\,r}(M) \,:\, |||\xi||| < 1+\delta\}$$

$$\mathcal{U}_\delta^n := \{\xi = \xi_n \circ \cdots \circ \xi_1 \,:\, \xi_i \in \mathcal{U}_\delta \,, \forall i\}\,,$$

and the random variable $\tau_n(\delta) := \inf\{s \,:\, \xi_s \notin \mathcal{U}_\delta^n\}$.

It is proved in Baxendale [1984] and Kifer [1988b] that for all $\epsilon > 0$

$$P\{\tau_n(\delta) \le \epsilon\} \le (P\{\tau_1(\delta) \le \epsilon\})^n\,.$$

Also, using a formula in Franks [1979, Lemma 3.2], we obtain inductively that for all ξ in $\mathcal{U}_\delta^n$,

$$|||\xi||| \le C^{n-1}(1+\delta)\Big[(1+\delta)^r + 1\Big]^{n-1}\,,$$

where the constant C only depends on r. From these estimates, we easily derive the following sublemmas:

Sublemma 1. *(Baxendale [1984], Kifer [1988b]). Fix $k > 0$. Then for all sufficiently small $\epsilon > 0$, the expectation*

$$E(|||\xi_\epsilon|||^k) < \infty\,.$$

Proof of Sublemma 1. Fix an arbitrary $\delta > 0$ and choose ϵ such that $P\{\tau_1(\delta) < \epsilon\}$ is sufficiently small. Let $\tau_0 = 0$, and define $A_n := \{\tau_{n-1}(\delta) \le \epsilon < \tau_n(\delta)\}$. Then

$$\begin{aligned} E|||\xi_\epsilon|||^k &\le \sum_{n=1}^\infty (\sup\{|||\xi||| \,:\, \xi \in \mathcal{U}_\delta^n\})^k \cdot P(A_n) \\ &\le \sum_{n=1}^\infty \Big[C^{n-1}(1+\delta)\big((1+\delta)^r + 1\big)^{n-1}\Big]^k \cdot \Big(P\{\tau_1(\delta) < \epsilon\}\Big)^{n-1} \\ &< \infty\,. \quad \square \end{aligned}$$

The proof of Sublemma 1 also gives the uniform integrability of $|||\xi_\epsilon|||^k$ as ϵ varies. We state that as Sublemma 2.

Sublemma 2. *Fix $k > 0$ and assume ϵ is small. Then given $\alpha > 0$, there exists $\beta > 0$ (independent of ϵ) such that*

$$\sup_{A:\nu_\epsilon(A)<\beta} E\big(|||\xi||| \cdot \chi_A\big)^k < \alpha \,.$$

Sublemma 3. *(Essentially in Baxendale [1984].) Fix $k > 0$. Then*

$$E|||\xi_\epsilon - Id|||^k \to 0 \quad \text{as } \epsilon \to 0 \,.$$

Proof of Sublemma 3. Write

$$E|||\xi_\epsilon - \mathrm{Id}|||^k = \sum_{n=1}^{\infty} E\big(|||\xi_\epsilon - \mathrm{Id}||| \cdot \chi_{A_n}\big)^k \,.$$

First let $\epsilon \to 0$ for fixed δ to get

$$\lim_{\epsilon \to 0} E|||\xi_\epsilon - \mathrm{Id}|||^k \leq \sup\{|||\xi - \mathrm{Id}||| \,:\, \xi \in \mathcal{U}_\delta\} \,.$$

The quantity on the right clearly tends to zero as $\delta \to 0$. $\square$

C. Statement of our results.

Theorem 2. *Let $f : M \to M$ be a $\mathcal{C}^r$ expanding map as described in Section 4.A, with expanding constant λ, and let $\mu_0 = \rho_0\, dm$ be its unique absolutely continuous invariant probability measure. Let $\{\mathcal{X}^\epsilon, \epsilon > 0\}$ be a small random perturbation of f of the type described in Section 4.B, with invariant probability measure $\mu_\epsilon = \rho_\epsilon\, dm$. Then:*

(1) *The dynamical system (f, μ_0) is stochastically stable under $\mathcal{X}^\epsilon$ in the space of $\mathcal{C}^{r-1}$ functions, i.e., the $\mathcal{C}^{r-1}$-norm of $\rho_\epsilon - \rho_0$ tends to zero as $\epsilon \to 0$.*

Let τ_0 and τ_ϵ be the rates of decay of correlation functions for f and $\mathcal{X}^\epsilon$ respectively, in the space of $\mathcal{C}^{r-1}$ functions. If, in addition, $\tau_0 > \lambda^{-(r-1)}$, then:

(2) *The rate of mixing for f is robust, i.e., $\tau_\epsilon \to \tau_0$ as $\epsilon \to 0$.*

We show in fact that

(3) *For each $\delta > 0$, outside of $\{|z| \leq \lambda^{-(r-1)} + \delta\}$, the spectrum of $\mathcal{L}_\epsilon$ converges to that of $\mathcal{L}$ as $\epsilon \to 0$.*

Remark. We believe, but do not know how to prove rigorously, that the correct scaling in ϵ for this kind of perturbation is $\|\rho_\epsilon - \rho_0\|_{r-2} = O(\sqrt{\epsilon})$.

D. Dynamical lemmas.

The setting and all notations are as in Sections 4.A and B, and except for the scaling statement the two lemmas needed are identical to those in Section 3. Once again, they are:

Lemma 6. *For fixed $n \in \mathbb{Z}^+$ and $\varphi \in \mathcal{C}^{r-1}$,*

$$\|\mathcal{L}_\epsilon^n \varphi - \mathcal{L}^n \varphi\| \to 0 \quad \text{as } \epsilon \to 0\,.$$

Lemma 7. *For $\Lambda > \lambda^{-(r-1)}$, there exists $N_0 \in \mathbb{Z}^+$ such that for all $n \geq N_0$ there exists $\epsilon(n) > 0$ such that for each $\epsilon < \epsilon(n)$,*

$$\|\mathcal{L}_\epsilon^n - \mathcal{L}^n\| < \Lambda^n\,.$$

We will use the proof of Lemma 7, with $r = 2$, to illustrate how the analysis in Section 3.D can be adapted to the present setting. The other proofs are handled similarly.

We use the random maps representation of $\mathcal{X}^\epsilon$, i.e., we consider the probability space (Ω, ν_ϵ) where Ω can be identified with Diff $^r(M)$ and ν_ϵ is the distribution of ξ_ϵ. We let $\xi_\epsilon(\omega)$ denote the diffeomorphism corresponding to $\omega \in \Omega$, and write $f_\omega = \xi_\epsilon(\omega) \circ f$. Using the notation in Section 3.D, we have $f_{\vec{\omega}}^n = f_{\omega_n} \circ \cdots \circ f_{\omega_1}$ if $\vec{\omega} = (\omega_1, \ldots, \omega_n) \in \Omega^n$, and

$$(\mathcal{L}_\epsilon^n \varphi)(x) = \int \cdots \int \nu_\epsilon(d\omega_1) \cdots \nu_\epsilon(d\omega_n)\, (\mathcal{L}_{\vec{\omega}}^n \varphi)(x)\,,$$

where

$$(\mathcal{L}_{\vec{\omega}}^n \varphi)(x) = \sum_{y : f_{\vec{\omega}}^n y = x} \varphi(y) \cdot \frac{1}{|\det Df_{\vec{\omega}}^n(y)|}\,.$$

Let n be fixed for now. For local considerations we will assume that we are in Euclidean space.

Sublemma 4.

$$\frac{d}{dx_i}(\mathcal{L}_\epsilon^n\varphi) = \int\cdots\int \nu_\epsilon(d\omega_1)\cdots\nu_\epsilon(d\omega_n)\,\frac{d}{dx_i}(\mathcal{L}_{\vec\omega}^n\varphi)\,.$$

Proof of Sublemma 4. We fix $x \in M$, and write

$$\frac{d}{dx_i}(\mathcal{L}_{\vec\omega}^n\varphi)(x) = \lim_{t\to 0}\Phi_t(\vec\omega)\,,$$

where

$$\Phi_t(\vec\omega) = \frac{1}{t}\Big\{\big(\mathcal{L}_{\vec\omega}^n\varphi\big)(x+t\,u_i) - \big(\mathcal{L}_{\vec\omega}^n\varphi\big)(x)\Big\} = \frac{d}{dx_i}\big(\mathcal{L}_{\vec\omega}^n\varphi\big)(x_t)\,,$$

for some x_t, where u_i is the unit vector in the i^{th} direction. Our assertion amounts to exchanging the order of the limit and integrals. To do that, we will produce $\Phi \in L^1(\Omega^n, \nu_\epsilon^n)$ with $|\Phi_t| \le |\Phi|$. Differentiating the expression for $\mathcal{L}_{\vec\omega}^n\varphi$ above, we observe that $\frac{d}{dx_i}(\mathcal{L}_{\vec\omega}^n\varphi)(x_t)$ is the sum of finitely many terms, each one of which is bounded in absolute value by a product of the form

$$C\cdot\|\varphi\|_1\cdot |||\xi_\epsilon(\omega_1)|||^{k_1}\cdots|||\xi_\epsilon(\omega_n)|||^{k_n}\,,$$

where C is a constant depending on f and n, and $k_1,\dots,k_n$ depend on n and the dimension of M. We set $\Phi(\vec\omega)$ to be the corresponding sum. It follows from Sublemma 1 that Φ is integrable. Hence the Dominated Convergence Theorem applies. $\square$

Consider first $\vec\omega = (\omega_1,\dots,\omega_n)$ where $f_{\vec\omega}^n$ is $\mathcal{C}^2$ very near f^n, say $\|f_{\vec\omega}^n - f^n\|_2 < \delta$ for some $\delta > 0$. We assume δ is small enough so that the inverse branches of $f_{\vec\omega}^n$ are easily identified with those of f^n. Then the same argument as in Section 3.D, line by line, gives

$$\mathcal{L}_{\vec\omega}^n\varphi = \mathcal{L}^n\varphi \pm c_{n,\delta}\|\varphi\|_1\,,$$

and

$$\frac{d}{dx_i}(\mathcal{L}_{\vec\omega}^n\varphi) = \frac{d}{dx_i}(\mathcal{L}^n\varphi) \pm c_{n,\delta}\|\varphi\|_1 \pm C\lambda^{-n}\|\varphi\|_1\,.$$

The strategy of our proof is as follows: first we choose n and then $\delta = \delta(n)$ so that for all $\vec{\omega}$ with the properties above, we have

$$\|\mathcal{L}^n_{\vec{\omega}}\varphi - \mathcal{L}^n\varphi\| \le {\Lambda'}^n \|\varphi\|\,, \text{ for some } \lambda^{-(r-1)} < \Lambda' < \Lambda\,.$$

We then choose $\epsilon \ll \delta$ such that if $\Omega_0 := \{\omega \,:\, \|f_\omega - f\|_2 \ge \delta\}$, then $\nu_\epsilon \Omega_0$ is very small, small enough that these "bad" $\vec{\omega}$ do not contribute significantly to $\|\mathcal{L}^n_\epsilon\varphi - \mathcal{L}^n\varphi\|$. More precisely, let

$$\Omega^n_0 := \{(\omega_1, \dots, \omega_n) \,:\, \omega_i \in \Omega_0\,,\, \forall i\}$$

and

$$\Omega^n_j := \{(\omega_1, \dots, \omega_n) \,:\, \omega_j \notin \Omega_0\}\,.$$

First we consider the $\mathcal{C}^0$-norm:

$$\begin{aligned}
|\mathcal{L}^n_\epsilon\varphi - \mathcal{L}^n\varphi| &= |\int_{\Omega^n} d\nu^n_\epsilon(\vec{\omega})\big(\mathcal{L}^n_{\vec{\omega}}\varphi - \mathcal{L}^n\varphi\big)| \\
&\le \int_{\Omega^n_0} |\mathcal{L}^n_{\vec{\omega}}\varphi - \mathcal{L}^n\varphi| + \sum_{j=1}^n \int_{\Omega^n_j} \Big(|\mathcal{L}^n_{\vec{\omega}}\varphi| + |\mathcal{L}^n\varphi|\Big)\,.
\end{aligned}$$

The Ω^n_0-term has been shown to be bounded above by $c_{n,\epsilon} \cdot \|\varphi\|_1$, and

$$\int_{\Omega^n_j} |\mathcal{L}^n\varphi| \le \|\mathcal{L}^n\| \cdot \|\varphi\|_1 \cdot \nu_\epsilon\Omega_0\,,$$

the last factor of which can be made small as $\epsilon \to 0$. It remains to estimate $\int_{\Omega^n_j} |\mathcal{L}^n_{\vec{\omega}}\varphi|$. Note that $\mathcal{L}^n_{\vec{\omega}}\varphi$ is a sum of finitely many terms of the form

$$\frac{\varphi(\cdot)}{|\det Df_{\omega_1}(\cdot)| \cdots |\det Df_{\omega_n}(\cdot)|}\,.$$

This expression is bounded above by

$$C \cdot |\varphi| \cdot |||\xi_\epsilon(\omega_1)|||^{k_1} \cdot \dots \cdot |||\xi_\epsilon(\omega_n)|||^{k_n}\,.$$

Its integral over Ω^n_j is therefore bounded above by

$$C \cdot |\varphi| \cdot \left(\prod_{i \ne j} E|||\xi_\epsilon|||^{k_i}\right) \cdot E\Big(|||\xi_\epsilon|||^{k_j} \cdot \chi_{\Omega_0}\Big)\,.$$

By Sublemma 2, the last factor can again be arranged to be arbitrarily small by choosing ϵ small. This proves

$$|\mathcal{L}_\epsilon^n \varphi - \mathcal{L}^n \varphi| \le c_{n,\epsilon} \cdot \|\varphi\|_1 \,.$$

A similar argument (see Sublemma 4) gives

$$|\frac{d}{dx_i} \mathcal{L}_\epsilon^n \varphi - \frac{d}{dx_i} \mathcal{L}^n \varphi| \le {\Lambda'}^n \|\varphi\|_1 + c_{n,\epsilon} \|\varphi\|_1 \le \Lambda^n \|\varphi\|_1 \,. \quad \square$$

Remark. We cannot use the convolution argument of Lemma 4 (1) to show Lemma 6. However, since $\varphi^{(k)}(y) = \varphi^{(k)}(z) + C(\varphi, d(y,z))$, where $\varphi \in \mathcal{C}^r$, $k \le r$ and the "error" $\|C(\varphi, d(y,z))\| \to 0$ as $d(y,z) \to 0$ for a fixed φ, we can prove Lemma 6 by a suitable adaptation of the proof of Lemma 7.

E. Proof of Theorem 2.

Use Section 2 and proceed as in Section 3.E.

5. Piecewise expanding maps of the interval

A. The unperturbed model.

We consider here $f : I \to I$, where $I = [0,1]$ and f is a continuous piecewise $\mathcal{C}^2$, piecewise expanding map. More precisely, we assume that there exists a partition $0 = a_0 < a_1 < \cdots < a_M = 1$ of I such that for each i, the restriction $f|_{[a_i, a_{i+1}]}$ can be extended to a $\mathcal{C}^2$ map with $\min |f'| \ge \lambda > 1$. The a_i are called the *turning points* of f. The continuity assumption on f is imposed only for simplicity of exposition. One could replace it by piecewise continuity and consider left-hand and right-hand limits of the turning points.

Recall that for $\varphi : I \to \mathbb{R}$, the total variation of φ on an interval $[a,b]$ is defined to be

$$\mathrm{var}_{\,[a,b]} \varphi = \sup\{\sum_{i=0}^{n} |\varphi(x_{i+1}) - \varphi(x_i)| \,:\, n \ge 1 \,, a \le x_0 < x_1 < \ldots < x_n \le b\} \,.$$

We use $|\varphi|_1 := \int_I |\varphi|$ to denote the L^1-norm of φ with respect to Lebesgue measure. Let $BV := \{\varphi : I \to \mathbb{C} \, : \, \mathrm{var}_I \varphi < \infty\}$. One often considers the Banach space $(BV, \|\cdot\|)$ where

$$\|\varphi\| = \mathrm{var}_I \varphi + |\varphi|_1 \, .$$

Let $\mathcal{L}$ be the Perron-Frobenius operator associated with f acting on $(BV, \|\cdot\|)$.

The spectrum of $\mathcal{L}$ in this setting has been studied by many people (Wong [1978], Rychlik [1983], Hofbauer–Keller [1982]). It has been shown that $\mathcal{L}$ is quasi-compact, its spectral radius is equal to one, it has unity as an eigenvalue, and its essential spectral radius is equal to

$$\Theta = \lim_{n\to\infty} (\sup(1/|(f^n)'|)^{1/n} \le 1/\lambda \, .$$

(The derivative of f is not well-defined at the turning points, but both limits $f'_+(a_i) = \lim_{x\downarrow a_i} f'(x)$ and $f'_-(a_i) = \lim_{x\uparrow a_i} f'(x)$ exist; we replace implicitly each occurrence of $f'(a_i)$ by the maximum of these two limits.)

Let ρ_0 be an eigenfunction for the eigenvalue 1, with $|\rho_0|_1 = 1$. Then ρ_0 is the density of an invariant probability measure μ_0 for f. We *assume* that f has no other absolutely continuous invariant probability measure, and that f is weak mixing with μ_0. Under these assumptions, it has been shown that 1 is the only point of $\sigma(\mathcal{L})$ on the unit circle, its generalized eigenspace is one-dimensional, and that $\tau_0 := \sup\{|z| \, : \, z \in \sigma(\mathcal{L}), z \neq 1\} < 1$ measures the exponential rate of decay of correlations for functions in BV (Hofbauer–Keller [1982], Keller [1984]).

In our analysis to follow, it will be necessary for us to work with some other norms in BV. For $0 < \gamma \le 1$, we define

$$\|\varphi\|_\gamma = \gamma \cdot \mathrm{var}_I \varphi + |\varphi|_1 \, .$$

Note that for any $0 < \gamma < \gamma'$ the norms $\|\cdot\|_\gamma$ and $\|\cdot\|_{\gamma'}$ are equivalent.

B. Type of perturbation: convolutions.

As in Section 3.B, we consider a small random perturbation $\mathcal{X}^\epsilon$ of f by convolution. Let us make the assumption that $f(I) \subset [\delta, 1-\delta]$, for some $\delta > 0$, so that we can avoid the problems at the boundary of I when f is perturbed. (There are other ways to deal with this.) We obtain as before a perturbed transfer operator $\mathcal{L}_\epsilon$ acting on $(BV, \|\cdot\|)$. As in the first two models, $\mathcal{L}_\epsilon$ has 1 as an eigenvalue with eigenfunction ρ_ϵ which is the density of an invariant probability measure μ_ϵ for $\mathcal{X}^\epsilon$.

If we had made the additional assumptions that θ_ϵ is continuous and of bounded variation, then we could easily prove that $\mathcal{L}_\epsilon$ is a compact operator. (First use continuity of θ_ϵ and the usual Ascoli argument to show that any sequence $\mathcal{L}_\epsilon \varphi_n = \int \theta_\epsilon(x - fy)\varphi_n(y)dy$, with $\|\varphi_n\| \leq 1$ has a subsequence which converges uniformly to a continuous function. Then use the fact that θ_ϵ is of bounded variation to prove that this subsequence is also a Cauchy sequence for the BV norm.)

Unfortunately, not all piecewise expanding maps are stochastically stable. A major difference between the situation here and that in Section 3 is that we do not have the kind of "shadowing" property used in the proof of Lemma 5. More precisely, let $\vec{t} = (t_1, \ldots, t_n)$ and $f^n_{\vec{t}}$ be as in Section 3.D. We count the smallest number of intervals on which f^n is monotone, for that measures in some way the number of "distinct orbits" of f. In general $f^n_{\vec{t}}$ may have many more intervals of monotonicity than f^n. See Figure 1 for an example in which a turning point fixed by f generates 2^n extra intervals of monotonicity for $f^n_{\vec{t}}$. Indeed, this example is not stochastically stable, not even in the sense of weak convergence of μ_ϵ (see Keller [1982, §6]).

We remark that the "shadowing" property used in our proof of Lemma 5 is not the usual shadowing property: we deal only with orbits of finite length but require a complete matching of backwards branches of the map. For more information on the usual shadowing for interval maps see Coven–Kan–Yorke [1988].

C. Statement of our results.

Figure 1 Comparing the fourth iterate of a map with $\mathcal{M} = \infty$ to a deterministically-perturbed iterate

It is clear from our discussion in the last subsection that our situation improves if the turning points do not get mapped near themselves. We say that f has no *periodic turning point* if $f^k(a_i) \neq a_i$ for all $k \geq 1$.

To make it easier to state our results, we will use the following language. For $\kappa_1 < \kappa_0$, we call the open annular region $A(\kappa_1, \kappa_0) := \{\kappa_1 < |z| < \kappa_0\}$ a *spectral gap* for $\mathcal{L}$ if $A(\kappa_1, \kappa_0) \cap \sigma(\mathcal{L}) = \emptyset$. Furthermore we will say that $A(\kappa_1, \kappa_0)$ satisfies

Assumption A. If $\max(\Theta, \kappa_1) < \kappa_0^2$.

Assumption B. If either $\max(2 \cdot \Theta, \kappa_1) < \kappa_0^2$,

or $\max((3/2) \cdot \Theta, \kappa_1) < \kappa_0^2$, and each θ_ϵ is symmetric.

(The kernel θ_ϵ used in our convolutions is called symmetric if $\theta_\epsilon(x) = \theta_\epsilon(-x)$, $\forall x$. The definition of Θ is given in Section 5.A.)

We first state our result assuming that f has no periodic turning points

Theorem 3. *Let $f : I \to I$ be as described in Section 5.A, with a unique absolutely continuous invariant probability measure $\mu_0 = \rho_0 \, dm$, and let $\mathcal{X}^\epsilon$ be a small random* *perturbation of f of the type described in Section 5.B with invariant probability measure $\rho_\epsilon\, dm$. We assume also that f has no periodic turning points. Then*

(1) *The dynamical system (f, μ_0) is stocastically stable under $\mathcal{X}^\epsilon$ in $L^1(dm)$, i.e., $|\rho_\epsilon - \rho_0|_1$ tends to 0 as $\epsilon \to 0$.*

Let τ_0 and τ_ϵ be the rates of decay of correlations functions for f and $\mathcal{X}^\epsilon$ respectively for test functions in BV.

(2) *If $\mathcal{L}$ has a spectral gap of the form $A(\tau, \tau_0)$, and $A(\tau, \tau_0)$ satisfies Assumption A, then $\tau_\epsilon \to \tau_0$ as $\epsilon \to 0$.*

We show in fact that

(3) *for every $\kappa_0 > 0$, if $\mathcal{L}$ has a spectral gap of the form $A(\kappa_1, \kappa_0)$ and Assumption A is satisfied, then there exists a small $\delta > 0$ such that outside of $\{|z| \le \kappa_0 - \delta\}$, the spectrum of $\mathcal{L}_\epsilon$ converges to that of $\mathcal{L}$ as $\epsilon \to 0$.*

Theorem 3'. *Let f and $\mathcal{X}^\epsilon$ be as in Theorem 3, except that we do not require that f has no periodic turning points. Then*

(1) *is true if either $\Theta < 1/2$ or $\Theta < 2/3$ and θ_ϵ is symmetric;*

(2) *and (3) are true if Assumption A is replaced by Assumption B.*

Assumptions A and B arise in part from our use of balanced norms in the proof of Lemma 9. We do not know to what extent they are needed — although it is clear from our discussion in Section 5.B that *some* hypothesis on f is necessary to give the type of results we want. We remark also that the hypothesis we use for proving stochastic stability is slightly weaker than that in Keller [1982,§6].

D. Dynamical lemmas.

The setting and notations are as in Sections 5.A and 5.B.

Lemma 8. *For a fixed $n \in \mathbb{Z}^+$ and $\varphi \in BV$*

$$|\mathcal{L}_\epsilon^n \varphi - \mathcal{L}^n \varphi|_1 = O(\epsilon) \ \text{ as } \epsilon \to 0\,.$$

Proof of Lemma 8. As in Section 3.D, it suffices to consider the case $n = 1$, the more general case follows by induction. Set $\psi = \mathcal{L}(\varphi)$. Then $\psi \in BV$. It is enough to show that

$$\int |\psi(x) - \psi(x-t)|\, dx = O(|t|)\,.$$

To show this, we can assume that ψ is monotone increasing. Then $\int |\psi(x) - \psi(x-t)|\, dx$ is the area of the subset of $I \times I$ between the graphs of $\psi(x)$ and $\psi(x-t)$. But this area is simply $|t| \cdot (\psi(1) - \psi(0))$. $\square$

Note that it is not true in general that $\mathrm{var}\,(\mathcal{L}_\epsilon \varphi - \mathcal{L}\varphi) \to 0$ as $\epsilon \to 0$.

We will use the notations $c_{n,\epsilon}$, $g = 1/|f'|$, and $f^n_{\vec{t}}$, $\mathcal{L}^n_{\vec{t}}$ of Section 3.D. We also write

$$g^n(y) = g(y) \cdot g(fy) \cdots g(f^{n-1}y)$$
$$g^n_{\vec{t}}(y_{\vec{t}}) = g(y_{\vec{t}}) \cdot g(f_{\vec{t}} y_{\vec{t}}) \cdots g(f^{n-1}_{\vec{t}} y_{\vec{t}})\,.$$

We note

$$M_i := \#\{k \,:\, k \geq 1, f^k(a_i) \in \{a_0, \ldots, a_M\}\}\,,$$

and $\mathcal{M} = \max M_i \leq M + 1$. Note that f is without periodic turning points if and only if $\mathcal{M} < \infty$.

Denote by $\mathcal{Z}_n$ the "partition" of I into (closed) intervals of monotonicity of f^n, and by $\mathcal{Z}_{n,\vec{t}}$ the "partition" of I into (closed) intervals of monotonicity for $f^n_{\vec{t}}$. Write $\mathcal{Z}_1 = \eta_1 \cup \ldots \cup \eta_M$. By definition an element $\eta(j_0, \ldots, j_{n-1})$ of $\mathcal{Z}_n$ is an interval of the form

$$\eta(j_0, \ldots, j_{n-1}) = \eta_{j_0} \cap f^{-1}(\eta_{j_1}) \cap \ldots \cap f^{-(n-1)}(\eta_{j_{n-1}})\,,$$

with nonempty interior; and an element $\eta'(j_0, \ldots, j_{n-1})$ of $\mathcal{Z}_{n,\vec{t}}$ is an interval of the form

$$\eta'(j_0, \ldots, j_{n-1}) = \eta_{j_0} \cap f^{-1}_{(t_1)}(\eta_{j_1}) \cap \ldots \cap f^{-(n-1)}_{(t_1, \ldots t_{n-1})}(\eta_{j_{n-1}})$$

with nonempty interior.

If $\mathcal{M} = 0$, it is not difficult to check that for fixed $n \geq 1$, there exists $\epsilon(n)$ such that, for all $\epsilon < \epsilon(n)$, the elements of $\mathcal{Z}_{n,\vec{t}}$ are in bijection with those of $\mathcal{Z}_n$. We say that two such intervals $\eta(j_0, \ldots, j_{n-1}) \in \mathcal{Z}_n$ and $\eta'(j_0, \ldots, j_{n-1}) \in \mathcal{Z}_{n,\vec{t}}$ are *associated* and that η' is *admissible.*

Consider now the case $1 \leq \mathcal{M}$ and a fixed value of n. For small enough ϵ, by continuity, with each $\eta(j_0, \ldots, j_{n-1}) \in \mathcal{Z}_n$ we can *associate* the element $\eta'(j_0, \ldots, j_{n-1}) \in \mathcal{Z}_{n,\vec{t}}$, which we call *admissible.* In general, when $\mathcal{M} \geq 1$, we do not obtain all intervals of $\mathcal{Z}_{n,\vec{t}}$ in this way. The remaining intervals are called *nonadmissible*, and we are going to show that there are not too many of them if $\mathcal{M} < \infty$. If $\eta(j_0, \ldots, j_{n-1})$ is empty then, by continuity, $\eta'(j_0, \ldots, j_{n-1})$ is also empty for small enough ϵ. It hence suffices to consider sets $\eta(j_0, \ldots, j_{n-1})$ which are reduced to a point x_0. (We consider the case $0 \neq x_0 \neq 1$, the remaining two cases are similar.) There are at most $2^{\mathcal{M}}$ different sequences for which $\eta(j_{0,p}, \ldots, j_{n-1,p}) = x_0$. Indeed, the only way an intersection can be reduced to a point is if there exist $0 \leq q_1 < q_2 < \ldots < q_L \leq n$ such that the iterates $f^{q_i}(x_0) = b_i$ lie in the turning set (assume no other iterates do). But then $f^{q_L - q_i}(b_i) = b_L$ and $q_L - q_1 \leq \mathcal{M}$, which proves our claim (because, there are at most $2^L \leq 2^{\mathcal{M}}$ different possibilities, which correspond to the at most $\mathcal{M}$ pairs of intervals on either side of the b_i). Hence, if ϵ is small enough, then at most $2^{\mathcal{M}}$ different nonadmissible elements, which are adjacent intervals of $\mathcal{Z}_{n,\vec{t}}$, can be generated between the admissible intervals $\eta'(k_0, \ldots, k_{n-1})$ and $\eta'(\ell_0, \ldots, \ell_{n-1})$, where $\eta(k_0, \ldots, k_{n-1})$ and $\eta(\ell_0, \ldots, \ell_{n-1})$ are the two intervals of $\mathcal{Z}_n$ which are to the left and the right of x_0. If $\mathcal{M} = \infty$, we can only show that there are at most 2^n such nonadmissible elements.

We now "trim" the intervals of $\mathcal{Z}_n$ and the admissible intervals of $\mathcal{Z}_{n,t}$.

Assume first that $\mathcal{M} = 0$ and ϵ is small enough. Let $\eta \in \mathcal{Z}_n$ and $\eta' \in \mathcal{Z}_{n,\vec{t}}$ be a pair of associated intervals of monotonicity. We decompose η and η' into two as follows: set $G(\eta, \eta') = f^n \eta \cap f^n_{\vec{t}} \eta'$ and $\eta_G = (f^n|_\eta)^{-1}(G)$, $\eta'_G = (f^n_{\vec{t}}|'_\eta)^{-1}(G)$; and let $\eta_B = \eta \setminus \eta_G$ and $\eta'_B = \eta' \setminus \eta'_G$. We again say that the intervals η_G and η'_G are *associated* and that

η_B and η'_B are their respective *co-respondents*. If we denote by B the union of all co-respondents η_B and by B' the union of all co-respondents η'_B, the measures of B and B' both tend to zero as ϵ tends to zero.

In the case where $1 \leq \mathcal{M}$, we have seen that with each interval $\eta \in \mathcal{Z}_n$ is associated one interval $\eta' \in \mathcal{Z}_{n,\vec{t}}$, and that to the right and to the left of η' there are at most $2^{\mathcal{M}}$ adjacent nonadmissible intervals $\eta'_p \in \mathcal{Z}_{n,\vec{t}}$ if $\mathcal{M} < \infty$, at most 2^n such intervals if $\mathcal{M} = \infty$. The intervals η and η' can be decomposed into $\eta' = \eta'_G \cup \xi_B$ and $\eta = \eta_G \cup \eta_B$ as just described in the case $\mathcal{M} = 0$. We again say that η_G and η'_G are *associated* and that η_B is the *co-respondent* of η_G. We define the *co-respondents* of η'_G to be ξ_B together with the first half of the non-admissible intervals immediately to the left and to the right of η' (with obvious modifications if there is an odd number of intervals, or if η' is at an extremity of I). Each non-admissible interval is hence the co-respondent of a unique η'_G. We denote by B the union of all the "bad" intervals η_B, and by B' the union of all co-respondents.

Lemma 9. *Assume that f has no periodic turning points and let $\Lambda^2 > \Theta$. Then, there exist $C > 0$ and $N_0 \in \mathbb{Z}^+$ such that, for each $n \geq N_0$ there exists $\epsilon(n) > 0$ such that for each $\epsilon < \epsilon(n)$,*

$$\|\mathcal{L}_\epsilon^n - \mathcal{L}^n\|_{\Lambda^n} < C \cdot \Lambda^n \,.$$

Proof of Lemma 9. In the proof, $\tilde{\Theta}$ denotes a constant slightly larger than Θ (we will have to increase $\tilde{\Theta}$ slightly a finite number of times in the argument). There exists an n_0 such that $g^n(x) \leq \tilde{\Theta}^n$ if $n \geq n_0$.

We have

$$\|\mathcal{L}^n\varphi - \mathcal{L}_{\vec{t}}^n\varphi\| \leq \|\mathcal{L}_{\vec{t}}^n(\varphi\chi_{B'})\| + \|\mathcal{L}^n(\varphi\chi_B)\| + \|\mathcal{L}^n(\varphi\chi_{(I\setminus B)} - \mathcal{L}_{\vec{t}}^n(\varphi\chi_{I\setminus B'})\| \,. \qquad (5.1)$$

We start with the details of the proof for the first "bad" term $\|\mathcal{L}_{\vec{t}}^n(\varphi\chi_{B'})\|$, the second "bad" term is obtained by similar (more classical) bounds. The third term will be considered in Equations (5.10) to (5.14) below.

For each $\eta'_B \in B'$ and for $x \in f^n_{\vec{t}}\eta'_B$, we have

$$\mathcal{L}^n_{\vec{t}}(\varphi\chi_{\eta'_B})(x) = \varphi(y_{\vec{t}}) \cdot g(y_{\vec{t}}) \cdots g(f^{n-1}y_{\vec{t}}) \, ,$$

where $y_{\vec{t}}$ is the unique element of η'_B such that $f^n_{\vec{t}}(y_{\vec{t}}) = x$. It follows that

$$|\mathcal{L}^n_{\vec{t}}(\varphi\chi_{\eta'_B})|_1 \le \int_{\eta'_B} |\varphi| \le \ell(\eta'_B) \cdot (\operatorname{var}\varphi + |\varphi|_1) \, , \tag{5.2}$$

where $\ell(\eta'_B)$ denotes the length of the interval η'_B.

Summing (5.2) over all intervals η'_B, we get

$$|\mathcal{L}^n_{\vec{t}}(\varphi\chi_{B'})|_1 \le c_{n,\epsilon} \cdot (\operatorname{var}\varphi + |\varphi|_1) \, . \tag{5.3}$$

For the variation, we have

$$\operatorname{var}\mathcal{L}^n_{\vec{t}}(\varphi\chi_{\eta'_B}) \le \operatorname{var}_{\eta'_B}\varphi \cdot \sup_{\eta'_B} g^n_{\vec{t}} + \sup_{\eta'_B}|\varphi| \cdot \operatorname{var}_{\eta'_B} g^n_{\vec{t}} + 2 \cdot \sup_{\eta'_B}|\varphi| \cdot \sup_{\eta'_B} g^n_{\vec{t}} \, . \tag{5.4}$$

Were it not for the last term of (5.4), everything would be much easier! To further bound the variation, we will use the following easily proved inequalities: if n is large enough, say $n \ge n_1$, and $\eta' \in \mathcal{Z}_{n,\vec{t}}$ for small enough ϵ, then

$$\begin{cases} \sup_{\eta'} g^n_{\vec{t}} & \le \tilde{\Theta}^n \\ \operatorname{var}_{\eta'} g^n_{\vec{t}} & \le \tilde{\Theta}^n \, . \end{cases} \tag{5.5}$$

Set $n_2 = \max(n_0, n_1)$ and assume first that $n = n_2$. The interval η'_B is a subset of some $\eta' \in \mathcal{Z}_{n,\vec{t}}$ and is a co-respondent of a unique good interval η'_G. ¿From (5.4) and (5.5), denoting by η'' the union of η'_G together with its at most $2^{\mathcal{M}-1}+1$ co-respondents, we obtain:

$$\begin{aligned} \operatorname{var}\mathcal{L}^n_{\vec{t}}(\varphi\chi_{\eta'_B}) &\le \operatorname{var}\mathcal{L}^n_{\vec{t}}(\varphi\chi_{\eta'}) \\ &\le \operatorname{var}_{\eta'}\varphi \cdot \tilde{\Theta}^n + \left(\operatorname{var}_{\eta''}\varphi + \inf_{\eta''}|\varphi|\right)\left(\operatorname{var}_{\eta'} g^n_{\vec{t}} + 2 \cdot \sup_{\eta'} g^n_{\vec{t}}\right) \\ &\le \operatorname{var}_{\eta''}\varphi \cdot 4\tilde{\Theta}^n + \left(\operatorname{var}_{\eta'} g^n_{\vec{t}} + 2 \cdot \sup_{\eta'} g^n_{\vec{t}}\right) \cdot \frac{1}{\ell(\eta'')}\int_{\eta''}|\varphi| \\ &\le \operatorname{var}_{\eta''}\varphi \cdot 4\tilde{\Theta}^n + D \cdot \int_{\eta''}|\varphi| \, , \end{aligned} \tag{5.6}$$

where $D = \sup_{\eta' \in \mathcal{Z}_{n,\vec{t}}} \big[\mathrm{var}_{\,\eta'} g_{\vec{t}}^{n_2} + 2 \cdot \sup_{\eta'} g_{\vec{t}}^{n_2} \big] / \ell_{n_2}$, with ℓ_{n_2} equal to the infimum of the lengths of admissible intervals in $\mathcal{Z}_{n_2,\vec{t}}$. Note that when ϵ tends to zero, ℓ_{n_2} tends to $\inf \ell(\eta)$, for η in $\mathcal{Z}_{n_2}$.

Summing (5.6) over all intervals η'_B, we get for $n = n_2$

$$\mathrm{var}\,(\mathcal{L}_{\vec{t}}^{n}(\varphi \chi_{B'})) \leq 4 \cdot 2^{\mathcal{M}} \cdot \tilde{\Theta}^n \cdot \mathrm{var}\,(\varphi) + 2^{\mathcal{M}} D \cdot |\varphi|_1 \,,$$

and, by increasing $\tilde{\Theta}$ slightly and assuming n_2 is large enough,

$$\begin{aligned} \mathrm{var}\,(\mathcal{L}_{\vec{t}}^{n}(\varphi \chi_{B'})) &\leq \sum_{\eta' \in \mathcal{Z}_{n,\vec{t}}} \mathrm{var}\,(\mathcal{L}_{\vec{t}}^{n}(\varphi \chi_{\eta'})) \\ &\leq \tilde{\Theta}^n \cdot \mathrm{var}\,(\varphi) + 2^{\mathcal{M}} D \cdot |\varphi|_1 \,. \end{aligned} \tag{5.7}$$

If $n > n_2$, write $n = q \cdot n_2 + r$ with $r < n_2$. If a vector $\vec{t}$ of length $2n_2$ is the concatenation of two vectors $\vec{u}$ and $\vec{v}$ of length n_2, and ξ, ζ are the unique intervals in $\mathcal{Z}_{n_2,\vec{u}}$, respectively $\mathcal{Z}_{n_2,\vec{v}}$ such that a given $\eta' \in \mathcal{Z}_{n,\vec{t}}$ is equal to $(f_{\vec{v}}^{n_2}|_{\xi})^{-1}(\xi) \cap \zeta$ then

$$\mathcal{L}_{\vec{t}}^{2n_2}(\varphi \chi'_{\eta}) = \mathcal{L}_{\vec{u}}^{n_2}(\chi_{\xi} \cdot \mathcal{L}_{\vec{v}}^{n_2}(\chi_{\zeta} \cdot \varphi)) \,.$$

In particular

$$\begin{aligned} \sum_{\xi \in \mathcal{Z}_{n_2,\vec{u}}} \mathcal{L}_{\vec{u}}^{n_2}(\chi_{\xi} \cdot \mathcal{L}_{\vec{v}}^{n_2}(\chi_{\zeta} \cdot \varphi)) &\leq \tilde{\Theta}^{n_2} \cdot \mathrm{var}\,(\mathcal{L}_{\vec{v}}^{n_2}(\chi_{\zeta} \cdot \varphi)) + 2^{\mathcal{M}} D \cdot |\mathcal{L}_{\vec{v}}^{n_2}(\chi_{\zeta} \cdot \varphi)|_1 \\ &\leq \tilde{\Theta}^{n_2} \cdot \mathrm{var}\,(\mathcal{L}_{\vec{v}}^{n_2}(\chi_{\zeta} \cdot \varphi)) + 2^{\mathcal{M}} D \cdot \int_{\zeta} |\varphi| \,. \end{aligned}$$

A standard induction argument yields

$$\mathrm{var}\,\mathcal{L}_{\vec{t}}^{n}(\varphi \chi_{B'}) \leq \tilde{\Theta}^n \cdot \mathrm{var}\,\varphi + D' \cdot |\varphi|_1 \,, \tag{5.8}$$

where D' is essentially $2^{\mathcal{M}} D / (1 - \tilde{\Theta})$ (see e.g. Rychlik [1983, Lemma 7, and Proposition 1]).

The problem we have to solve now is the fact that the term $D' \cdot |\varphi|_1$ in (5.8) is not small. To do this, we follow the "balancing" idea suggested to us by Collet [1991] (see also e.g. Young [1992]). We rewrite (5.3) and (5.8) using our new norm $\|\cdot\|_\gamma$

$$\begin{aligned} |\mathcal{L}_{\vec{t}}^{n}(\varphi \chi_{B'})|_1 &\leq c_{n,\epsilon} \cdot (\gamma \cdot \mathrm{var}\,\varphi + |\varphi|_1) \\ \gamma \cdot \mathrm{var}\,(\mathcal{L}_{\vec{t}}^{n}(\varphi \chi_{B'})) &\leq \gamma \cdot \tilde{\Theta}^n \cdot \mathrm{var}\,\varphi + \gamma \cdot D' \cdot |\varphi|_1 \,. \end{aligned}$$

and thus

$$\|\mathcal{L}_{\vec{t}}^n(\varphi\chi_{B'})\|_\gamma \le (c_{n,\epsilon} + \tilde{\Theta}^n + \gamma \cdot D') \cdot \|\varphi\|_\gamma \le (\tilde{\Theta}^n + D' \cdot \gamma) \cdot \|\varphi\|_\gamma \,. \tag{5.9}$$

We now bound the difference $\|\mathcal{L}^n \varphi\chi_{(I\setminus B)} - \mathcal{L}_{\vec{t}}^n(\varphi\chi_{(I\setminus B')})\|$. We first consider the supremum norm to control the L^1 part. Let us fix some point x in $f_{\vec{t}}^n(I \setminus B')$. By assumption, there exist two nonempty lists of intervals $\eta'_{G,j} \subset I \setminus B'$, and $\eta_{G,j} \subset I \setminus B$ ($j = 1, \ldots k(x)$) such that $x \in f^n(\eta_{G,j}) = f_{\vec{t}}^n(\eta'_{G,k})$ for $j = 1, \ldots k(x)$. Fixing j and denoting by $y_{\vec{t}}$, respectively y the unique n-preimage of x in $\eta = \eta_{G,j}$, respectively $\eta' = \eta'_{G,j}$, we have $d(y, y_{\vec{t}}) = c_{n,\epsilon}$ and hence

$$\begin{aligned}
\mathcal{L}_{\vec{t}}^n(\varphi\chi_{\eta'})(x) &= \varphi(y_{\vec{t}})g(y_{\vec{t}}) \ldots g(f^{n-1}y_{\vec{t}}) \\
&\le (\varphi(y) + \mathrm{var}_{\,\eta\cup\eta'}\varphi) \cdot (g^n(y) + c_{n,\epsilon}) \\
&\le \mathcal{L}^n(\varphi\chi_\eta)(x) + \tilde{\Theta}^n \cdot \mathrm{var}_{\,\eta\cup\eta'}\varphi + c_{n,\epsilon} \cdot (\mathrm{var}_{\,\eta\cup\eta'}\varphi + \sup|\varphi|) \,.
\end{aligned} \tag{5.10}$$

We have an analogous lower bound. Summing over j and using again (5.5), we get:

$$\begin{aligned}
|\mathcal{L}_{\vec{t}}^n(\varphi\chi_{(I\setminus B')}) - \mathcal{L}^n(\varphi\chi_{(I\setminus B)})|_1 &\le \mathrm{const}|\mathcal{L}_{\vec{t}}^n(\varphi\chi_{(I\setminus B')})(x) - \mathcal{L}^n(\varphi\chi_{(I\setminus B)})(x)| \\
&\le \tilde{\Theta}^n \mathrm{var}\,\varphi + c_{n,\epsilon}|\varphi|_1 \,.
\end{aligned} \tag{5.11}$$

The "trimming" was not really needed for the bound (5.11) on the L^1-norm since $f^n(B) \cup f_{\vec{t}}^n(B')$ has a measure tending to zero as ϵ tends to zero, but it will be crucial for the next bound.

Let us consider a pair $(\eta, \eta') = (\eta_{G,j}, \eta'_{G,j})$ as above. We can assume that $\eta' \to \eta$ as $\epsilon \to 0$. If ϵ is small enough, the union $\eta \cup \eta'$ is thus an interval, which intersects at most two other such intervals $\eta_{G,k} \cup \eta'_{G,k}$. Defining the bijection $\Psi : \eta' \to \eta$ by $\Psi(y_{\vec{t}}) = y$, we obtain

$$
\begin{aligned}
\operatorname{var}(\mathcal{L}^n_{\vec{t}}(\varphi\chi_{\eta'}) - \mathcal{L}^n(\varphi\chi_\eta)) &= \operatorname{var}(g^n_{\vec{t}}\varphi\chi_{\eta'} - (g^n\varphi)\circ\Psi\chi_{\eta'}) \\
&\le \operatorname{var}((g^n_{\vec{t}}\varphi\chi_{\eta'} - g^n_{\vec{t}}(\varphi\circ\Psi)\chi_{\eta'}) + \operatorname{var}(g^n_{\vec{t}}(\varphi\circ\Psi)\chi_{\eta'} - (g^n\varphi)\circ\Psi\chi_{\eta'}) \\
&\le \operatorname{var}_{\eta'}(g^n_{\vec{t}}(\varphi - \varphi\circ\Psi)) + \operatorname{var}_{\eta=\Psi\eta'}(\varphi(g^n_{\vec{t}}\circ\Psi^{-1} - g^n)) \\
&\quad + 2\sup_{\eta'}(g^n_{\vec{t}}(\varphi-\varphi\circ\Psi)) + 2\sup_{\eta}(\varphi(g^n_{\vec{t}}\circ\Psi^{-1} - g^n)) \\
&\le \sup_{\eta'} g^n_{\vec{t}}\cdot\operatorname{var}_{\eta'}(\varphi-\varphi\circ\Psi) + \operatorname{var}_{\eta'} g^n_{\vec{t}}\cdot\sup_{\eta'}|\varphi-\varphi\circ\Psi| \\
&\quad + \sup_{\eta}|\varphi|\cdot\operatorname{var}_{\eta'}(g^n_{\vec{t}} - g^n\circ\Psi) + \operatorname{var}_{\eta}\varphi\cdot\sup_{\eta'}|g^n_{\vec{t}} - g^n\circ\Psi| \\
&\quad + 2\sup_{\eta'} g^n_{\vec{t}}\cdot\sup_{\eta'}|\varphi-\varphi\circ\Psi| + 2\sup_{\eta}|\varphi|\cdot\sup_{\eta'}|g^n_{\vec{t}} - g^n\circ\Psi| \\
&\le 2\tilde{\Theta}^n\cdot\operatorname{var}_{\eta'\cup\eta}(\varphi) + \tilde{\Theta}^n\cdot\operatorname{var}_{\eta\cup\eta'}\varphi + \sup_{\eta}|\varphi|\cdot c_{n,\epsilon} + \operatorname{var}_{\eta}\varphi\cdot 2\tilde{\Theta}^n \\
&\quad + 2\tilde{\Theta}^n\cdot\operatorname{var}_{\eta\cup\eta'}\varphi + 2\sup_{\eta}|\varphi|\cdot c_{n,\epsilon}\,,
\end{aligned}
\tag{5.12}
$$

where we have used that f is $\mathcal{C}^2$ in the last inequality to get $\operatorname{var}_{\eta'}(g^n_{\vec{t}} - g^n\circ\Psi) \le c_{n,\epsilon}$.

Summing the above inequality over all $(\eta'_{G,j}, \eta_{G,j})$, we get

$$
\operatorname{var}(\mathcal{L}^n(\varphi\chi_{I\setminus B}) - \mathcal{L}^n_{\vec{t}}(\varphi\chi_{I\setminus B'})) \le \tilde{\Theta}^n\cdot(\operatorname{var}\varphi + |\varphi|_1)\,. \tag{5.13}
$$

¿From (5.11) and (5.13) we find:

$$
\begin{aligned}
\|\mathcal{L}^n(\varphi\chi_{I\setminus B}) - \mathcal{L}^n_{\vec{t}}(\varphi\chi_{I\setminus B'})\|_\gamma &\le \gamma^{-1}\cdot\tilde{\Theta}^n\cdot(\operatorname{var}\varphi + |\varphi|_1) \\
&\le \gamma^{-1}\cdot\tilde{\Theta}^n\cdot\|\varphi\|_\gamma\,.
\end{aligned}
\tag{5.14}
$$

If $\gamma = \Lambda^n$ then, since $\Lambda^2 > \tilde{\Theta}$,

$$
D'\cdot\gamma + \tilde{\Theta}^n + \gamma^{-1}\cdot\tilde{\Theta}^n \le C\cdot\Lambda^n\,.
$$

Integrating (5.14) and (5.9), together with the analogue of (5.9) for the unperturbed operator, over the t_i ends the proof of Lemma 9. $\square$

Lemma 9'. *Assume that $\Lambda^2 > (3/2) \cdot \Theta$ if θ is symmetric, and $\Lambda^2 > 2 \cdot \Theta$ otherwise. There exists $C > 0$ and $N_0 \in \mathbb{Z}^+$ such that, for each $n \geq N_0$ there exists $\epsilon(n) > 0$ such that, for each $\epsilon < \epsilon(n)$,*

$$\|\mathcal{L}_\epsilon^n - \mathcal{L}^n\|_{\Lambda^n} < C \cdot \Lambda^n \, ,$$

Proof of Lemma 9'. We shall follow the proof of Lemma 9, noting only the modifications which are necessary when $\mathcal{M} = \infty$.

Assume first that $\Lambda^2 > 2 \cdot \tilde{\Theta}$. We see that the only important change in occurs when we sum (5.6) over the intervals η'_B. Since each good interval η'_G has at most $2^{n-1}+1$ co-respondents, the sum yields for $n = n_2$:

$$\mathrm{var}\,(\mathcal{L}_{\vec{t}}^n(\varphi\chi_{B'})) \leq 4 \cdot 2^n \cdot \tilde{\Theta}^n \cdot \mathrm{var}\,(\varphi) + 2^n D \cdot |\varphi|_1 \, .$$

For general $n = q \cdot n_2 + r$, the same induction argument as in the proof of Lemma 9 allows us to replace Inequality (5.8) by

$$\mathrm{var}\,\mathcal{L}_{\vec{t}}^n(\varphi\chi_{B'}) \leq (2 \cdot \tilde{\Theta})^n \cdot \mathrm{var}\,\varphi + 2^n \cdot 2D \cdot |\varphi|_1 \, .$$

Inequality (5.9) hence becomes

$$\|\mathcal{L}_{\vec{t}}^n(\varphi\chi_{B'})\|_\gamma \leq (c_{n,\epsilon} + (2 \cdot \tilde{\Theta})^n + \gamma \cdot 2^n \cdot D) \cdot \|\varphi\|_\gamma \leq ((2 \cdot \tilde{\Theta})^n + \gamma \cdot 2^n D) \cdot \|\varphi\|_\gamma \, .$$

Inequality (5.14) does not have to be changed and we are done.

Assume now that each θ_ϵ is symmetric, and that $\Lambda^2 > (3/2) \cdot \Theta$. Again inequality (5.14) does not have to be changed, and it suffices to get a bound replacing (5.9). Let η'_G be a trimmed admissible interval for $f_{\vec{t}}^n$ which is associated with $\eta_G \subset \eta \in \mathcal{Z}_n$, where a boundary point b of η is periodic. We claim that there exists a sequence $S = \{s_j\}_{j=1,\ldots n}$ of signs $s_j = \in \{+,-\}$ such that η'_G has at most $2^{k(S)}$ nonadmissible co-respondents η'_B, where $0 \leq k(S) \leq n$ is the numbers of coordinates t_i of $\vec{t}$ such that the sign of $t_j = s_j$. Indeed, take s_j to be $+$ or $-$, depending on whether the j^{th} iterate of b is a local maximum or a local minimum respectively for f^n. (For example, in the map of Figure 1, the sequence of signs is $s_j = +$ for all j.)

We first sum (5.6) over the bad intervals η'_B for which $k(\eta'_B)$ is equal to some fixed k and call this partial sum A_k. Since θ_ϵ is symmetric, we have

$$\int \theta_\epsilon(t_1)\ldots\theta_\epsilon(t_n)A_k \le \binom{n}{k}\frac{2^k}{2^n}\cdot\left(\tilde\Theta^n\cdot\operatorname{var}\varphi + D\cdot|\varphi|_1\right),$$

hence, using $\sum_{k=1}^n\binom{n}{k}2^k = 3^n-1$,

$$\operatorname{var}\mathcal{L}^n_\epsilon(\varphi\chi_{B'}) \le \sum_k \int \theta_\epsilon(t_1)\ldots\theta_\epsilon(t_n)A_k \le ((3/2)\cdot\tilde\Theta)^n\cdot\operatorname{var}\varphi + (3/2)^n D\cdot|\varphi|_1\,.$$

We thus obtain

$$\begin{aligned}\|\mathcal{L}^n_\epsilon(\varphi\chi_{B'})\|_\gamma &\le (c_{n,\epsilon} + ((3/2)\cdot\tilde\Theta)^n + \gamma\cdot(3/2)^n\cdot D')\cdot\|\varphi\|_\gamma\\ &\le \left[\left((3/2)\cdot\tilde\Theta\right)^n + \gamma\cdot(3/2)^n D'\right]\cdot\|\varphi\|_\gamma\,,\end{aligned}$$

which yields the claim. $\square$

We have implicitly used the following inequality in the proofs of Lemma 9 and Lemma 9': assume that $\psi(x,t)$ is a function of two variables such that the function $t\mapsto\theta_\epsilon(t)\psi(x,t)$ is in $L^1(dm)$ for each fixed x, then

$$\begin{aligned}|\int dt\,\theta_\epsilon(t)\psi(\cdot,t)|_1 &\le \int dt\,\theta_\epsilon(t)|\psi(\cdot,t)|_1\\ \operatorname{var}_x\Big(\int dt\,\theta_\epsilon(t)\psi(x,t)\Big) &\le \int dt\,\theta_\epsilon(t)\operatorname{var}_x\psi(x,t)\,.\end{aligned}$$

As in the first two models, we have not used in the proofs the expanding condition as stated, but only the slightly weaker assumption $\Theta<1$.

E. Perturbation lemmas for abstract operators.

Because of the need to introduce the norms $\|\cdot\|_\gamma$, we need a slightly refined version of Section 2. Again, $(X,\|\cdot\|)$ is a complex Banach space, and $\{T_\epsilon,\epsilon\ge 0\}$ is a family of bounded linear operators on X. We make exactly the same assumption about T_0: there exist two real numbers $0<\kappa_1<\kappa_0\le 1$ such that the spectrum of T_0 decomposes as $\Sigma_0\cup\Sigma_1$ where

$$\begin{aligned}\kappa_0 &= \inf\{|z|\,:\,z\in\Sigma_0\}\\ \kappa_1 &= \sup\{|z|\,:\,z\in\Sigma_1\}\,,\end{aligned}\tag{A'.1}$$

and use the notations X_i, π_i, $i = 0, 1$ of Section 2.

We assume further that there is another norm $|\cdot|$ on X such that $|x| \le \|x\|$ for all x, and a family of norms $\|\cdot\|_\gamma$, with $0 < \gamma \le 1$ with

$$\|\cdot\|_\gamma = \gamma\|\cdot\| + (1-\gamma)|\cdot|\,.$$

(In particular $\gamma\|\cdot\| \le \|\cdot\|_\gamma \le \|\cdot\|$ and $|\cdot| \le \|\cdot\|_\gamma$.)

Condition (A.2) is replaced by the assumption that there exists κ with $(\kappa_1/\kappa_0) < \kappa < \kappa_0$ such that for each large enough $N \in \mathbb{Z}^+$ there exists $\epsilon(n)$ such that for all $0 < \epsilon < \epsilon(n)$

$$\|T_\epsilon^N - T_0^N\|_{\kappa^N} \le \kappa^N\,. \tag{A'.2}$$

We shall need two sublemmas:

Sublemma 5. *Assume that*

$$dim\, X_0 < \infty\,. \tag{A'.3}$$

Then for any $\kappa_0' < \kappa_0$, there exists N_0 such that for all $n \ge N_0$, any $0 < \gamma \le 1$, and any $x \in X_0$

$$\|T_0^n x\|_\gamma \ge (\kappa_0')^n \|x\|_\gamma\,.$$

Proof of Sublemma 5. First note that there exists $N_0 \ge 1$ such that for all $n \ge N_0$, and any $x \in X_0$

$$|T_0^n x| \ge (\kappa_0')^n |x|\,.$$

Indeed, the vector space X_0 is invariant under T_0. Since it is finite-dimensional, it is closed in the $|\cdot|$-norm, and the spectrum of $T_0|_{X_0}$ in the $|\cdot|$-norm is a finite set Σ of eigenvalues. Obviously, $\Sigma = \Sigma_0$ (an eigenfunction of T_0 in X_0 is of bounded $\|\cdot\|$-norm).

It now suffices to observe that for $x \in X_0$, $0 < \gamma \le 1$, and $n \ge N_0$

$$\|T_0^n x\|_\gamma \ge |T_0^n x| \ge (\kappa_0')^n |x| \ge \text{const} \cdot (\kappa_0')^n \|x\| \ge \text{const} \cdot (\kappa_0')^n \|x\|_\gamma\,,$$

where we have used the fact that the norms $|\cdot|$ and $\|\cdot\|$ are equivalent on the finite-dimensional subspace X_0. $\square$

Sublemma 6. *If* (A'.3) *holds, then there exists a constant* C *such that for any* $0 < \gamma \leq 1$*, we have* $\|\pi_0\|_\gamma \leq C$ *and* $\|\pi_1\|_\gamma \leq C+1$.

Proof of Sublemma 6. For $x \in X$, we have

$$\|\pi_0 x\|_\gamma \leq \|\pi_0 x\| \leq \text{const}|\pi_0| \cdot |x| \leq \text{const}|\pi_0| \cdot \|x\|_\gamma \,,$$

where we have used the fact that the norms $|\cdot|$ and $\|\cdot\|$ are equivalent on the finite-dimensional space X_0. To finish, observe that $\pi_0+\pi_1 = I$ so that $\|\pi_1\|_\gamma \leq \|\pi_0\|_\gamma+1$. $\square$

We can now prove:

Lemma 1'. *Assume that* (A'.1)–(A'.2) *hold, then the conclusion of Lemma 1 from Section 2 is true, except that* κ' *cannot be made arbitrarily close to* κ*.*

Proof of Lemma 1'. Since any $\|\cdot\|_\gamma$-norm is equivalent to $\|\cdot\|$, we can follow the proof of Lemma 1, replacing each occurrence of $\|\cdot\|$ by $\|\cdot\|_{\kappa^N}$ and noting which changes have to be made.

Let $\kappa_1 < \kappa_1' < \kappa < \kappa' < \kappa_0' < \kappa_0$ and let $\kappa' < |\lambda| < \kappa_0'$. We will show that $\lambda \notin \sigma(T_\epsilon)$ if κ' is close enough to κ_0. By Sublemma 5, for $x \in X_0$ and $\kappa_0' < \hat{\kappa}_0 < \kappa_0$,

$$\|T_0^N x - \lambda x\|_{\kappa^N} \geq \text{const} \cdot (\hat{\kappa}_0)^N \|x\|_{\kappa^N} \,.$$

Since $\|\cdot\| \leq \kappa^{-N} \|\cdot\|_{\kappa^N}$, for $x \in X_1$

$$\begin{aligned}\|T_0^N x - \lambda x\|_{\kappa^N} &\geq \big(-(\frac{\kappa_1'}{\kappa})^N + (\kappa')^N\big)\|x\|_{\kappa^N} \\ &\geq \text{const} \cdot (\kappa')^N \|x\|_{\kappa^N} \,,\end{aligned}$$

if κ' is close enough to κ_0 (use the inequality $\kappa_1 < \kappa \cdot \kappa_0$).

Hence, for large enough N,

$$\|R(T_0^N, \lambda^N)\|_{\kappa^N} \leq \frac{\text{const}\big(\|\pi_0\|_{\kappa^N} + \|\pi_1\|_{\kappa^N}\big)}{(\kappa')^N} \leq \frac{1}{\kappa^N} \,,$$

where we have used Sublemma 6 in the last inequality. $\square$

Lemma 2 from Section 2 holds in the present setting, with convergence in the sense of the $\|\cdot\|_{\kappa^N}$-norm (i.e., for any $\delta > 0$ there are $N \in \mathbb{Z}^+$ and $\epsilon(N)$ such that, for each $\epsilon < \epsilon(N)$, $\|\pi_0 - \pi_0^\epsilon\|_{\kappa^N} < \delta$), and the same proof.

For $n \geq 1$, we use the notation $C_n(\epsilon)$ from Section 2, replacing the norm $\|\cdot\|$ with $\|\cdot\|_\gamma$ (we again have $C_n(\epsilon) \leq \kappa^n$ for small enough ϵ) and we also write

$$C_n^*(\epsilon) := \sup_{\substack{x \in X_0 \\ x \neq 0}} \frac{|T_\epsilon^n x - T_0^n x|}{|x|} \,.$$

Lemma 3'. *Assume that* (A'.1)-(A'.3) *hold and that* $|T_\epsilon|$ *is uniformly bounded. Let* d *denote the maximum algebraic multiplicity of* $z \in \sigma(T_0|_{X_0})$ *and let* κ' *and* κ_0' *be the constants from Lemma 1'. Fix* N *large and consider* $\epsilon < \epsilon(N)$*. Then*

(1) *HD-distance*$(\sigma(T_0|_{X_0}), \sigma(T_\epsilon|_{X_0^\epsilon})) \leq$ *const* $\cdot\, (C_1^*(\epsilon) + \frac{C_N(\epsilon)}{{\kappa_0'}^N})^{1/d}$.

(2) *If* $\hat{x}_0 \in X_0$ *is an eigenvector for* T_0 *with* $T_0 \hat{x}_0 = \nu_0 \hat{x}_0$*, then* T_ϵ *has an eigenvector* $\hat{x}_0^\epsilon \in X_0^\epsilon$ *with eigenvalue* ν_0^ϵ *which is const* $\cdot\, (C_1^*(\epsilon) + \frac{C_N(\epsilon)}{{\kappa_0'}^N})^{1/d}$*-near* ν_0 *such that*

$$|\hat{x}_0^\epsilon - \hat{x}_0| \leq \mathit{const} \cdot (C_1^*(\epsilon) + \frac{C_N(\epsilon)}{{\kappa_0'}^N})^{1/d} \,.$$

Proof of Lemma 3'. As in Lemma 3, we can show that $X_0^\epsilon = \mathrm{graph}(S_\epsilon)$ for a linear $S_\epsilon : X_0 \to X_1$ with $\|S_\epsilon\|_{\kappa^N} \to 0$ as $\epsilon \to 0$ and $N \to \infty$. To estimate $\|S_\epsilon\|_{\kappa^N}$, use Sublemma 5 again to observe that (2.4) can be replaced by

$$\|S_\epsilon\|_{\kappa^N} \leq \frac{\|\pi_1\|_{\kappa^N} \left(\left((\frac{\kappa_1'}{\kappa})^N + \kappa^N \right) \|S_\epsilon\|_{\kappa^N} + C_N(\epsilon) \right)}{(\kappa_0')^N - \|\pi_0\|_{\kappa^N}(1 + \|S_\epsilon\|_{\kappa^N})\kappa^N} \tag{5.15}$$

Hence, $\|S_\epsilon\|_{\kappa^N} \leq \mathrm{const} \cdot C_N(\epsilon)/(\kappa_0')^N$ (where we have used again $\kappa_1 < \kappa \cdot \kappa_0$ and Sublemma 6).

Define again $\hat{T}_\epsilon : X_0 \to X_0$ by

$$\hat{T}_\epsilon(x) = \pi_0 \circ T_\epsilon(x, S_\epsilon x) \,.$$

For $x \in X_0$, we have

$$
\begin{aligned}
|\hat{T}_\epsilon x - T_0 x| &\le |\pi_0| \cdot (|T_\epsilon x - T_0 x| + |T_\epsilon S_\epsilon x|) \\
&\le |\pi_0| \cdot (C_1^*(\epsilon) + |T_\epsilon| \cdot \frac{C_N(\epsilon)}{{\kappa_0'}^N}) \cdot \|x\|_{\kappa^N} \\
&\le |\pi_0| \cdot (C_1^*(\epsilon) + |T_\epsilon| \cdot \frac{C_N(\epsilon)}{{\kappa_0'}^N}) \cdot \|x\| \\
&\le \text{const} \cdot |\pi_0| \cdot (C_1^*(\epsilon) + |T_\epsilon| \cdot \frac{C_N(\epsilon)}{{\kappa_0'}^N}) \cdot |x| \,.
\end{aligned}
$$

There is a similar bound for $|\pi_1 \circ T_\epsilon(x, S_\epsilon x) - \pi_1 T_0 x|$ with $x \in X_0$. We finish as in Section 2. $\square$

Proof of Theorem 3 and Theorem 3'. By Lemma 9 (respectively 9') and Assumption A (respectively B), the Assumptions (A'.1)–(A'.3) hold for $T_0 = \mathcal{L}$, $T_\epsilon = \mathcal{L}_\epsilon$, $X = BV$, $\|\cdot\|$ the BV-norm, $\|\cdot\|_\gamma$ the balanced norms, $|\cdot|$ the L^1-norm, κ_0, and κ_1 as in Assumptions A or B, and $\kappa < \kappa_0$ with

$$
\kappa^2 > \begin{cases} \max(\kappa_1, \Theta)\,, & \text{if there are no periodic turning points,} \\ \max(\kappa_1, 3/2 \cdot \Theta)\,, & \text{if each } \theta_\epsilon \text{ is symmetric,} \\ \max(\kappa_1, 2 \cdot \Theta)\,, & \text{otherwise.} \end{cases}
$$

(The fact that the L^1-norm is strictly speaking only a norm when one quotients out functions of bounded variation φ for which $|\varphi|_1 = 0$ is not a problem, see Proposition 1 in Baladi–Keller [1990].) The norm of $\mathcal{L}_\epsilon : L^1 \to L^1$ is equal to 1, and it follows from Lemmas 9 and 9' that $\mathcal{L}_\epsilon$ is quasi-compact so that $\rho_\epsilon \in BV$ (see e.g. Keller [1982, p. 315]). Theorem 3 and Theorem 3' hence follow from Lemma 3', Lemma 8 and the results stated in Sections 5.A and 5.B. $\square$

References

1. V. Baladi and G. Keller, *Zeta functions and transfer operators for piecewise monotone transformations*, Comm. Math. Phys. **127** (1990), 459–477.
2. P. Baxendale, *Brownian motions in the diffeomorphism group*, Compositio Math. **53** (1984), 19–50.
3. M. Benedicks and L.-S. Young, *Absolutely continuous invariant measures and random perturbations for certain one-dimensional maps*, Ergodic Theory Dynamical Systems **12** (1992), 13–37.
4. P. Collet, *Ergodic properties of some unimodal mappings of the interval*, Preprint Mittag-Leffler (1984).

5. ______, *Some ergodic properties of maps of the interval*, Dynamical Systems and Frustrated Systems (R. Bamon, J.-M. Gambaudo and S. Martinez, eds.), 1991 (to appear).
6. P. Collet and S. Isola, *On the essential spectrum of the transfer operator for expanding Markov maps*, Comm. Math. Phys. **139** (1991), 551–557.
7. E.M. Coven, I. Kan and J.A. Yorke, *Pseudo-orbit shadowing in the family of tent maps*, Trans. Amer. Math. Soc. **308** (1988), 227–241.
8. J. Franks, *Manifolds of $\mathcal{C}^r$ mappings and applications to differentiable dynamical systems*, Studies in Analysis, Adv. Math. Suppl. Stud. **4** (1979), 271–291.
9. F. Hofbauer and G. Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations*, Math. Z. **180** (1982), 119–140.
10. T. Kato, *Perturbation theory for Linear Operators (Second edition)*, Springer-Verlag, Berlin Heidelberg New York, 1976.
11. G. Keller, *Stochastic stability in some chaotic dynamical systems*, Monatsh. Math. **94** (1982), 313–333.
12. ______, *On the rate of convergence to equilibrium in one-dimensional systems*, Comm. Math. Phys. **96** (1984), 181–193.
13. Y. Kifer, *On small random perturbations of some smooth dynamical systems*, Math. USSR-Izv. **8** (1974), 1083–1107.
14. ______, *Ergodic Theory of Random Transformations*, Birkhäuser, Boston, Basel, 1986.
15. ______, *Random Perturbations of Dynamical Systems*, Birkhäuser, Boston, Basel, 1988a.
16. ______, *A note on integrability of $\mathcal{C}^r$ norms of stochastic flows and applications*, Stochastic Mechanics and Stochastic Processes, Proc. Conf. Swansea/UK 1986, Springer Verlag (Lecture Notes in Math. 1325), Berlin, 1988b, pp. 125–131.
17. H. Kunita, *Stochastic Flows and Stochastic Differential Equations*, Cambridge University Press, Cambridge, 1990.
18. R. Mañé, *Ergodic Theory and Differentiable Dynamics*, Springer-Verlag, Berlin Heidelberg New York, 1987.
19. D. Mayer, *On a ζ function related to the continued fraction transformation*, Bull. Soc. Math. France **104** (1976), 195–203.
20. M. Pollicott, Private communication (1991).
21. D. Ruelle, *Zeta functions for expanding maps and Anosov flows*, Invent. Math. **34** (1976), 231–242.
22. ______, *Locating resonances for Axiom A dynamical systems*, J. Stat. Phys. **44** (1986), 281–292.
23. ______, *The thermodynamic formalism for expanding maps*, Comm. Math. Phys. **125** (1989), 239–262.
24. ______, *An extension of the theory of Fredholm determinants*, Inst. Hautes Études Sci. Publ. Math. **72** (1990), 175–193.
25. M. Rychlik, *Bounded variation and invariant measures*, Studia Math. **LXXVI** (1983), 69–80.
26. J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, London, 1965.
27. S. Wong, *Some metric properties of piecewise monotonic mappings of the unit interval*, Trans. Amer. Math. Soc. (1978), 493–500.
28. K. Yosida, *Functional Analysis (Sixth Edition)*, Springer-Verlag (Grundlehren der mathematischen Wissenschaften 123), Berlin Heidelberg New York, 1980.
29. L.-S. Young, *Decay of correlations for certain quadratic maps* (1992) (to appear Comm. Math. Phys.).

UMPA, ENS Lyon (CNRS, UMR 128), 46, Allée d'Italie, F-69364 Lyon Cedex, France
E-mail address: baladi@umpa.ens-lyon.fr

Department of Mathematics, UCLA, Los Angeles, CA 90024, USA
E-mail address: lsy@math.ucla.edu