Open access • Journal Article • DOI:10.1007/BF02098487

On the spectra of randomly perturbed expanding maps — Source link

Viviane Baladi, Lai Sang Young

Institutions: École normale supérieure de Lyon, University of Arizona

Published on: 01 Sep 1993 - Communications in Mathematical Physics (Springer-Verlag)

Commun. Math. Phys. 156, 355–385 (1993)

Communications in Mathematical Physics

On the Spectra of Randomly Perturbed Expanding Maps

V. Baladi[1,*] **and L.-S. Young**[2,**]

[1] CNRS, UMR 128, UMPA, ENS Lyon, 46, Allée d'Italie, F-69364 Lyon Cedex, France,
E-mail address: baladi@umpa.ens-lyon.fr
[2] Department of Mathematics, University of Arizona, Tucson, AZ 85721, USA, and
Department of Mathematics, UCLA, Los Angeles, CA 90024, USA
E-mail address: lsy@math.arizona.edu and lsy@math.ucla.edu

Received November 13, 1992; in revised form February 25, 1993

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 the robustness of the spectra of their Perron-Frobenius operators.

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 $\mathscr{L}$ 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 $\mathscr{L}$ between the eigenvalue equal to unity and the "next largest eigenvalue." Corresponding to the noisy situation is a noisy Perron-Frobenius operator $\mathscr{L}_\varepsilon$, which should not be too different from $\mathscr{L}$ for small ε. By standard perturbation arguments for linear operators, the eigenfunction corresponding to the eigenvalue 1 for $\mathscr{L}_\varepsilon$ should be near that for $\mathscr{L}$, proving stochastic stability.

Also, since the "second largest" eigenvalue of $\mathscr{L}$ 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

* 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

gaps exist, this reasoning can be extended to some other eigenvalues of $\mathscr{L}$ (the "resonances" of Ruelle [1986]).

One obvious way to make this heuristic argument rigorous would be to show that $\mathscr{L}_\varepsilon$ converges to $\mathscr{L}$ in the topology of operator norms. That, unfortunately, is almost never true. In general, the relation between $\mathscr{L}_\varepsilon$ and $\mathscr{L}$ depends on the dynamics as well as the function space in question. The purpose of this paper is to examine the nature of this perturbation for the following three models:

Our first model consists of expanding maps of the circle, which we perturb by taking convolutions with a fixed kernel. The function space on which our Perron-Frobenius operators act is the space of $\mathscr{C}^r$ functions. Our second model is a slight generalization of the first: we consider expanding maps of Riemannian manifolds followed by stochastic flows. Our third model consists of piecewise expanding maps of the interval, which we assume to be mixing. The perturbations are the same as those in the first model, but our test functions are only of bounded variation. All three models, when unperturbed, have the exponential mixing property.

For the first two models we prove that $\mathscr{L}_\varepsilon$ converges to $\mathscr{L}$ in a strong enough sense to guarantee the convergence of the spectrum on certain regions of the complex plane. (There is a disk containing the essential spectrum of $\mathscr{L}$ on which we have little control.) The situation in the third model is somewhat more delicate. We have the same results provided we further restrict the domain of convergence. As explained earlier, these convergence results allow us to read off immediately properties such as stochastic stability, robustness of the rate of mixing, etc.

Not all of our results are new. Stochastic stability, particularly in the sense of weak convergence of measures, has been proved for various dynamical systems. See e.g. Kifer [1988a]. Stability in the bounded variation case is first proved in Keller [1982]. Kifer has a result in the opposite direction [1988a]. He proves the collapse of the spectrum of a related unitary operator for hyperbolic toral automorphisms. (This operator has continuous spectrum.) More references will be given later on.

This paper is organized as follows. In Sect. 2 we prove some simple perturbation lemmas for abstract operators. We deal with our three models in Sect. 3, 4, and 5, proving some dynamical lemmas that relate $\mathscr{L}_\varepsilon$ to $\mathscr{L}$. We then obtain our desired conclusions by appealing to the results in Sect. 2. We hope that this method of proof goes beyond the situations considered in the present article.

In a forthcoming paper by the first named author some of the results here will be brought to greater 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 $\mathscr{L}$ on certain regions of the complex plane.

We express our thanks to Pierre Collet and François Ledrappier for very useful conversations. V. Baladi acknowledges the hospitality and financial support of the 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 $\mu_0 = \varrho_0 dm$, 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 $\mathscr{B}$ denote the σ-algebra of Borel sets of M and $\mathscr{P}$ the space of Borel probability measures on M. Recall that a random perturbation of f is a family of Markov chains $\mathscr{X}^\varepsilon$ (with small $\varepsilon > 0$) defined on the measure space $(M, \mathscr{B})$, with transition probabilities $\{P^\varepsilon(x, \cdot)\}$ in $\mathscr{P}$, i.e., $P\{\mathscr{X}^\varepsilon_{n+1} \in E : \mathscr{X}^\varepsilon_n = x\} = P^\varepsilon(x, E)$. We assume that the following conditions are satisfied:

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

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

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

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

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

$$\mu_\varepsilon(E) = \int P^\varepsilon(x, E) \, d\mu_\varepsilon(x) , \quad \forall E \in \mathscr{B} .$$

(For more details, see e.g. Kifer [1988a]. Note that the assumption that $P^\varepsilon(x, \cdot)$ has a density with respect to Lebesgue is not essential for most of the results below.)

We say that (f, μ_0) is *stochastically stable* under the perturbation $\mathscr{X}^\varepsilon$ if μ_ε tends to μ_0 weakly as $\varepsilon \to 0$. 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 $(\mathscr{F}, \|\cdot\|)$ is a Banach space of functions $\varrho: M \to \mathbb{R}$ containing ϱ_0 and ϱ_ε, then we say that (f, μ_0) is *stochastically stable in* $(\mathscr{F}, \|\cdot\|)$ if $\|\varrho_\varepsilon - \varrho_0\|$ tends to zero as $\varepsilon \to 0$. (See e.g. Keller [1982] and Collet [1984] for certain interval maps, with $\mathscr{F} = L^1(dm)$.)

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* $(\mathscr{F}, \|\cdot\|)$ if τ_0 is the smallest number such that the following holds: for each $\tau > \tau_0$ and each pair $\varphi, \psi \in \mathscr{F}$, there exists $C = C(\tau, \|\varphi\|, \|\psi\|)$ such that

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

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

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

$$\left| \int \left(\int \varphi(y) P^\varepsilon_n(x, dy) \right) \cdot \psi(x) d\mu_\varepsilon(x) - \int \varphi \, d\mu_\varepsilon \int \psi \, d\mu_\varepsilon \right| \le C\tau^n , \quad \forall n \ge 1 .$$

We say that the rate of mixing of (f, μ_0) in $\mathscr{F}$ 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 $(\mathscr{F}, \|\cdot\|)$ as above, and for $\varphi \in \mathscr{F}$, we define

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

Or, equivalently, if $\varphi \in \mathscr{F}$ is the density of a signed measure μ on M, then $\mathscr{L}\varphi$ 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 $E \in \mathscr{B}$. We assume that $\mathscr{L}: \mathscr{F} \to \mathscr{F}$ is a well-defined bounded operator, and that $\varrho_0 \in \mathscr{F}$. Then 1 is an eigenvalue of $\mathscr{L}$, 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 $\mathscr{L}$ is quasi-compact, i.e., its essential spectral radius $\operatorname{ess\,sp}(\mathscr{L})$ is strictly less than its spectral radius. In particular, for every $\tau > \operatorname{ess\,sp}(\mathscr{L})$, the set $\sigma(\mathscr{L}) \cap \{z : |z| \geq \tau\}$ 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 it has been shown that the spectrum of $\mathscr{L}$ can be written as $\sigma(\mathscr{L}) = \sigma_0 \cup \{1\}$, where 1 is a simple eigenvalue (i.e. it has a one-dimensional generalized eigenspace) 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

$$\int (\varphi \circ f^n) \cdot \psi \, d\mu_0 = \int \varphi \cdot \mathscr{L}^n(\psi\,\varrho_0)\, dm,$$

we have

$$\left| \int (\varphi \circ f^n)\psi \, d\mu_0 - \int \varphi \, d\mu_0 \int \psi \, d\mu_0 \right| = \left| \int \varphi \left[\mathscr{L}^n(\psi\,\varrho_0) - \left(\int \psi \varrho_0 \, dm \right) \varrho_0 \right] dm \right| .$$

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

$$\begin{aligned} &\leq C \cdot \|\mathscr{L}^n(\psi\varrho_0) - \pi(\psi\varrho_0)\| \\ &\leq C' \cdot \tau^n , \end{aligned}$$

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 eigenspace of 1. Thus we have $\tau_0 = |\sigma_0|$.

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

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

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

(1) Does $\|\varrho_\varepsilon - \varrho_0\| \to 0$?
(2) Does $\tau_\varepsilon \to \tau_0$ (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 $\|\varrho_\varepsilon - \varrho_0\|$ or $|\tau_\varepsilon - \tau_0|$ scale with ε as $\varepsilon \to 0$?

2. Perturbation Lemmas for Abstract Operators

Let $(X, \|\cdot\|)$ be a complex Banach space, and let $\{T_\varepsilon, \varepsilon \geq 0\}$ 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

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

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 $\kappa < \kappa_0$ such that for each sufficiently large $n \in \mathbb{Z}^+$, there exists $\varepsilon(n)$ such that for all $0 < \varepsilon < \varepsilon(n)$,*

$$\|T_\varepsilon^n - T_0^n\| \leq \kappa^n. \tag{A.2}$$

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

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

such that if

$$\kappa_1^\varepsilon := \sup\{|z| : z \in \Sigma_1^\varepsilon\} \quad \text{and} \quad \kappa_0^\varepsilon := \inf\{|z| : z \in \Sigma_0^\varepsilon\},$$

then $\kappa_1^\varepsilon < \kappa_0^\varepsilon$.

It will become clear later on that (A.2) agrees with the nature of our perturbations. Note that we do not assume that $T_\varepsilon^n x$ converges to $T_0^n x$ as $\varepsilon \to 0$ for fixed n and/or x, nor do we assume that for fixed ε we know anything about $\|T_\varepsilon^n - T_0^n\|$ for all large n.

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

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

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

$$x \in X_0 \Rightarrow \|T_0^N x\| \geq (\kappa_0'')^N \|x\|,$$
$$x \in X_1 \Rightarrow \|T_0^N x\| \leq (\kappa_1')^N \|x\|.$$

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

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

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

Assuming $\|T_\varepsilon^N - T_0^N\| < \kappa^N$, 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 $x \in X$, $\|x\| = 1$,

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

so that it suffices to bound $\|R(T_0^N, \lambda^N)|_{X_i}\|$, $i = 0, 1$.

For $x \in X_0$, we have

$$\begin{aligned}\|T_0^N x - \lambda^N x\| &\geq \|T_0^N x\| - |\lambda|^N \|x\| \\ &\geq ((\kappa_0'')^N - (\kappa_0')^N)\,\|x\| \\ &\geq C \cdot (\kappa_0'')^N \|x\|\,,\end{aligned}$$

where C is a constant depending only on κ_0' and κ_0''. This gives

$$\|R(T_0^N, \lambda^N)|_{X_0}\| \leq \frac{1}{C(\kappa_0'')^N}\,.$$

Similarly, for $x \in X_1$, we have

$$\|T_0^N x - \lambda^N x\| \geq ((\kappa')^N - (\kappa_1')^N)\,\|x\|\,,$$

proving

$$\|R(T_0^N, \lambda^N)|_{X_1}\| \leq \frac{1}{C(\kappa')^N}\,.$$

Hence, for large enough N,

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

Define

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

Note that $\kappa_1^\varepsilon \leq \kappa'$, which can be made arbitrarily near $\max(\kappa, \kappa_1)$ by choosing ε small.

Let $\pi_0^\varepsilon : X_0^\varepsilon \oplus X_1^\varepsilon \to X_0^\varepsilon$ 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* $\|\pi_0 - \pi_0^\varepsilon\| \to 0$ *as* $\varepsilon \to 0$.

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 $\varepsilon < \varepsilon(N)$.

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 $(\Sigma_0^\varepsilon)^N$ are contained in the annular region bounded by C, and we have

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

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

$$\begin{aligned}\|\pi_0 - \pi_0^\varepsilon\| &\leq \frac{1}{2\pi} \int\limits_C \|R(T_0^N, \lambda) - R(T_\varepsilon^N, \lambda)\| d\lambda \\ &\leq \frac{1}{2\pi} \cdot l(B_{\hat\kappa^N}) \cdot \max_{\lambda \in B_{\hat\kappa^N}} \|R(T_0^N, \lambda) - R(T_\varepsilon^N, \lambda)\| \\ &\quad + \text{the corresponding term for } B_{r_0^N} \\ &= : (1) + (2)\,.\end{aligned} \tag{2.3}$$

Using (2.1) we have

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

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

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

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

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

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

Lemma 3. *Assume that* (A.1)–(A.2) *hold, that* $\|T_\varepsilon\|$ *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 $\varepsilon<\varepsilon(N)$:

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

(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 $\hat x_0^\varepsilon\in X_0^\varepsilon$ with eigenvalue ν_0^ε which is* $\text{const}\cdot\left(C_1(\varepsilon)+\frac{C_N(\varepsilon)}{\kappa_0'^N}\right)^{1/d}$*-near ν_0 such that*

$$\|\hat x_0^\varepsilon-\hat x_0\|\le \text{const}\cdot\left(C_1(\varepsilon)+\frac{C_N(\varepsilon)}{\kappa_0'^N}\right)^{1/d}.$$

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

Proof of Lemma 3. First we show that $X_0^\varepsilon=\text{graph}(S_\varepsilon)$ for some linear $S_\varepsilon:X_0\to X_1$ with $\|S_\varepsilon\|\to 0$ as $\varepsilon\to 0$. To see this, consider ε small and let $x\in X_0^\varepsilon$. Since $\|x-\pi_0 x\|\le\|\pi_0^\varepsilon-\pi_0\|\,\|x\|$, 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 $x,x'\in X_0^\varepsilon$ and $\pi_0x=\pi_0x'$ then $x=x'$.

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

$$\|S_\varepsilon\|\le\frac{\|\pi_1T_\varepsilon^N(x_0,S_\varepsilon x_0)\|}{\|\pi_0T_\varepsilon^N(x_0,S_\varepsilon x_0)\|}\,.$$

This is

$$\leq \frac{\|\pi_1\| \left(((\kappa_1')^N + \kappa^N) \|S_\varepsilon\| + C_N(\varepsilon) \right)}{(\kappa_0')^N - \|\pi_0\| (1 + \|S_\varepsilon\|) \cdot \kappa^N} . \tag{2.4}$$

from which we see that

$$\|S_\varepsilon\| \leq \text{const} \frac{C_N(\varepsilon)}{(\kappa_0')^N} .$$

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

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

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

$$\begin{aligned} \|\hat{T}_\varepsilon x - T_0 x\| &\leq \|\pi_0\| \cdot (\|T_\varepsilon x - T_0 x\| + \|T_\varepsilon S_\varepsilon x\|) \\ &\leq \text{const} \cdot \left(C_1(\varepsilon) + \|T_\varepsilon\| \cdot \frac{C_N(\varepsilon)}{\kappa_0'^N} \right). \end{aligned}$$

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

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 $\mathscr{C}^r$ transformation of S^1 ($2 \leq r < \infty$) which is expanding, i.e., $|f'| \geq \lambda > 1$. 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 $\mathscr{F} = \mathscr{C}^{r-1}(S^1)$ and let $\|\cdot\|$ be the usual $\mathscr{C}^{r-1}$-norm. Let $\mathscr{L} : \mathscr{F} \to \mathscr{F}$ be the Perron-Frobenius operator associated with f:

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

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

We remark that if the map f is $\mathscr{C}^\infty$ or $\mathscr{C}^\omega$, we can let $\mathscr{L}$ act on the Fréchet space $\mathscr{C}^\infty(S^1)$ of $\mathscr{C}^\infty$ functions, respectively the Banach space $\mathscr{C}^\omega(S^1)$ of real analytic functions endowed with the supremum norm. Using the fact (Ruelle [1989]) that, for a $\mathscr{C}^r$ map, the eigenfunctions of $\mathscr{L}$ acting on $\mathscr{C}^{r'}$ for $1 \leq r' < r - 1$ are all elements of $\mathscr{C}^{r-1}(S^1)$, it makes sense to speak of the eigenvalues of $\mathscr{L}$ when acting on $\mathscr{C}^\infty(S^1)$, even though $\mathscr{C}^\infty(S^1)$ is not a Banach space. In particular, one can view $\mathscr{L} : \mathscr{C}^\infty(S^1) \to \mathscr{C}^\infty(S^1)$ as a "compact" operator. If $r = \omega$, the operator $\mathscr{L}$ 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 $\mathscr{C}^r(S^1)$, 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.

B. *Type of Perturbation: Convolutions*

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

$$\theta_\varepsilon \ge 0\,, \quad \operatorname{supp} \theta_\varepsilon \subset [-\varepsilon, \varepsilon]\,, \quad \text{and} \quad \int \theta_\varepsilon = 1\,.$$

Consider the random perturbation $\mathscr{X}^\varepsilon$, where the transition probabilities $P^\varepsilon(x, dy)$ have densities $\theta_\varepsilon(y - fx)$. (Note that the density depends only on the difference $y - fx$.) Equivalently, using Fubini's Theorem, 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, Chap. IV]).

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

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

Analogous operators have been used by Keller [1982, Sect. 5] and Collet [1984] among others. The operator $\mathscr{L}_\varepsilon$ is clearly linear and bounded on $\mathscr{C}^{r-1}(S^1)$. Also, it is quasi-compact and the density ϱ_ε is in $\mathscr{C}^{r-1}$ (Ruelle [1990]).

If we had made the additional assumption that θ_ε is $\mathscr{C}^{r-1}$, then $\mathscr{L}_e$ would be a compact operator on $\mathscr{C}^{r-1}(S^1)$. This follows from the fact that a kernel operator

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

with $\mathscr{C}^0$ kernel $K(\cdot, \cdot)$ is compact (see e.g. Yosida [1980, p. 277]).

C. Statement of our Results

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

Theorem 1. *Let $f : S^1 \to S^1$ be a $\mathscr{C}^r$ expanding map ($r \geq 2$) of the circle as defined in Sect. 3.A, with expanding constant λ, and let $\mu_0 = \varrho_0\, dm$ be its unique absolutely continuous invariant probability measure. Let $\mathscr{X}^\varepsilon$ be a small random perturbation of f of the type described in Sect. 3.B, with invariant measure $\mu_\varepsilon = \varrho_\varepsilon\, dm$. Then:*
(1) The dynamical system (f, μ_0) is stochastically stable under $\mathscr{X}^\varepsilon$ in the space of $\mathscr{C}^{r-1}$ functions, i.e., $\|\varrho_\varepsilon - \varrho_0\|_{r-1}$ tends to 0 as $\varepsilon \to 0$. Moreover, we have $\|\varrho_\varepsilon - \varrho_0\|_{r-2} = O(\varepsilon)$.

Let τ_0 and τ_ε be the rates of decay of correlations for f and $\mathscr{X}^\varepsilon$ respectively, in the space of $\mathscr{C}^{r-1}$ functions.
(2) If $\tau_0 > \lambda^{-(r-1)}$, then the rate of mixing is robust, i.e., $\tau_\varepsilon \to \tau_0$ as $\varepsilon \to 0$. Furthermore, if $\tau_0 > \lambda^{-(r-2)}$ then $|\tau_\varepsilon - \tau_0| = O(\varepsilon^{1/d})$ for some integer $d \geq 1$.

We show in fact that
(3) For each $\delta > 0$, the spectrum of $\mathscr{L}_\varepsilon$ restricted to $\{|z| > \lambda^{-(r-1)} + \delta\}$, converges to that of $\mathscr{L}$ (restricted to the same domain) as $\varepsilon \to 0$.

The proofs below yield the same results for small *deterministic* perturbations by translations (i.e., maps $f^\varepsilon = f + t$ with $|t| \leq \varepsilon$), as well as for perturbations of $\mathscr{C}^r$ expanding 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 Sect. 2). The setting and notations are as in Sect. 3.A and 3.B.

Lemma 4. (1) *For a fixed $n \in \mathbb{Z}^+$ and $\varphi \in \mathscr{C}^{r-1}$,*

$$\|\mathscr{L}_\varepsilon^n \varphi - \mathscr{L}^n \varphi\| \to 0 \quad \text{as } \varepsilon \to 0 .$$

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

$$\|\mathscr{L}_\varepsilon^n \varphi - \mathscr{L}^n \varphi\|_{r-2} = O(\varepsilon) , \qquad \varepsilon \to 0 .$$

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

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

(The induction hypothesis need only be applied to φ and $\mathscr{L}^{n-1}\varphi$.)
(1) Since $\mathscr{L}_\varepsilon \varphi = \theta_\varepsilon * \varphi$, each derivative satisfies $D^k(\mathscr{L}_\varepsilon \varphi) = \theta_\varepsilon * D^k(\mathscr{L}\varphi)$. It hence suffices to consider $\mathscr{C}^0$-norms. But if ψ is continuous the convolution $\theta_\varepsilon * \psi$ converges uniformly to ψ.

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

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

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

The key lemma follows:

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

$$\|\mathscr{L}_\varepsilon^n - \mathscr{L}^n\| < \Lambda^n\,.$$

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

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

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

$$f_{\vec{t}}^n(z) = f(\ldots(f(f(z) + t_1) + t_2)\ldots) + t_n\,,$$

we have

$$\begin{aligned}(\mathscr{L}_\varepsilon^n\varphi)(x) &= \int \ldots \int dt_1 \ldots dt_n \theta_\varepsilon(t_1)\ldots\theta_\varepsilon(t_n) \\ &\qquad \sum_{y_{\vec{t}}:\, f_{\vec{t}}^n(y_{\vec{t}}) = x} \varphi(y_{\vec{t}}) g(y_{\vec{t}}) \ldots g(f_{\vec{t}}^{n-1} y_{\vec{t}}) \\ &= \int \ldots \int dt_1 \ldots dt_n \theta_\varepsilon(t_1)\ldots\theta_\varepsilon(t_n) \sum_{y_{\vec{t}}:\, f_{\vec{t}}^n(y_{\vec{t}}) = x} (\mathscr{L}_{\vec{t}}^n \varphi)_{y_{\vec{t}}} \\ &= \int \ldots \int dt_1 \ldots dt_n \theta_\varepsilon(t_1)\ldots\theta_\varepsilon(t_n) (\mathscr{L}_{\vec{t}}^n\varphi)(x)\,,\end{aligned}$$

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

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 $(t_1, \ldots, t_n)$ with $|t_i| \le \varepsilon$, there is a natural bijection between the set $\{y: f^n(y) = x\}$ and the set $\{y_{\vec{t}}: f^n_{\vec{t}}(y_{\vec{t}}) = x\}$. Moreover, for each pair $(y, y_{\vec{t}})$ corresponding to a choice of an inverse branch of f^n at x we have

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

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

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

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

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

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

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

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

$$\begin{aligned}A &= \sum_{j=0}^{n-1} \varphi(y_{\vec{t}})g(y_{\vec{t}}) \dots [g'(f^j_{\vec{t}} y_{\vec{t}})g(f^j_{\vec{t}} y_{\vec{t}}) \dots g(y_{\vec{t}})]g(f^{j+1}_{\vec{t}} y_{\vec{t}}) \dots g(f^{n-1}_{\vec{t}} y_{\vec{t}}) \\ &= \sum_j (\varphi(y) \pm c_{n,\varepsilon}|\varphi'|)(g(y) \dots [g'(f^j(y)) \dots] \dots g(f^{n-1}y) \pm c_{n,\varepsilon}) \\ &= \left(\text{the corresponding part for } \frac{d}{dx}(\mathscr{L}^n\varphi)_y\right) \pm c_{n,\varepsilon}(|\varphi| + |\varphi'|)\,.\end{aligned} \tag{3.4}$$

For the second part, we get

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

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

$$(\mathscr{L}^n_\varepsilon\varphi)' = (\mathscr{L}^n\varphi)' \pm c_{n,\varepsilon}\|\varphi\|_1 \pm 2\|\varphi\|_1\lambda^{-n} \sum_{y:\, f^n(y)=x} \prod g(f^j(y))\,. \tag{3.6}$$

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

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

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

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

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

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

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

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

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

E. Proof of Theorem 1

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

To prove (1), we let $\Sigma_0 = \{1\}$. Lemma 5 together with the fact that (f,μ_0) is exact tell us that conditions (A.1) to (A.3) in Sect. 2 are met. We also know that $\|\mathscr{L}_\varepsilon\|$ is uniformly bounded, that 1 is always an eigenvalue of $\mathscr{L}_\varepsilon$ 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 $\kappa_0' < 1$, $\|\varrho_\varepsilon - \varrho_0\| = O\bigg(\Big(C_1(\varepsilon) + \frac{C_N(\varepsilon)}{\kappa_0'^N}\Big)^{1/d}\bigg)$ which tends to zero as $\varepsilon\to 0$ by Lemma 4 (1), proving stochastic stability in $(\mathscr{C}^{r-1}(S^1), \|\cdot\|)$. Since $C_N(\varepsilon) := \|\mathscr{L}_\varepsilon^N\varrho_0 - \varrho_0\|$, the speed with which $C_N(\varepsilon)$ tends to 0 depends on the modulus of continuity of $D^{r-1}\varrho_0$. In particular, if we rewrite everything with $X = \mathscr{C}^{r-2}(S^1)$ and $\|\cdot\|$ the $\mathscr{C}^{r-2}$ norm, then $D^{r-2}\varrho_0$ is Lipschitz and we have by Lemma 4 (2) $C_N(\varepsilon) = O(\varepsilon)$. This completes the proof of (1).

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

To see how $|\tau_\varepsilon - \tau_0|$ scales with ε, we let $\mathscr{L}$ act on $(\mathscr{C}^{r-2}(S^1), \|\cdot\|_{r-2})$ instead of $(\mathscr{C}^{r-1}(S^1), \|\cdot\|_{r-1})$. Since the eigenfunctions of $\mathscr{L}$ are always $\mathscr{C}^{r-1}$ the rates of decay of correlations 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 $\mathscr{L}$ acts, the definition of Σ_0 remains unchanged. In fact, X_0 stays the same (Ruelle [1989]). In the definition of $C_N(\varepsilon)$, we are now dealing with $\mathscr{C}^{r-2}$ norms for functions in X_0, a finite dimensional subspace of $\mathscr{C}^{r-1}(S^1)$. By Lemma 4 (2), we have $C_N(\varepsilon) = O(\varepsilon)$. Hence $|\tau_\varepsilon - \tau_0| = O(\varepsilon^{1/d})$.

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

4. Expanding Maps of Manifolds Followed by Stochastic Flows

This is a generalization of Sect. 3.

A. The Unperturbed Model

Here, M is a $\mathscr{C}^\infty$ compact, connected Riemannian manifold without boundary, and $f: M \to M$ is a $\mathscr{C}^r$ map for some $2 \le r < \infty$. We assume that f is *expanding*, i.e., there exists $\lambda > 1$ such that for all x in M and all v in $T_x M$, we have $|Df_x v| \ge \lambda |v|$. 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 $\mu_0 = \varrho_0\, dm$ with respect to which f is exact (see e.g. Mañé [1987]).

Let $\mathscr{F} = \{\varphi: M \to \mathbb{R}: \varphi \text{ is } \mathscr{C}^{r-1}\}$. For $\varphi \in \mathscr{F}$, we define $\|\varphi\|$ to be the $\mathscr{C}^{r-1}$ norm of φ, defined using a set of charts that will remain fixed throughout. The Perron-Frobenius operator $\mathscr{L}: \mathscr{F} \to \mathscr{F}$ 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

$$\operatorname{ess\,sp}(\mathscr{L}) \le \lambda^{-(r-1)}.$$

B. Type of Perturbation: Time-ε-Maps of Stochastic Flows

Let $X_0, X_1, \dots, X_m$ be $\mathscr{C}^\infty$ 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 $\mathscr{X}^\varepsilon$, our ε-perturbation of f, to be $\xi_\varepsilon \circ f$, i.e., $\mathscr{X}^\varepsilon$ is the Markov chain whose transition probabilities are given by

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

If the vector fields $X_0, \dots, X_m$ span the tangent space of M, then condition (2) from Sect. 1 is satisfied.

As in the last section, we wish to view $\mathscr{X}^\varepsilon$ as the composition of random maps. To do that we realize the solution of (4.1) as a stochastic process

$$\xi_t: \Omega \to \operatorname{Diff}^\infty(M),$$

where (Ω, P) is a probability space and $\{\xi_t\}$ satisfies

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

(ii) for $t_0 < t_1 < \dots < t_n$, 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.) Our perturbed process $\mathscr{X}^\varepsilon$ can then be viewed as the random map

$$\dots \circ f_{\omega_2} \circ f_{\omega_1},$$

where $\omega_1, \omega_2, \dots, \in \Omega$ are independent and $f_{\omega_i} := \xi_\varepsilon(\omega_i) \circ f$.

Using this representation of $\mathscr{X}^\varepsilon$, we can write the perturbed Perron-Frobenius operator $\mathscr{L}_\varepsilon : \mathscr{C}^{r-1}(M) \to \mathscr{C}^{r-1}(M)$ as

$$(\mathscr{L}_\varepsilon \varphi)(x) = \int P(d\omega)\,(\mathscr{L}_\omega \varphi)(x),$$

where

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

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

In the remainder of this subsection we summarize a few technical estimates about the $\mathscr{C}^r$-norms of ξ_ε that will be needed later on. For $\xi \in \mathrm{Diff}^r(M)$, we define the $\mathscr{C}^r$-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

$$\mathscr{U}_\delta : = \{\xi \in \mathrm{Diff}^r(M) : |||\xi||| < 1 + \delta\},$$
$$\mathscr{U}_\delta^n : = \{\xi = \eta_n \circ \ldots \circ \eta_1 : \eta_i \in \mathscr{U}_\delta, \forall i\},$$

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

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

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

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

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

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 $\varepsilon > 0$, the expectation*

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

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

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

The proof of Sublemma 1 also gives the uniform integrability of $|||\xi_\varepsilon|||^k$ 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 for all $A \subset \Omega$ with $P(A) < \beta$,*

$$E(|||\xi_\varepsilon||| \cdot \chi_A)^k < \alpha.$$

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

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

Proof of Sublemma 3. Write

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

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

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

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

C. Statement of our Results

Theorem 2. *Let $f:M \to M$ be a $\mathscr{C}^r$ expanding map as described in Sect. 4.A, with expanding constant λ, and let $\mu_0 = \varrho_0\,dm$ be its unique absolutely continuous invariant probability measure. Let $\{\mathscr{X}^\varepsilon, \varepsilon > 0\}$ be a small random perturbation of f of the type described in Sect. 4.B, with invariant probability measure $\mu_\varepsilon = \varrho_\varepsilon\,dm$. Then:*
(1) The dynamical system (f, μ_0) is stochastically stable under $\mathscr{X}^\varepsilon$ in the space of $\mathscr{C}^{r-1}$ functions, i.e., the $\mathscr{C}^{r-1}$-norm of $\varrho_\varepsilon - \varrho_0$ tends to zero as $\varepsilon \to 0$.
Let τ_0 and τ_ε be the rates of decay of correlations for f and $\mathscr{X}^\varepsilon$ respectively, in the space of $\mathscr{C}^{r-1}$ functions. If, in addition, $\tau_0 > \lambda^{-(r-1)}$, then:
(2) The rate of mixing for f is robust, i.e., $\tau_\varepsilon \to \tau_0$ as $\varepsilon \to 0$.
We show in fact that
(3) For each $\delta > 0$, outside of $\{|z| \le \lambda^{-(r-1)} + \delta\}$, the spectrum of $\mathscr{L}_\varepsilon$ converges to that of $\mathscr{L}$ as $\varepsilon \to 0$.

Remark. We conjecture that the correct scaling in ε for this kind of perturbation is $\|\varrho_\varepsilon - \varrho_0\|_{r-2} = O(\sqrt{\varepsilon})$.

D. Dynamical Lemmas

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

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

$$\|\mathscr{L}_\varepsilon^n \varphi - \mathscr{L}^n \varphi\| \to 0 \quad as\ \varepsilon \to 0\,.$$

Lemma 7. *Let $\Lambda > \lambda^{-(r-1)}$ be given. Then there exists $N_0 \in \mathbb{Z}^+$ such that for all $n \ge N_0$ there exists $\varepsilon(n) > 0$ such that for each $\varepsilon < \varepsilon(n)$,*

$$\|\mathscr{L}_\varepsilon^n - \mathscr{L}^n\| < \Lambda^n\,.$$

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

Using the random maps representation of $\mathscr{L}^\varepsilon$ and the notation in Sect. 3.D, we have $f^n_{\vec\omega} = f_{\omega_n} \circ \ldots \circ f_{\omega_1}$ if $\vec\omega = (\omega_1, \ldots, \omega_n) \in \Omega^n$, and

$$(\mathscr{L}^n_\varepsilon \varphi)(x) = \int \ldots \int P(d\omega_1) \ldots P(d\omega_n) (\mathscr{L}^n_{\vec\omega} \varphi)(x),$$

where

$$(\mathscr{L}^n_{\vec\omega})(x) = \sum_{y\,:\,f^n_{\vec\omega} y = x} \varphi(y) \cdot \frac{1}{|\det Df^n_{\vec\omega}(y)|}.$$

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

Sublemma 4.

$$\frac{d}{dx_i}(\mathscr{L}^n_\varepsilon \varphi) = \int \ldots \int P(d\omega_1) \ldots P(d\omega_n) \frac{d}{dx_i} (\mathscr{L}^n_{\vec\omega} \varphi).$$

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

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

where

$$\Phi_t(\vec\omega) = \frac{1}{t}\{(\mathscr{L}^n_{\vec\omega}\varphi)(x + tu_i) - (\mathscr{L}^n_{\vec\omega}\varphi)(x)\} = \frac{d}{dx_i} (\mathscr{L}^n_{\vec\omega}\varphi)(x_t),$$

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 $\Phi \in L^1(\Omega^n, P^n)$ with $|\Phi_t| \le |\Phi|$. Differentiating the expression for $\mathscr{L}^n_{\vec\omega}\varphi$ above, we observe that $\frac{d}{dx_i} (\mathscr{L}^n_{\vec\omega}\varphi)(x_t)$ is the sum of finitely many terms, each one of which is bounded in absolute value by a product of the form

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

where C is a constant depending on f and n, and $k_1, \ldots, k_n$ 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. □

Consider first $\vec\omega = (\omega_1, \ldots, \omega_n)$, where $f^k_{\vec\omega}$ is $\mathscr{C}^2$ very near f^k for $1 \le k \le n$, say $\|f^k_{\vec\omega} - f^k\|_2 < \delta$ for some $\delta > 0$. We assume δ is small enough so that the inverse branches of $f^n_{\vec\omega}$ are easily identified with those of f^n. Then the same argument as in Sect. 3.D, line by line, gives

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

and

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

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

$$\|\mathscr{L}^n_{\vec\omega}\varphi - \mathscr{L}^n\varphi\| \le \Lambda'^n\|\varphi\|$$

for some $\lambda^{-(r-1)} < \Lambda' < \Lambda$. We then choose $\varepsilon \ll \delta$ such that if $\Omega_0 := \{\omega : \|f_\omega - f\|_2 \geq \delta\}$, then $P(\Omega_o)$ is very small, small enough that these "bad" $\vec{\omega}$ do not contribute significantly to $\|\mathscr{L}_\varepsilon^n \varphi - \mathscr{L}^n \varphi\|$. More precisely, let

$$\Omega_0^n := \{(\omega_1, \dots, \omega_n) : \omega_i \notin \Omega_0, \forall i\} \quad \text{and} \quad \Omega_j^n := \{(\omega_1, \dots, \omega_n) : \omega_j \in \Omega_0\}.$$

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

$$|\mathscr{L}_\varepsilon^n \varphi - \mathscr{L}^n \varphi| = \left| \int_{\Omega^n} dP^n(\vec{\omega}) (\mathscr{L}_{\vec{\omega}}^n \varphi - \mathscr{L}^n \varphi) \right|$$
$$\leq \int_{\Omega_0^n} |\mathscr{L}_{\vec{\omega}}^n \varphi - \mathscr{L}^n \varphi| + \sum_{j=1}^n \int_{\Omega_j^n} (|\mathscr{L}_{\vec{\omega}}^n \varphi| + |\mathscr{L}^n \varphi|).$$

The Ω_0^n-term has been shown to be bounded above by $c_{n,\varepsilon} \cdot \|\varphi\|_1$, and

$$\int_{\Omega_j^n} |\mathscr{L}^n \varphi| \leq \|\mathscr{L}^n\| \cdot \|\varphi\|_1 \cdot P(\Omega_0),$$

the last factor of which can be made small as $\varepsilon \to 0$. It remains to estimate $\int_{\Omega_j^n} |\mathscr{L}_{\vec{\omega}}^n \varphi|$. Note that $\mathscr{L}_{\vec{\omega}}^n \varphi$ is a sum of finitely many terms of the form

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

This expression is bounded above by

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

Its integral over Ω_j^n is therefore bounded above by

$$C \cdot |\varphi| \cdot \left(\prod_{i \neq j} E|||\xi_\varepsilon|||^{k_i} \right) \cdot E\left(|||\xi_\varepsilon|||^{k_j} \cdot \chi_{\Omega_0} \right).$$

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

$$|\mathscr{L}_\varepsilon^n \varphi - \mathscr{L}^n \varphi| \leq c_{n,\varepsilon} \cdot \|\varphi\|_1.$$

A similar argument (see Sublemma 4) gives

$$\left| \frac{d}{dx_i} \mathscr{L}_\varepsilon^n \varphi - \frac{d}{dx_i} \mathscr{L}^n \varphi \right| \leq \Lambda'^n \|\varphi\|_1 + c_{n,\varepsilon} \|\varphi\|_1 \leq \Lambda^n \|\varphi\|_1. \quad \square$$

E. Proof of Theorem 2

Use Sect. 2 and proceed as in Sect. 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 $\mathscr{C}^2$, piecewise expanding map. More precisely, we assume that there exists a partition $0 = a_0 < a_1 < \ldots < a_M = 1$ of I such that for each i, the restriction $f|_{[a_i, a_{i+1}]}$ can be extended to a $\mathscr{C}^2$ map with $\min |f'| \geq \lambda > 1$. 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

$$\operatorname*{var}_{[a,b]} \varphi = \sup \left\{ \sum_{i=0}^{n} |\varphi(x_{i+1}) - \varphi(x_i)| : n \geq 1,\, a \leq x_0 < x_1 < \ldots < x_n \leq b \right\}.$$

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

$$\|\varphi\| = \operatorname*{var}_I \varphi + |\varphi|_1 .$$

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

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

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

[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 $|\varrho_0|_1 = 1$. 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 respect to μ_0. Under these assumptions, it has been shown that 1 is the only point of $\sigma(\mathscr{L})$ on the unit circle, its generalized eigenspace is one-dimensional, and that $\tau_0 := \sup\{|z| : z \in \sigma(\mathscr{L}), z \neq 1\} < 1$ 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 $0 < \gamma \leq 1$, we define

$$\|\varphi\|_\gamma = \gamma \cdot \operatorname*{var}_I \varphi + |\varphi|_1 .$$

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

B. Type of Perturbation: Convolutions

As in Sect. 3.B, we consider a small random perturbation $\mathscr{X}^\varepsilon$ 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 $\mathscr{L}_\varepsilon$ acting on $(BV, \|\cdot\|)$. As in the first two models, $\mathscr{L}_\varepsilon$ has 1 as an eigenvalue with eigenfunction ϱ_ε which is the density of an invariant probability measure μ_ε for $\mathscr{X}^\varepsilon$ (Lemma 19 in Keller [1982]).

It is known that not all piecewise expanding maps are stochastically stable. A major difference between the situation here and that in Sect. 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 Sect. 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 Fig. 1 for an example in which a turning point fixed by f generates $2^n - 2$ extra intervals of monotonicity for $f^n_{\vec{t}}$. This example is not stochastically stable, not even in the sense of weak convergence of μ_ε (see Keller [1982, Sect. 6] and also Blank [1992]).

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

From our discussion in the last subsection we see 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 $k \geq 1$. The kernel θ_ε used in our convolutions is called symmetric if $\theta_\varepsilon(x) = \theta_\varepsilon(-x)$, $\forall x$. The definition of Θ is given in Sect. 5.A.

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

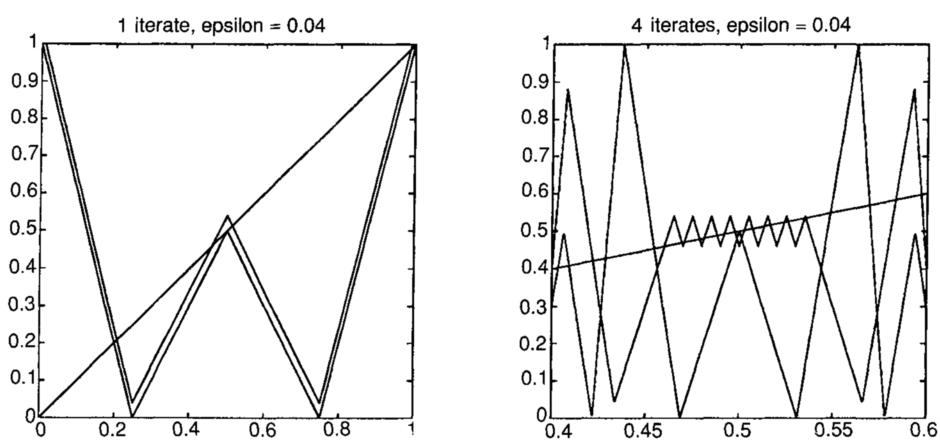

Fig. 1. The fourth iterate of a map with a fixed turning point compared to the fourth iterate of a perturbation

Theorem 3. *Let $f: I \to I$ be as described in Sect. 5.A, with a unique absolutely continuous invariant probability measure $\mu_0 = \varrho_0\, dm$, and let $\mathscr{X}^\varepsilon$ be a small random perturbation of f of the type described in Sect. 5.B with invariant probability measure $\varrho_\varepsilon\, dm$. We assume also that f has no periodic turning points. Then*
(1) *The dynamical system (f, μ_0) is stochastically stable under $\mathscr{X}^\varepsilon$ in $L^1(dm)$, i.e., $|\varrho_\varepsilon - \varrho_0|_1$ tends to 0 as $\varepsilon \to 0$.*
Let τ_0 and τ_ε be the rates of decay of correlations for f and $\mathscr{X}^\varepsilon$ respectively for test functions in BV.
(2) *If $\tau_0^2 > \Theta$ then $\tau_\varepsilon \to \tau_0$ as $\varepsilon \to 0$.*
We show in fact that
(3) *if we let $\tau = \min\{|z| : z \in \sigma(\mathscr{L}), |z| > \sqrt{\Theta}\}$, then there exists $\delta > 0$ such that the spectrum of $\mathscr{L}_\varepsilon$ restricted to $\{|z| \geq \tau - \delta\}$, converges to that of $\mathscr{L}$ (restricted to the same domain) as $\varepsilon \to 0$.*

Theorem 3′. *Let f and $\mathscr{X}^\varepsilon$ 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 $\sqrt{\Theta}$ is replaced by $\sqrt{2\Theta}$; or $\sqrt{\Theta}$ by $\sqrt{(3/2)\Theta}$ if θ_ε is symmetric.*
The square roots arise from our use of balanced norms in the proofs of Lemmas 9 and 9′. We do not know to what extent they are needed. We do not know either if we can weaken the replacement of Θ by 2Θ [or $(3/2)\Theta$] in Theorem 3′. However, it is clear that *some* hypothesis on f or on the nature of our perturbations is necessary to give the type of results we want (see Sect. 5.B). We remark also that the hypothesis we use for proving stochastic stability is slightly weaker than that in Keller [1982, Sect. 6] or Kifer [1988a, Chap. IV] (in the latter reference, only weak convergence is shown and the assumption that $\lambda > 2$ is implicitly used, see also Blank [1992]).

D. Dynamical Lemmas

The setting and notations are as in Sect. 5.A and 5.B. We have the obvious lemma:

Lemma 8. *For fixed $n \geq 1$ and $\varphi \in L^1$,*

$$|\mathscr{L}_\varepsilon^n \varphi - \mathscr{L}^n \varphi|_1 \to 0 \quad \text{as } \varepsilon \to 0\,.$$

It is not true in general that $\mathrm{var}(\mathscr{L}_\varepsilon \varphi - \mathscr{L}\varphi) \to 0$ as $\varepsilon \to 0$ for a fixed $\varphi \in BV$.
We will use the notations $c_{n,\varepsilon}$, $g = 1/|f'|$, and $f_{\vec{t}}^n, \mathscr{L}_{\vec{t}}^n$ of Sect. 3.D. We also write

$$g^n(y) = g(y) \cdot g(fy) \ldots g(f^{n-1}y)\,,$$
$$g_{\vec{t}}^n(y_{\vec{t}}) = g(y_{\vec{t}}) \cdot g(f_{\vec{t}}\, y_{\vec{t}}) \ldots g(f_{\vec{t}}^{n-1} y_{\vec{t}})\,,$$

We let

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

and $\mathscr{M} = \max M_i \leq M + 1$. Note that f is without periodic turning points if and only if $\mathscr{M} < \infty$.
Denote by $\mathscr{Z}_n$ the "partition" of I into (closed) intervals of monotonicity of f^n, and by $\mathscr{Z}_{n,\vec{t}}$ the "partition" of I into (closed) intervals of monotonicity for $f_{\vec{t}}^n$. Write $\mathscr{Z}_1 = \eta_1 \cup \ldots \cup \eta_M$. By definition an element $\eta(j_0, \ldots, j_{n-1})$ of $\mathscr{Z}_n$ 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 $\eta'(j_0,\dots,j_{n-1})$ of $\mathscr{Z}_{n,\vec t}$ is an interval of the form

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

with nonempty interior.

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

Consider now the case $\mathscr{M} \geq 1$. We fix η and assume that ε is sufficiently small for this value of n. Consider $f^n_{\vec t}$, where each $|t_i| < \varepsilon$. We associate elements of $\mathscr{Z}_n$ with those in $\mathscr{Z}_{n,\vec t}$ as before, but in general this will not account for all the elements of $\mathscr{Z}_{n,\vec t}$. An element of $\mathscr{Z}_{n,\vec t}$ without a counterpart in $\mathscr{Z}_n$ is called *nonadmissible.*

Let us look at how nonadmissible elements are created. Let a_i be a turning point, and let $q > 0$ be the first time $f^j a_i$ returns to the turning set. From the definition of $\mathscr{Z}_{q,\vec t}$, we see that the two intervals adjacent to a_i in $\mathscr{Z}_{q,\vec t}$ are admissible, but that $\mathscr{Z}_{q+1,\vec t}$ may have two nonadmissible intervals adjacent to a_i. This is due to the fact that $f^q(a_i-\delta, a_i+\delta)$ lies on one side of some turning point $a_{i'}$, while $f^q_{\vec t}(a_i-\delta, a_i+\delta)$ may intersect both sides of $a_{i'}$. We think of these two newly created nonadmissible intervals as so short that their dynamics up to time n is tied to that of a_i.

If there is no q', with $q < q' < n$, such that $f^{q'} a_i$ is in the turning set again, then in $\mathscr{Z}_{n,\vec t}$ these two nonadmissible intervals will be the only ones between the admissible intervals nearest to a_i. If, however, such a q' exists, then the same mechanism as before may create two new nonadmissible intervals for $\mathscr{Z}_{q'+1,\vec t}$. In addition to that, each one of the already existent nonadmissible intervals near a_i may get divided again, giving rise to a total of $2^2 + 2 = 6$ nonadmissible intervals near a_i in $\mathscr{Z}_{q'+1,\vec t}$.

Continuing this reasoning, if a_i returns to the turning set L times before time n, then the maximum number of nonadmissible intervals created near a_i is $2^{(L+1)} - 2$. Also, if $f^k x = a_i$ for $k < n$, then an imprint of the picture at a_i is made at x, giving rise to other nonadmissible intervals between admissible ones in $\mathscr{Z}_{n,\vec t}$. These are the *only* ways in which nonadmissible intervals are created.

To sum up, we have the following estimates. If f has no periodic turning points, i.e. if $\mathscr{M} < \infty$, then between any two admissible intervals in $\mathscr{Z}_{n,\vec t}$ there are at most $2^{\mathscr{M}+1}$ nonadmissible ones. If f has periodic turning points, then the maximum number of contiguous nonadmissible intervals is at most $2^n - 2$.

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

Assume first that $\mathscr{M} = 0$ and ε is small enough. Let $\eta \in \mathscr{Z}_n$ and $\eta' \in \mathscr{Z}_{n,\vec t}$ be a pair of associated intervals of monotonicity. We decompose η and η' into two parts 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 $\eta_B = \eta \backslash \eta_G$ and $\eta'_B = \eta' \backslash \eta'_G$. 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*. We denote by B the union of all co-respondents η_B and by B' the union of all co-respondents η'_B. Then, for fixed n the measures of B and B' both tend to zero as ε tends to zero.

In the case where $\mathscr{M} \geq 1$, we decompose associated intervals $\eta \in \mathscr{Z}_n$ and $\eta' \in \mathscr{Z}_{n,\vec t}$ into $\eta' = \eta'_G \cup \xi_B$ and $\eta = \eta_G \cup \eta_B$ as described in the case $\mathscr{M} = 0$. 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 half of the nonadmissible intervals immediately to the left and half of those to the right of η'. 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 $\Theta < \Lambda^2 < 1$. Then, there exist $C > 0$ and $N_0 \in \mathbb{Z}^+$ such that for each $n \geq N_0$ there exists $\varepsilon(n) > 0$ such that for each $\varepsilon < \varepsilon(n)$,*

$$\|\mathscr{L}_\varepsilon^n - \mathscr{L}^n\|_{\Lambda^n} < C \cdot \Lambda^n .$$

Recall that $\|\cdot\|_{\Lambda^n}$ is the balanced norm with weight Λ^n (see Sect. 5.A).

Proof of Lemma 9. In the proof, $\tilde{\Theta}$ denotes a generic 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

$$\|\mathscr{L}^n\varphi - \mathscr{L}_{\vec{t}}^n\varphi\| \leq \|\mathscr{L}_{\vec{t}}^n(\varphi\chi_{B'})\| + \|\mathscr{L}^n(\varphi\chi_B)\| + \|\mathscr{L}^n(\varphi\chi_{(I\setminus B)} - \mathscr{L}_{\vec{t}}^n(\varphi\chi_{I\setminus B'})\| \cdot \tag{5.1}$$

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

For each $\eta'\hat{B} \in B'$ and for $x \in f_{\vec{t}}^n\eta'\hat{B}$, we have

$$\mathscr{L}_{\vec{t}}^n(\varphi\chi_{\eta'\hat{B}})(x) = \varphi(y_{\vec{t}}) \cdot g(y_{\vec{t}}) \ldots g(f^{n-1}y_{\vec{t}}),$$

where $y_{\vec{t}}$ is the unique element of $\eta'\hat{B}$ such that $f_{\vec{t}}^n(y_{\vec{t}}) = x$. It follows that

$$|\mathscr{L}_{\vec{t}}^n(\varphi\chi_{\eta'\hat{B}})|_1 \leq \int_{\eta'\hat{B}} |\varphi| \leq l(\eta'\hat{B}) \cdot (\operatorname{var}\varphi + |\varphi|_1), \tag{5.2}$$

where $l(\eta'\hat{B})$ denotes the length of the interval $\eta'\hat{B}$.

Summing (5.2) over all intervals $\eta'\hat{B}$, we get

$$|\mathscr{L}_{\vec{t}}^n(\varphi\chi_{B'})|_1 \leq c_{n,\varepsilon} \cdot (\operatorname{var}\varphi + |\varphi|_1). \tag{5.3}$$

For the variation, we have for any interval $\eta' \in \mathscr{Z}_{n,\vec{t}}$,

$$\operatorname{var}\mathscr{L}_{\vec{t}}^n(\varphi\chi_{\eta'}) \leq \operatorname*{var}_{\eta'}\varphi \cdot \sup_{\eta'} g_{\vec{t}}^n + \sup_{\eta'}|\varphi| \cdot \operatorname*{var}_{\eta'} g_{\vec{t}}^n + 2 \cdot \sup_{\eta'}|\varphi| \cdot \sup_{\eta'} g_{\vec{t}}^n . \tag{5.4}$$

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

$$\begin{cases} \sup_{\eta'} g_{\vec{t}}^n \leq \tilde{\Theta}^n \\ \operatorname*{var}_{\eta'} g_{\vec{t}}^n \leq \tilde{\Theta}^n . \end{cases} \tag{5.5}$$

(The first inequality is obvious, the second is proved by induction.)

Set $n_2 = \max(n_0, n_1)$ and assume first that $n = n_2$. The interval $\eta'\hat{B}$ is a subset of some $\eta' \in \mathscr{Z}_{n,\vec{t}}$ and is a co-respondent of a unique good interval $\eta'\hat{B}$. From (5.4) and (5.5), denoting by η'' the smallest interval containing $\eta'\hat{G}$ and η', we obtain:

$$\begin{aligned}\operatorname{var}\mathscr{L}_{\vec{t}}^{n}(\varphi\chi_{\eta'\hat{B}}) &\le \operatorname{var}\mathscr{L}_{\vec{t}}^{n}(\varphi\chi_{\eta'})\\ &\le \operatorname*{var}_{\eta'}\varphi\cdot\tilde{\Theta}^{n} + \Big(\operatorname*{var}_{\eta''}\varphi + \inf_{\eta''}|\varphi|\Big)\Big(\operatorname*{var}_{\eta'} g_{\vec{t}}^{n} + 2\cdot\sup_{\eta'} g_{\vec{t}}^{n}\Big)\\ &\le \operatorname*{var}_{\eta''}\varphi\cdot 4\tilde{\Theta}^{n} + \left(\frac{\operatorname*{var}_{\eta'} g_{\vec{t}}^{n} + 2\cdot\sup_{\eta'} g_{\vec{t}}^{n}}{l(\eta'')}\right)\cdot l(\eta'')\inf_{\eta''}|\varphi|\\ &\le \operatorname*{var}_{\eta''}\varphi\cdot 4\tilde{\Theta}^{n} + D\cdot l(\eta'')\inf_{\eta''}|\varphi|\,,\end{aligned}\tag{5.6}$$

where $D = \sup_{\eta'\in\mathscr{Z}_{n,\vec{t}}}[\operatorname*{var}_{\eta'} g_{\vec{t}}^{n_2} + 2\cdot\sup_{\eta'} g_{\vec{t}}^{n_2}]/l_{n_2}$, with l_{n_2} equal to the infimum of the lengths of admissible intervals in $\mathscr{Z}_{n_2,\vec{t}}$. Note that when ε tends to zero, l_{n_2} tends to $\inf l(\eta)$, for η in $\mathscr{Z}_{n_2}$, and observe that $l(\eta'')\inf_{\eta''}|\varphi| \le \int_{\eta''}|\varphi|$.

Summing (5.6) over all intervals $\eta'\hat{B}$, and using the fact that the good intervals $\eta'\hat{G}$ are overcounted at most $2^{\mathscr{M}}$ times, we get for $n = n_2$,

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

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

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

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 $\mathscr{Z}_{n_2,\vec{u}}$, respectively $\mathscr{Z}_{n_2,\vec{v}}$ such that a given $\eta' \in \mathscr{Z}_{n,\vec{t}}$ is equal to $(f_{\vec{v}}^{n_2}|_{\xi})^{-1}(\xi)\cap\zeta$, then

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

In particular

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

A standard induction argument yields

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

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

The problem we have to deal with now is 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]. Not knowing which γ to choose for now, we rewrite (5.3) and (5.8) using our new norm $\|\cdot\|_\gamma$:

$$|\mathscr{L}_{\vec{t}}^n(\varphi\chi_{B'})|_1 \le c_{n,\varepsilon} \cdot (\gamma \cdot \operatorname{var}\varphi + |\varphi|_1),$$
$$\gamma \cdot \operatorname{var}(\mathscr{L}_{\vec{t}}^n(\varphi\chi_{B'})) \le \gamma \cdot \tilde{\Theta}^n \cdot \operatorname{var}\varphi + \gamma \cdot D' \cdot |\varphi|_1 .$$

Together, they give

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

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

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

We have an analogous lower bound. Summing over j, we get:

$$\begin{aligned}
|\mathscr{L}_{\vec{t}}^n(\varphi\chi_{(I\backslash B')}) - \mathscr{L}^n(\varphi\chi_{(I\backslash B)})|_1 &\le \sup|\mathscr{L}_{\vec{t}}^n(\varphi\chi_{(I\backslash B')})(x) - \mathscr{L}^n(\varphi\chi_{(I\backslash B)})(x)| \\
&\le \tilde{\Theta}^n \operatorname{var}\varphi + c_{n,\varepsilon}|\varphi|_1 .
\end{aligned} \tag{5.11}$$

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.

Consider an associated pair (η_G, η'_G) which for simplicity of notation we write as (η, η'). Defining the bijection $\Psi: \eta' \to \eta$ by $\Psi(y_{\vec{t}}) = y$, we obtain

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

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

We have also used the fact that $\eta' \to \eta$ as $\varepsilon \to 0$, so that $\eta \cup \eta'$ is a connected interval.

Summing the above inequalities over all elements of $\mathscr{Z}_n$, and noting that intervals of the form $\eta \cup \eta'$ intersect at most two of their neighbors, we get

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

From (5.11) and (5.13) we find:

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

Adding (5.9), the analogue of (5.9) for $\mathscr{L}^n$ and (5.14), and integrating over $\vec{t}$, we obtain

$$\|\mathscr{L}_\varepsilon^n - \mathscr{L}^n\|_\gamma \leq 2(\tilde{\Theta}^n + D'\gamma) + \gamma^{-1}\tilde{\Theta}^n .$$

Remembering that $\Lambda^2 > \tilde{\Theta}$, we see that if we let $\gamma = \Lambda^n$, then the right side of the above inequality is bounded above by $C \cdot \Lambda^n$. This completes the proof of Lemma 9. □

Lemma 9′. *Let Λ be such that $\Theta < \min(\Lambda, 2\Lambda^2)$. Then there exist $C > 0$ and $N_0 \in \mathbb{Z}^+$ such that for each $n \geq N_0$ there exists $\varepsilon(n) > 0$ such that for each $\varepsilon < \varepsilon(n)$,*

$$\|\mathscr{L}_\varepsilon^n - \mathscr{L}^n\|_{\Lambda^n} < C \cdot (2\Lambda)^n .$$

If each θ_ε is symmetric, then for Λ such that $\Theta < \min(\Lambda, (3/2)\Lambda^2)$ we have the better inequality

$$\|\mathscr{L}_\varepsilon^n - \mathscr{L}^n\|_{\Lambda^n} < C \cdot \left(\tfrac{3}{2}\Lambda\right)^n .$$

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

We see that the only important change occurs when we sum (5.6) over the intervals $\eta'\hat{B}$. Since each good interval $\eta'\hat{G}$ has at most 2^{n-1} co-respondents, the sum yields for $n = n_2$:

$$\operatorname{var}(\mathscr{L}_{\vec{t}}^n(\varphi\chi_{B'})) \leq 2 \cdot (2\tilde{\Theta})^n \cdot \operatorname{var}(\varphi) + 2^{n-1}D \cdot |\varphi|_1 .$$

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

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

Inequality (5.9) hence becomes

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

Inequality (5.14) does not have to be changed. Summing up, we have

$$\|\mathscr{L}_\varepsilon^n\varphi - \mathscr{L}^n\varphi\|_\gamma \leq ((2\tilde{\Theta})^n + \gamma^{-1}\tilde{\Theta}^n + \gamma 2^n D')\,\|\varphi\|_\gamma ,$$

and hence the inequality as claimed.

Assume now that each θ_ε is symmetric. 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 $\eta_G \subset \eta \in \mathscr{Z}_n$, where a boundary point b of η is periodic. We claim that there exists a sequence $S = \{s_j\}_{j=1,\dots,n}$ of signs $s_j =\in \{+,-\}$ such that η'_G has at most $2^{k(S)}$ nonadmissible co-respondents η'_B, where $0 \le k(S) \le n$ 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 Fig. 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

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

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

$$\operatorname{var}\mathscr{L}_\varepsilon^n(\varphi\chi_{B'}) \le \sum_k \int \theta_\varepsilon(t_1)\dots\theta_\varepsilon(t_n)A_k \le ((3/2)\cdot\tilde{\Theta})^n\cdot\operatorname{var}\varphi + (3/2)^n D\cdot|\varphi|_1 .$$

We thus obtain

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

which yields the claim. □

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 $t\mapsto\theta_\varepsilon(t)\psi(x,t)$ is in $L^1(dm)$ for each fixed x, then

$$\left|\int dt\,\theta_\varepsilon(t)\psi(\cdot,t)\right|_1 \le \int dt\,\theta_\varepsilon(t)\,|\psi(\cdot,t)|_1 ,$$

$$\operatorname{var}_x\left(\int dt\,\theta_\varepsilon(t)\psi(x,t)\right) \le \int dt\,\theta_\varepsilon(t)\operatorname{var}_x\psi(x,t) .$$

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: a Modified Version of Sect. 2

Because of the need to introduce the norms $\|\cdot\|_\gamma$, we need a slightly refined version of Sect. 2. Again, $(X, \|\cdot\|)$ is a complex Banach space, and $\{T_\varepsilon, \varepsilon\ge 0\}$ is a family of bounded linear operators on X. We assume that T_0 satisfies conditions (A.1) and (A.3) in Sect. 2, i.e., $\sigma(T_0) = \Sigma_0\cup\Sigma_1$ with

$$\kappa_1 := \sup\{|z| : z\in\Sigma_1\} < \inf\{|z| : z\in\Sigma_0\} =: \kappa_0 ,$$

and $\dim X_0 < \infty$. We further assume that Σ_1 can be written as the union of isolated sets

$$\Sigma_1 = \Sigma_{1,0}\cup\Sigma_{1,1} , \tag{A′.1}$$

where $\Sigma_{1,0}$ could be empty and $\dim X_{1,0}$ is at most finite. (The notations $\pi_{1,0}, \pi_{1,1}$, $X_{1,0}$ and $X_{1,1}$ have the obvious meanings.) Let

$$\kappa_{11} := \sup\{|z| : z \in \Sigma_{1,1}\}.$$

We assume 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 $(\kappa_{11}/\kappa_0) < \kappa < \kappa_0$ such that for each large enough $N \in \mathbb{Z}^+$ there exists $\varepsilon(N)$ such that for all $0 < \varepsilon < \varepsilon(N)$,

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

We shall need two sublemmas:

Sublemma 5. *Assume* (A.1), (A$'$.1), *and* (A.3). *Then for any* $\kappa_0' < \kappa_0$, $\kappa_1' > \kappa_1$, *there exists* N_0 *such that for all* $n \ge N_0$, *any* $0 < \gamma \le 1$, *and any* $x \in X_0$, $y \in X_{1,0}$,
(1) $\|T_0^n x\|_\gamma \ge (\kappa_0')^n \|x\|_\gamma$,
(2) $\|T_0^n y\|_\gamma \le (\kappa_1')^n \|y\|_\gamma$.

Proof of Sublemma 5. We prove (1). Since X_0 is finite dimensional, all norms are equivalent. We choose N_0 such that for all $n \ge N_0$ and $x \in X_0$,

$$|T_0^n x| \ge (\kappa_0')^n |x| \quad \text{and} \quad \|T_0^n x\| \ge (\kappa_0')^n \|x\|.$$

The same inequality then holds for $\|\cdot\|_\gamma$ which is a weighted average of $|\cdot|$ and $\|\cdot\|$. □

Sublemma 6. *If* (A.1), (A$'$.1), *and* (A.3) *hold, then there exists a constant* C *such that for any* $0 < \gamma \le 1$, *we have* $\|\pi_0\|_\gamma \le C$, $\|\pi_{1,0}\|_\gamma \le C$, *and* $\|\pi_1\|_\gamma \le 2C + 1$.

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

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

where we have used again the fact that the norms $|\cdot|$ and $\|\cdot\|$ are equivalent on the finite-dimensional space X_0. We proceed in the same way for $\|\pi_{1,0}\|_\gamma$. To finish, observe that $\pi_0 + \pi_{1,0} + \pi_1 = I$ so that $\|\pi_1\|_\gamma \le \|\pi_0\|_\gamma + \|\pi_{1,0}\|_\gamma + 1$. □

We can now prove:

Lemma 1$'$. *Assume* (A.1), (A.3), (A$'$.1), *and* (A$'$.2), *then the conclusion of Lemma 1 from Sect. 2 is true.*

Proof of Lemma 1$'$. Let

$$\kappa_1 < \kappa_1' < \kappa' < \kappa_0' < \kappa_0'' < \kappa_0,$$
$$\kappa_{11} < \kappa_{11}' < \kappa < \kappa',$$
$$\frac{\kappa_{11}'}{\kappa} < \kappa'.$$

Let N be large enough for various purposes. In particular, we require (see Sublemma 5) that

$$x \in X_0 \Rightarrow \|T_0^N x\|_{\kappa^N} \geq (\kappa_0'')^N \|x\|_{\kappa^N} ,$$
$$x \in X_{1,0} \Rightarrow \|T_0^N x\|_{\kappa^N} \leq (\kappa_1')^N \|x\|_{\kappa^N} ,$$
$$x \in X_{1,1} \Rightarrow \|T_0^N x\| \leq (\kappa_{11}')^N \|x\| .$$

We let $\varepsilon < \varepsilon(N)$ and will show that $\lambda \notin \sigma(T_\varepsilon)$ for λ with $\kappa' < |\lambda| < \kappa_0'$ (if κ' is close enough to κ_0).

We proceed as in Lemma 1, using $\|\cdot\|_{\kappa^N}$ in the place of $\|\cdot\|$ and estimating $\|R(T_0^N, \lambda^N)\|_{\kappa^N}$ by projecting onto X_0, $X_{1,0}$, and $X_{1,1}$. It follows from our choice of constants that for $x \in X_0$, we have

$$\|T_0^N - \lambda x\|_{\kappa^N} \geq \text{const} \cdot (\kappa_0'')^N \|x\|_{\kappa^N} ,$$

and for $x \in X_{1,0}$, we have

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

As for $x \in X_{1,1}$, we have

$$\|T_0^N x\|_{\kappa^N} \leq \|T_0^N x\| \leq (\kappa_{11}')^N \|x\| \leq \left(\frac{\kappa_{11}'}{\kappa}\right)^N \|x\|_{\kappa^N} ,$$

from which it follows that

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

These estimates together with Sublemma 6 give

$$\|R(T_0^N, \lambda^N)\|_{\kappa^N} \leq \frac{1}{\kappa^N} . \quad \square$$

Note that, unlike the situation in Sect. 2, κ' cannot be taken arbitrarily near κ.

Lemma 2 from Sect. 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 $\varepsilon(N)$ such that, for each $\varepsilon < \varepsilon(N)$, $\|\pi_0 - \pi_0^\varepsilon\|_{\kappa^N} < \delta$), and the same proof.

Define

$$C_1^*(\varepsilon) := \sup_{\substack{x \in X_0 \\ x \neq 0}} \frac{|T_\varepsilon x - T_0 x|}{|x|}$$

and assume that

$$C_1^*(\varepsilon) \to 0 \quad \text{as } \varepsilon \to 0 . \tag{A$'$.4}$$

Lemma 3$'$. *Assume* (A.1), (A.3), (A$'$.1), (A$'$.2), (A$'$.4), *and that* $|T_\varepsilon|$ *is uniformly bounded. Then*

$$\sigma(T_\varepsilon|_{X_0^\varepsilon}) \to \sigma(T_0|_{X_0})$$

as $\varepsilon \to 0$.

Proof of Lemma 3$'$. As in Lemma 3, we show that $X_0^\varepsilon = \text{graph}(S_\varepsilon)$ for some linear $S_\varepsilon : X_0 \to X_1$ with $\|S_\varepsilon\|_{\kappa^N} \to 0$ as $N \to \infty$ and $\varepsilon \to 0$, $\varepsilon < \varepsilon(N)$.

Define $\hat{T}_\varepsilon : X_0 \to X_0$ as before. To prove our claim, it suffices to show that $|\hat{T}_\varepsilon - T_0| \to 0$ as $\varepsilon \to 0$.

Now for $x \in x_0$, with $|x| = 1$,

$$|\hat{T}_\varepsilon x - T_0 x| \leq |\pi_0| \cdot (|T_\varepsilon x - T_0 x| + |T_\varepsilon S_\varepsilon x|)$$
$$\leq |\pi_0| \cdot (C_1^*(\varepsilon) + |T_\varepsilon| \cdot |S_\varepsilon x|),$$

and it only remains to show that $|S_\varepsilon x| \to 0$. This is true because

$$|S_\varepsilon x| \leq \|S_\varepsilon\|_{\kappa^N} \|x\|_{\kappa^N}$$
$$\leq \|S_\varepsilon\|_{\kappa^N} (\kappa^N \cdot \text{const} \cdot |x| + (1 - \kappa^N) \cdot |x|). \quad \square$$

Proof of Theorem 3. Obviously we wish to apply the results above to $T_0 = \mathscr{L}$, $T_\varepsilon = \mathscr{L}_\varepsilon$, $X = BV$ etc. We will indicate how to prove assertion (3). Let $\Theta < \Theta' < \Theta''$ be such that Θ'' is arbitrarily near Θ. We let

$$\Sigma_{1,1} = \{z \in \sigma(\mathscr{L}) : |z| \leq \Theta'\},$$
$$\Sigma_{1,0} = \{z \in \sigma(\mathscr{L}) : \Theta' < |z| < \sqrt{\Theta''}\},$$
$$\Sigma_0 = \{z \in \sigma(\mathscr{L}) : |z| \geq \sqrt{\Theta''}\}.$$

and choose $\kappa = \Lambda$ near $\sqrt{\Theta''}$ such that $\Theta' < \kappa^2 < \kappa\sqrt{\Theta''}$. The norm of $\mathscr{L}_\varepsilon : L^1 \to L^1$ is equal to 1, and it follows from Lemma 9 that $\mathscr{L}_\varepsilon$ is quasi-compact so that $\varrho_\varepsilon \in BV$ (see e.g. Keller [1982, p. 315]). Theorem 3 hence follows from Lemma 3′ and the results stated in Sect. 5.A and 5.B.

(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].) $\square$

Proof of Theorem 3′. Again we prove (3). We let $\kappa_{11} \leq \Theta'$ be as above. Here, however, we consider only $\kappa_0 > \sqrt{2\Theta''}$ and let $\kappa = \Lambda$ be very slightly smaller than $\kappa_0/2$. Then $\Theta < 2\kappa^2 < \kappa$ which is the hypothesis of Lemma 9′. Lemma 9′ does not yield (A′.2) but only the weaker bound

$$\|T_\varepsilon^N - T_0^N\|_{\kappa^N} \leq (2\kappa)^N.$$

However, since we can assume that the constant κ' in the proof of Lemma 1′ satisfies $\kappa_{11} < \kappa < 2\kappa < \kappa' < \kappa_0$, we obtain an improved version of (5.15):

$$\|R(T_0^N, \lambda^N)\|_{\kappa^N} \leq \frac{\text{const}(\|\pi_0\|_{\kappa^N} + \|\pi_{1,0}\|_{\kappa^N} + \|\pi_1\|_{\kappa^N})}{(\kappa')^N} \leq \frac{1}{(2\kappa)^N}.$$

The other requirement on κ in (A′.2), namely that $\kappa_{11} < \kappa\kappa_0$, is also satisfied. The conclusion of Lemma 1′ is thus still valid. (The proof of Lemma 2 can be modified in a similar fashion.) We finish as in Theorem 3.

If the functions θ_ε are symmetric, we can replace each factor 2 by 3/2 in the above choices. $\square$

References

1. Baladi, V., Keller, G.: Zeta functions and transfer operators for piecewise monotone transformations. Commun. Math. Phys. **127**, 459–477 (1990)
2. Baxendale, P.: Brownian motions in the diffeomorphism group. Compositio Math. **53**, 19–50 (1984)
3. Benedicks, M., Young, L.-S.: Absolutely continuous invariant measures and random perturbations for certain one-dimensional maps. Ergodic Theory Dynamical Systems **12**, 13–37 (1992)
4. Blank, M.L.: Chaotic mappings and stochastic Markov chains. Mathematical Physics X. Proceedings of the X$^{\text{th}}$ Congress on Mathematical Physics, Leipzig, Germany, 1991, Schmüdgen, K. (ed.) Berlin, Heidelberg, New York: Springer 1992, pp. 341–345
5. Collet, P.: Ergodic properties of some unimodal mappings of the interval. Preprint Mittag-Leffler (1984)
6. Collet, P.: Some ergodic properties of maps of the interval. Dynamical Systems and Frustrated Systems. Bamon, R., Gambaudo, J.-M., Martinez, S. (eds.) 1991 (to appear)
7. Collet, P., Isola, S.: On the essential spectrum of the transfer operator for expanding Markov maps. Commun. Math. Phys. **139**, 551–557 (1991)
8. Coven, E.M., Kan, I., Yorke, J.A.: Pseudo-orbit shadowing in the family of tent maps. Trans. Am. Math. Soc. **308**, 227–241 (1988)
9. Franks, J.: Manifolds of $\mathscr{C}^r$ mappings and applications to differentiable dynamical systems. Studies in Analysis, Adv. Math. Suppl. Stud. **4**, 271–291 (1979)
10. Hofbauer, F., Keller, G.: Ergodic properties of invariant measures for piecewise monotonic transformations. Math. Z. **180**, 119–140 (1982)
11. Keller, G.: Stochastic stability in some chaotic dynamical systems. Monatsh. Math. **94**, 313–333 (1982)
12. Keller, G.: On the rate of convergence to equilibrium in one-dimensional systems. Commun. Math. Phys. **96**, 181–193 (1984)
13. Kifer, Y.: On small random perturbations of some smooth dynamical systems. Math. USSR-Izv. **8**, 1083–1107 (1974)
14. Kifer, Y.: Ergodic theory of random transformations. Boston, Basel: Birkhäuser 1986
15. Kifer, Y.: Random perturbations of dynamical systems. Boston, Basel: Birkhäuser 1988a
16. Kifer, Y.: A note on integrability of $\mathscr{C}^r$ norms of stochastic flows and applications. Stochastic Mechanics and Stochastic Processes, Proc. Conf. Swansea/UK 1986, Lecture Notes in Math. **1325**, Berlin, Heidelberg, New York: Springer 1988b, pp. 125–131
17. Kunita, H.: Stochastic flows and stochastic differential equations. Cambridge: Cambridge University Press 1990
18. Mañé, R.: Ergodic theory and differentiable dynamics. Berlin, Heidelberg, New York: Springer 1987
19. Mayer, D.: On a ζ function related to the continued fraction transformation. Bull. Soc. Math. France **104**, 195–203 (1976)
20. Ruelle, D.: Zeta functions for expanding maps and Anosov flows. Invent. Math. **34**, 231–242 (1976)
21. Ruelle, D.: Locating resonances for Axiom A dynamical systems. J. Stat. Phys. **44**, 281–292 (1986)
22. Ruelle, D.: The thermodynamic formalism for expanding maps. Commun. Math. Phys. **125**, 239–262 (1989)
23. Ruelle, D.: An extension of the theory of Fredholm determinants. Inst. Hautes Études Sci. Publ. Math. **72**, 175–193 (1990)
24. Rychlik, M.: Bounded variation and invariant measures. Studia Math. **LXXVI**, 69–80 (1983)
25. Wilkinson, J.H.: The Algebraic Eigenvalue Problem. London: Oxford University Press 1965
26. Wong, S.: Some metric properties of piecewise monotonic mappings of the unit interval. Trans. Am. Math. Soc. **246**, 493–500 (1978)
27. Yosida, K.: Functional Analysis (Sixth Edition). Grundlehren der mathematischen Wissenschaften **123**. Berlin, Heidelberg, New York: Springer 1980

Communicated by J.-P. Eckmann