Strong stochastic stability and rate of mixing for unimodal maps

Viviane Baladi and Marcelo Viana

December 1994

Abstract. We consider small random perturbations of a large class of nonuniformly hyperbolic unimodal maps and prove stochastic stability in the strong sense (L^1-convergence of invariant densities) and uniform bounds for the exponential rate of decay of correlations. Our method is based on an analysis of the spectrum of a modified Perron-Frobenius operator for a tower extension of the Markov chain.

1. Introduction

Let $I \subset \mathbb{R}$ be a compact interval and $f : I \to I$ be a smooth unimodal map with $f(I) \subset \operatorname{int}(I)$. The prototype we have in mind are the quadratic maps $f(x) = -x^2 + a$ but our arguments and conclusions hold in the general context of maps with negative Schwarzian derivative and nondegenerate critical point. Let $c \in I$ be the critical point of f and $c_k = f^k(c)$ for $k \geq 0$. Throughout this paper we assume that

(A1) $|f^k(c) - c| \geq e^{-\alpha k}$ for all $k \geq H_0$,
(A2) $|(f^k)'(c_1)| \geq \lambda_c^k$ for all $k \geq H_0$,
(A3) f is topologically mixing on the interval bounded by c_1 and c_2,

where $H_0 \geq 1$, $1 < \lambda_c < 2$, and $0 < \alpha$ with $e^{2\alpha} < \sqrt{\lambda_c}$ are fixed constants.

Conditions (A1), (A2) are inspired by Benedicks-Carleson [BC], where it is proved that they are satisfied by quadratic maps for a positive measure set of values of the parameter a. Moreover, they imply the conclusion of Jakobson's theorem [Ja]: The map admits a (unique) invariant Borel probability measure m_0 which is absolutely continuous with respect to Lebesgue measure on I. This invariant measure is ergodic and describes the typical asymptotics of orbits of f, in the sense that $\frac{1}{n}\sum_{j=0}^{n-1}\delta_{f^j(x)} \to m_0$ for Lebesgue almost all $x \in I$. Assumption (A3) is used only in Section 5 and we discuss it there (quadratic maps satisfy all three conditions simultaneously, for a positive measure set of values of a).

Our purpose is to show that (A1)–(A3) ensure stability of the dynamics under random perturbations of the map: The asymptotics are only slightly affected when one replaces f^n by $(f + t_n) \circ \cdots \circ (f + t_1)$, with $t_1, \dots, t_n$ chosen at random in a small interval $[-\epsilon, \epsilon]$ following some probability distribution θ_ϵ. This contrasts with the structural instability

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

of these maps: For g arbitrarily close to f the asymptotic behaviour of g^n may be very different from that of f^n (e.g. it may be of periodic type).

Stability under random perturbations may be expressed more precisely as follows. For each small $\epsilon > 0$ we consider the Markov chain χ^ϵ on the σ-algebra of Borel subsets of I whose transition probabilities are given by $P^\epsilon(x, E) = \int_E \theta_\epsilon(y - fx)\, dy$. Our conditions on the probability density θ_ϵ are stated in (2.2)–(2.4). Then (see Section 2) for each $\epsilon > 0$ there exists a unique probability measure m_ϵ which is stationary under χ^ϵ, i.e.,

$$m_\epsilon(E) = \int P^\epsilon(x, E)\, dm_\epsilon(x) \text{ for every Borel set } E.$$

Moreover, m_ϵ is absolutely continuous with respect to Lebesgue measure and satisfies $\frac{1}{n}\sum_{j=0}^{n-1} \delta_{x_j} \to m_\epsilon$ for most random trajectories $x_j = (f + t_j) \circ \cdots \circ (f + t_1)(x)$.

We want to call f stochastically stable if these asymptotic distributions m_ϵ converge to the invariant probability m_0 of f as the noise level ϵ goes to zero. More precisely, we say that f is *weakly stochastically stable* under χ^ϵ if $m_\epsilon \to m_0$ in the weak*-topology. This is the same as having $\rho_\epsilon \to \rho_0$ in the weak sense, where ρ_0 and ρ_ϵ are the Radon-Nikodym derivatives of m_0 and m_ϵ with respect to Lebesgue measure. We say that f is *strongly stochastically stable* under χ^ϵ if $m_\epsilon \to m_0$ in the strong (norm) topology or, equivalently, if ρ_ϵ converges to ρ_0 in $L^1(dx)$.

Obviously, every strongly stable system is also weakly stable. A simple example of a sequence of functions in $[0, 1]$ which is weakly convergent but not L^1-convergent is $g_n(x) = (-1)^{[nx]}$, where $[z]$ is the integer part of z. This example illustrates a main advantage of strong stochastic stability over its weak analog: preventing large oscillations of the ρ_ϵ around the limit density ρ_0. For uniformly bounded sequences of functions having uniformly bounded variation, it is not difficult to check that weak convergence implies strong convergence. This provides a (very partial) explanation for the role of the variation in the theorem below.

Another important stochastic parameter we analyse here is the exponential rate of decay of correlations, which measures the mixing character of the dynamics. Let $\mathcal{F}$ be some Banach space of test functions on I (we shall always consider $\mathcal{F} = BV(I)$, the space of functions with bounded variation). We say that (f, m_0) has *exponential decay of correlations* in $\mathcal{F}$ if there exists $0 < \tau < 1$ and for any $\varphi, \psi \in \mathcal{F}$ there exists some $C = C(\tau, \|\varphi\|, \|\psi\|) > 0$ satisfying

$$\left| \int (\varphi \circ f^n)\psi\, dm_0 - \int \varphi\, dm_0 \int \psi\, dm_0 \right| \le C\tau^n \qquad \text{for all } n \ge 1.$$

Then the *rate of decay of correlations of* (f, m_0) in $\mathcal{F}$ is the infimum τ_0 of all such numbers τ. Analogously, we define the *rate of decay of correlations of* $(\chi^\epsilon, m_\epsilon)$ in $\mathcal{F}$ to be the infimum τ_ϵ over all $\tau > 0$ such that

$$\left| \int \left(\int \varphi(y) P_n^\epsilon(x, dy) \right) \psi(x)\, dm_\epsilon(x) - \int \varphi\, dm_\epsilon \int \psi\, dm_\epsilon \right| \le C\tau^n \text{ for all } n \ge 1,$$

with $C = C(\tau, \|\varphi\|, \|\psi\|)$, and where $P_n^\epsilon(x, dy)$ denotes the n-step transition probability.

We shall now state our main result. Here we call a differentiable map $f : I \to I$ unimodal if it has a unique critical point c and $c \in \text{int}\,(I)$. We take f to be C^4 and to have Schwarzian derivative $Sf < 0$, recall that $Sf = (f'''/f') - (3/2)(f''/f')^2$. We also let c be nondegenerate, i.e., $f''(c) \neq 0$ (but our arguments may be adapted easily to the case when c is only nonflat, meaning $f^{(2\ell)}(c)$ exists and is nonzero for some $\ell \geq 1$). Finally, we suppose $f(I) \subset \text{int}\,(I)$ and that f admits an extension to some compact interval $J \supset I$, preserving all the previous properties and satisfying $f(\partial J) \subset \partial J$.

Main Theorem. *Let $f : I \to I$ be a unimodal map with negative Schwarzian derivative and nondegenerate critical point as above, and let $(\chi^\epsilon)_\epsilon$ be random perturbations of f as introduced before. If f satisfies (A1)–(A3) then*

1. *(Strong stochastic stability.) The density ρ_ϵ of the unique invariant probability measure m_ϵ of χ^ϵ converges in $L^1(dx)$ to the density ρ_0 of the unique absolutely continuous invariant probability measure m_0 of f.*
2. *(Uniform rates of decay of correlations.) The systems $(\chi^\epsilon, m_\epsilon)$ and (f, m_0) have exponential decay of correlations in the space $BV(I)$ of functions with bounded variation, and their rates of decay are uniformly bounded: There exists $\bar\tau < 1$ depending only on f such that $\tau_\epsilon \leq \max(\sqrt{\tau_0}, \bar\tau) < 1$ for small enough $\epsilon > 0$.*

Stochastic stability and decay of correlations have been investigated for many dynamical systems, see e.g. Kifer [Ki2] and references therein. Let us focus on quadratic maps. Katok-Kifer [KK] proved weak stochastic stability under a uniform hyperbolicity assumption (nonrecurrence of the critical point). Then Benedicks-Young [BY1] showed that a large set of nonuniformly hyperbolic maps are weakly stochastically stable (they use a different form of assumptions (A1)–(A3) above). In fact, abundance of stochastic stability (in the strong sense) among nonuniformly hyperbolic quadratic maps had also been obtained in an unpublished work of Collet [Co]. Exponential decay of correlations was proved independently by Keller–Nowicki [KN] and by Young [Yo], for classes of nonuniformly hyperbolic maps related to ours.

Our basic approach in the proof of the main theorem is inspired by Baladi-Young [BaY] who obtained similar results for some uniformly hyperbolic systems. Indeed, we introduce certain transfer operators $\mathcal{L}_0$ and $\mathcal{L}_\epsilon$ associated with f and χ_ϵ, respectively, and derive the statements in the theorem from showing that these operators are quasi-compact (the peripheral spectrum is discrete or, in precise terms, the essential spectral radius is strictly smaller than the spectral radius) and that $\mathcal{L}_\epsilon$ is "close" to $\mathcal{L}_0$ for small $\epsilon > 0$. As a by-product, this method permits us to recover and unify in the present setting many of the results mentioned previously, including the existence of absolutely continuous invariant measures [Ja] and the exponential decay of correlations [KN, Yo]. We also expect it to be useful in more general situations, e.g. for higher-dimensional systems such as those in [BC], [BY2].

Let us sketch in more detail how this basic strategy will be carried out, describing the main new ingredients necessary in the present situation to overcome the lack of hyperbolicity. In Section 2, we construct a tower extension $\hat f : \hat I \to \hat I$ for the map f.

Towers are now a standard tool in 1-dimensional dynamics and were also used, e.g., in [KN, Yo]. However, neither of these constructions can be used directly in a random setting such as ours: Our tower must also support extensions $\hat{\chi}^\epsilon$ of the Markov chains χ^ϵ. In Section 2 we also introduce transfer operators $\mathcal{L}_0$ and $\mathcal{L}_\epsilon$, acting on a Banach space $BV(\hat{I})$ of functions of bounded variation. For the definition of $\mathcal{L}_0$ we must use a convenient cocycle $w_0 \colon \hat{I} \to [0, \infty)$:

$$\mathcal{L}_0 \varphi(x) = \frac{1}{w_0(x)} \sum_{\hat{f}(y)=x} w_0(y) \frac{\varphi(y)}{|\hat{f}'(y)|}$$

(this corresponds to a change of coordinates and is required to remove the poles of $1/|\hat{f}'|$ and to enforce the expansion during the "recovery" phases of orbits). Perturbed cocycles w_ϵ and perturbed operators $\mathcal{L}_\epsilon$, corresponding to $\hat{\chi}^\epsilon$, are also defined, involving averages over past (random) orbits. This seems to be the first time that perturbed cocycles are introduced.

Building on several preliminary results obtained in Section 3, we derive our main estimates in Section 4. We show that $\mathcal{L}_0$ satisfies a Lasota-Yorke [LY] type inequality, i.e., that there are $C > 0$ and $\sigma > 1$ such that for all $n \geq 0$,

$$\operatorname{var} \mathcal{L}_0^n \varphi \leq C \sigma^{-n} (\operatorname{var} \varphi + \sup |\varphi|) + C \int |\varphi| w_0 \, dx \, .$$

Estimates of this type are also central to [KN] and [Yo]. We also prove a similar fact for $\mathcal{L}_\epsilon$. Combined with our other bounds, this yields that $\mathcal{L}_\epsilon$ is close to $\mathcal{L}_0$ in the following sense: There are $C > 0$ and $\bar{\tau} < 1$, and for each $n \geq 1$ there are $\epsilon(n) > 0$ and a norm $\| \cdot \|_{(n)}$, such that $\|\mathcal{L}_\epsilon^n - \mathcal{L}_0^n\|_{(n)} \, l \leq C \bar{\tau}^n$ for $\epsilon \leq \epsilon(n)$.

Ergodic properties of our systems may then be deduced from the accumulated knowledge on these operators. This is done in Section 5, and follows well-known lines. First, if $\hat{\rho}_0$ is a (normalized) fixed function of $\mathcal{L}_0$ then $\hat{m}_0 = w_0 \hat{\rho}_0 \, dx$ is an absolutely continuous invariant probability measure for $\hat{f}$, and it projects down to the invariant measure m_0 of f. Moreover, after lifting the correlation functions to the tower, one sees that the gap in the spectrum of $\mathcal{L}_0$ separating 1 from the second largest eigenvalue is directly related with the rate of decay of correlations of the system (f, m_0). Similar statements hold for positive ϵ. Finally, using the above closeness between $\mathcal{L}_0$ and $\mathcal{L}_\epsilon$, and applying nonstandard perturbation results from [BaY], we obtain the claims in our main theorem.

Acknowledgements. V. Baladi is grateful to IMPA/CNPq-Brazil and M. Viana is grateful to FIM/ETH-Switzerland for their kind hospitality and support during the preparation of this paper.

2. The tower

Throughout, the notation C represents a generic (large) positive constant, and $C(\cdot)$, respectively $c(\cdot)$, is some positive function tending to infinity, respectively tending to zero, with its argument. We also use $C_n(\cdot)$, respectively $c_n(\cdot)$, to denote a sequence of positive functions which, for each fixed value of n, converge to infinity, respectively to zero, with the argument.

We make frequent use of the following easy inequalities. Let I, J be compact intervals, $\psi_1, \psi_2 : I \to \mathbb{C}$ be functions of bounded variation, and $h : I \to J$ be a homeomorphism.

(a) $\operatorname{var}_I(\psi_1 + \psi_2) \leq \operatorname{var}_I \psi_1 + \operatorname{var}_I \psi_2$,
(b) $\operatorname{var}_I(\psi_1 \cdot \psi_2) \leq \operatorname{var}_I \psi_1 \sup_I |\psi_2| + \sup_I |\psi_1| \operatorname{var}_I \psi_2$,
(c) $\operatorname{var}_I(\psi_1 \circ h) = \operatorname{var}_J \psi_1$,
(d) $\operatorname{var}_I(\psi_1 \chi_{I'}) \leq \operatorname{var}_I \psi_1 + 2 \sup_{I'} |\psi_1|$ for each interval $I' \subset I$.

Moreover, given $\phi : I \times J \to \mathbb{C}$, $\psi : J \to [0, +\infty)$ such that $\phi(x, \cdot)\psi(\cdot) \in L^1(J, dx)$ for each fixed x, and $\phi(\cdot, t)$ has bounded variation for each fixed t, then

(e) $\operatorname{var}_I \int \phi(x,t)\psi(t)\, dt \leq \int (\operatorname{var}_I \phi(x,t))\psi(t)\, dt$.

The dynamics.

We always take $f : I \to I$ to be as in Section 1. Without any restriction, we take the critical point c to coincide with zero; sometimes we denote $a = c_1$. Condition (A3) will not be used until Section 5. We fix $\lambda > 1$ and $\rho > e^\alpha$ so that

$$e^\alpha \lambda \rho < \sqrt{\lambda_c}\,. \tag{2.1}$$

Other constants $0 < \delta \ll \alpha$ and $1 < \sigma \leq \lambda$ will be introduced later on.

Now we fix some small ϵ_0 so that $f_t(I) \subset \operatorname{int}(I)$ for all $|t| \leq \epsilon_0$. Here $f_t(x) = f(x) + t$, and we also write $f^n_{\vec{t}} = f^n_{t_n \cdots t_1} = f_{t_n} \circ \cdots \circ f_{t_1}$ for each $n \geq 1$ and $\vec{t} = (t_1, \ldots, t_n)$. As explained before, we are interested in Markov chains χ^ϵ, with $0 < \epsilon < \epsilon_0$, whose transition probabilities $P^\epsilon(x, \cdot)$ have densities $\theta_\epsilon(y - fx)$. Each θ_ϵ is a probability distribution on $[-\epsilon, \epsilon]$, i.e., a nonnegative function with

$$\operatorname{supp} \theta_\epsilon \subset [-\epsilon, \epsilon] \qquad \text{and} \qquad \int \theta_\epsilon(x)\, dx = 1\,. \tag{2.2}$$

We also assume the θ_ϵ to satisfy

$$M = \sup_\epsilon (\epsilon \sup |\theta_\epsilon|) < \infty \tag{2.3}$$

and, denoting $J_\epsilon = \{t \mid \theta_\epsilon(t) > 0\}$,

$$J_\epsilon \text{ is an interval containing } 0 \text{ and } \phi_\epsilon = \log(\theta_\epsilon|_{J_\epsilon}) \text{ is concave.} \tag{2.4}$$

The technical condition (2.4) is introduced here in order to simplify some of our arguments, a weaker regularity assumption should suffice. In any case, it holds in most interesting cases, e.g. Gaussian and uniform distributions. Note that (2.2)–(2.4) are automatic if θ_ϵ has the form $\theta_\epsilon(t) = (1/\epsilon)\theta(t/\epsilon)$ for some θ satisfying (2.2) and (2.4).

Clearly, ϕ_ϵ is concave if $(\theta_\epsilon|_{J_\epsilon})$ is concave. On the other hand, $(\theta_\epsilon|_{J_\epsilon})$ is at most two-to-one if ϕ_ϵ is concave: Otherwise, there would be a point y with at least three preimages by θ_ϵ, and therefore there would be some z with at least three preimages by ϕ_ϵ, so that ϕ_ϵ would have to be constant, a contradiction.

It follows from our assumptions that, for all small enough ϵ, the Markov chain χ^ϵ has a unique invariant probability measure m_ϵ (we do not need (A3) for this) and this measure is absolutely continuous with respect to Lebesgue. (See [BY1, Part II] for a proof of uniqueness. We do not assume, as they do, that θ_ϵ is bounded from below, but if ϵ is small enough it is still true that an invariant measure for χ^ϵ must contain $c = 0$ in the interior of its support.) Uniqueness also implies that m_ϵ satisfies an ergodicity property: The product measure $m_\epsilon \times \theta_\epsilon^{\mathbb{N}}$ is ergodic (and invariant) with respect to the map on $I \times \mathbb{R}^{\mathbb{N}}$ defined by $(x, t_1, t_2, \dots) \mapsto (f_{t_1}(x), t_2, t_3, \dots)$, see [Ki1, Theorem 2.1]. It follows, using the ergodic theorem, that the Birkhoff averages of random trajectories $x_j = f_{t_j \dots t_1}(x)$ converge to m_ϵ for (Lebesgue) almost every $(x, t_1, \dots, t_j, \dots) \in \operatorname{supp} m_\epsilon \times \operatorname{supp} \theta_\epsilon^{\mathbb{N}}$, as already mentioned in the Introduction.

The tower.

We now construct a tower extension $\hat f : \hat I \to \hat I$ of f, as well as its deterministic perturbations $\hat f_t$, for $|t| < \epsilon < \epsilon_0$. Let $\alpha < \beta_1 < \beta_2 < 2\alpha$ be two constants; note that (2.1) above implies $e^{\beta_i/2}\lambda\rho < \sqrt{\lambda_c}$ for $i = 1, 2$. The *tower* $\hat I$ is the union $\hat I = \cup_{k \ge 0} E_k$ of levels $E_k = B_k \times \{k\}$ satisfying the following properties. The ground floor interval $B_0 = [a_0, b_0]$ is just the interval I. For $k \ge 1$, the interval $B_k = [a_k, b_k]$ is such that

$$[c_k - e^{-\beta_2 k}, c_k + e^{-\beta_2 k}] \subset B_k \subset [c_k - e^{-\beta_1 k}, c_k + e^{-\beta_1 k}] \, .$$

Observe that $0 = c \notin B_k$ for all $k \ge H_0$, where H_0 is given by (A1). For future use we introduce $B_0^+ = [0, b_0]$, $B_0^- = [a_0, 0]$ and $E_0^\pm = B_0^\pm \times \{0\}$.

Now we fix some small $\delta > 0$, in particular, we assume that

$$|f^j(x) - c_j| < \min\{|c_j| e^{-\alpha j}, \, e^{-\beta_2 j}\} \qquad \text{for all } 1 \le j \le H_0 \text{ and } |x| \le \delta \tag{2.5}$$

(the other conditions on δ are stated later in this subsection, in Lemmas 1 and 2, in the proof of the Sublemma, cf. (4.15), and in the proof of Corollary 2, cf. (5.1)). Then we also set $\tilde B_0^\pm = B_0^\pm \cap (-\delta, \delta)$ and $\tilde E_0^\pm = \tilde B_0^\pm \times \{0\}$.

Given $x \ne c$ we shall denote by x_- the unique point $x_- \ne x$ with $f(x_-) = f(x)$. It is no restriction to suppose that there is an uncountable set of arbitrarily small values of $\delta > 0$ for which $(-\delta)_- \le \delta$ (just change coordinates $x \mapsto -x$ otherwise) and we do so. Let us write $\{0, \pm\delta\} \cup \{a_j \mid j \ge 0\} \cup \{b_j \mid j \ge 0\} = \{e_0 = 0, e_1 = \delta, e_2 = -\delta, e_3, \dots\}$. We may, and do, require additionally that for all $j \ge 1$ and $k, \ell \ge 0$

$$f^j(e_k) \ne e_\ell \, . \tag{2.6}$$

Indeed, our assumptions on f imply that $f^j(0) \ne 0$ for all $j \ge 1$. We choose δ such that $\pm\delta$ do not belong to the critical orbit, and that $f^j(\pm\delta) \notin \{0, \pm\delta\}$ for all $j \ge 1$ (these are co-countable conditions). For each $\ell \ge 3$ we impose the co-countable conditions that

e_ℓ is not f-periodic, $e_\ell \notin \cup_j\{f^j(e_0),\dots,f^j(e_{l-1})\}$ and $f^j(e_\ell)\notin\{e_0,\dots,e_{\ell-1}\}$ for all $j\ge 1$.

For $(x,k)\in E_k$ and $|t|<\epsilon$ we set

$$\hat f_t(x,k)=\begin{cases}(f_t(x),k+1) & \text{if } k\ge 1 \text{ and } f_t(x)\in B_{k+1}\,,\\ (f_t(x),k+1) & \text{if } k=0 \text{ and } x\in(-\delta,\delta)\,,\\ (f_t(x),0) & \text{otherwise,}\end{cases}$$

and we define $\hat f^n_{t_n\dots t_1}$ as above. (We write $\hat f=\hat f_0$.) Denoting $\pi:\hat I\to I$ the projection to the first factor we have $f_t\circ\pi=\pi\circ\hat f_t$ on $\hat I$.

Define $H(\delta)=H(\delta,\epsilon_0)$ to be the minimal $k\ge 1$ such that there exist some $x\in(-\delta,\delta)$ and some $\vec t=(t_1,\dots,t_k,t_{k+1})\in J^{k+1}_{\epsilon_0}$ such that $\hat f^{k+1}_{\vec t}(x,0)\in E_0$. We observe that $H(\delta)$ can be made arbitrarily large by choosing small enough δ and ϵ_0 (by continuity). In particular, we assume that $H(\delta)\ge\max(2,H_0)$, cf. (2.5). We define the Markov chain $\hat\chi^\epsilon$ by considering the transition probabilities $\hat P^\epsilon((x,k),E)=\sum_{j=0}^\infty\int_{\pi E}\hat\theta_\epsilon((y,j),\hat f(x,k))\,dy$ where $\hat\theta_\epsilon((y,j),\hat f(x,k))=0$ if $\hat f_{y-fx}(x,k)\notin E_j$, and $\hat\theta_\epsilon((y,j),\hat f(x,k))=\theta_\epsilon(y-fx)$ otherwise (in which case $\hat f_{y-fx}(x,k)=(y,j)$; when there is no ambiguity, in particular when $j=k+1$, we simply write $\theta_\epsilon(y-fx)$).

The cocycles.

We wish to consider transfer operators $\mathcal L$ and $\mathcal L_\epsilon$ related to the (unique) absolutely continuous invariant probability measure of $\hat f$ and each $\hat\chi^\epsilon$. For this, it is useful to introduce cocycles in order to suppress the singularity of the weights $1/|f_t'|$.

We first give the definition of the unperturbed cocycle $w=w_0:\hat I\to\mathbb R$. If $k>\ell\ge 1$ then for each $(x,k)\in E_k\cap\mathrm{Im}\,(\hat f^\ell)$ there is a unique $(y,k-\ell)\in E_{k-\ell}$ such that $\hat f^\ell(y,k-\ell)=(x,k)$ and $f^j(y)$ has the same sign as $c_{k-\ell+j}$ for $0\le j<\ell$ (the second condition is needed only if $k-\ell<H_0$). We write $\hat f_+^{-\ell}(x,k)=y$. If $(x,k)\in E_k\cap\mathrm{Im}\,(\hat f^k)$ we also define $\hat f_+^{-k}(x,k)=y$ where $(y,0)$ is the unique point in $\tilde E_0^+$ with $\hat f^k(y,0)=(x,k)$. We set

$$w_0(x,k)=\begin{cases}\frac{\lambda^k}{|(f^k)'(\hat f_+^{-k}(x,k))|} & \text{if } (x,k)\in\mathrm{Im}\,(\hat f^k)\,,\\ 0 & \text{otherwise.}\end{cases}$$

(In particular, $w_0(x,0)\equiv 1$.) Note, for further use, that the support of the cocycle w_0 in E_k is an interval for each $k\ge 1$, with endpoints in the set $\partial E_k\cup\{\hat f^k(0,0),\hat f^k(\pm\delta,0)\}$. For $k\ge 1$ and $(x,k)\in\mathrm{Im}\,\hat f$ we have $w_0(x,k)=\lambda w_0(\hat f_+^{-1}(x,k),k-1)/|f'(\hat f_+^{-1}(x,k))|$.

The perturbed cocycle w_ϵ is defined by:

$$w_\epsilon(x,k)=\begin{cases}1 & \text{if } k=0\,,\\ \lambda\int_{\tilde B_0^+}\theta_\epsilon(x-fy)\,dy & \text{if } k=1\,.\\ \lambda\int_{B_{k-1}}w_\epsilon(y,k-1)\theta_\epsilon(x-fy)\,dy & \text{if } k\ge 2\,.\end{cases}$$

Defining $(x_{t_\ell,\cdots,t_1}, k-\ell) = \hat{f}^{-\ell}_{t_\ell\cdots t_1,+}(x,k)$, with $k \geq \ell \geq 1$, as in the unperturbed case,

$$w_\epsilon(x,1) = \lambda \int |f'(x_t)|^{-1}\theta_\epsilon(t)\, dt\,,$$
$$w_\epsilon(x,k) = \lambda \int w_\epsilon(x_t, k-1)|f'(x_t)|^{-1}\theta_\epsilon(t)\, dt\,, \text{ for } k \geq 2\,,$$

(integration is over the t such that x_t is defined, with $x_t \in [0,\delta)$ and $x_t \in B_{k-1}$, respectively). Introducing the notation $d\vec{\theta}_\epsilon(\vec{t}) = \theta_\epsilon(t_1)\cdots\theta_\epsilon(t_{k-1})\, dt_1 \ldots dt_{k-1}$, we also have for $k \geq 2$

$$w_\epsilon(x,k) = \lambda^{k-1} \int w_\epsilon(x_{t_{k-1}\cdots t_1}, 1)\, |(f^{k-1}_{t_{k-1}\ldots t_1})'(x_{t_{k-1}\cdots t_1})|^{-1}\, d\vec{\theta}_\epsilon(\vec{t})\,,$$

where the integral is over the $\vec{t} \in J_\epsilon^{k-1}$ such that $x_{t_{k-1}\cdots t_1} \in B_1$ exists.

Our assumptions imply that θ_ϵ converges to the Dirac function as ϵ tends to zero. It follows that $w_\epsilon(x,k)$ converges pointwise to $w_0(x,k) = w(x,k)$ as $\epsilon \to 0$. Moreover, for small enough ϵ, and for all $k \geq 0$, the support of w_ϵ in E_k is an interval with endpoints close to the endpoints of the support of $w = w_0$ in E_k. Writing dx for Lebesgue measure on $\hat{I}$, we introduce the positive measures $\mu_0 = w_0\, dx$ and $\mu_\epsilon = w_\epsilon\, dx$. It will follow from our analysis, e.g. the proof of Lemma 7, that these measures are finite.

We use the cocycles w_ϵ to define nonnegative weights g_t on $\hat{I}$, for $0 \leq t < \epsilon$, by

$$g_t(y,k) = \frac{w_\epsilon(y,k)}{w_\epsilon(\hat{f}_t(y,k))}\frac{1}{|f'(y)|}$$

(if the denominator is nonzero, otherwise we leave $g_t(y,k)$ undefined). Note that whenever $w_0(y,k) \neq 0$ we have $g_0(y,k) = (\lambda^k/|(f^k)'(\hat{f}_+^{-k}(y,k))|)\cdot(1/|f'(y)|)$ if $\hat{f}(y,k) \in E_0$, $g_0(y,k) = (|f'(y_-)|/\lambda|f'(y)|)$ if $(y,k) \in \tilde{E}_0^-$, and $g_0(y,k) = 1/\lambda$ in all other cases. We shall use the notation $g = g_0$, $g^{(n)} = \prod_{j=0}^{n-1}(g \circ \hat{f}^j)$, and similarly for $g_{\vec{t}}^{(n)}$.

The transfer operators.

Now we introduce a linear transfer operator $\mathcal{L} = \mathcal{L}_0$ acting on functions $\varphi : \hat{I} \to \mathbb{C}$ as follows. For $k \geq 1$ let $(\tilde{a}_k, k)$, $(\tilde{b}_k, k)$ be the endpoints of the interval $\operatorname{Im} \hat{f}^k \cap E_k$, with $\tilde{a}_k < \tilde{b}_k$. Given (x,k) such that either $k = 0$ or $k \geq 1$ with $\tilde{a}_k < x < \tilde{b}_k$ (in both cases $w(x,k) \neq 0$), we set

$$\mathcal{L}\varphi(x,k) = \frac{1}{w(x,k)} \sum_{\hat{f}(y,j)=(x,k)} \frac{\varphi(y,j)w(y,j)}{|f'(y)|} = \sum_{\hat{f}(y,j)=(x,k)} \varphi(y,j)g(y,j)\,.$$

Moreover, we set $\mathcal{L}\varphi(x,k) = \limsup_{y\uparrow\tilde{b}_k} \Re\mathcal{L}\varphi(y,k) + i \limsup_{y\uparrow\tilde{b}_k} \Im\mathcal{L}\varphi(y,k)$, if $k \geq 1$ and $x \geq \tilde{b}_k$, and similarly if $k \geq 1$ and $x \leq \tilde{a}_k$.

Analogously, denoting $(\tilde{a}_k^\epsilon, k)$, $(\tilde{b}_k^\epsilon, k)$ the endpoints of $\left(\cup_{\vec{t}\in J_\epsilon^k} \mathrm{Im}\, \hat{f}_{\vec{t}}^k\right) \cap E_k$, we define

$$\mathcal{L}_t\varphi(x,k) = \frac{1}{w_\epsilon(x,k)} \sum_{\hat{f}_t(y,j)=(x,k)} \frac{\varphi(y,j)w_\epsilon(y,j)}{|f'(y)|} = \sum_{\hat{f}_t(y,j)=(x,k)} \varphi(y,j)g_t(y,j)$$

and

$$\mathcal{L}_\epsilon\varphi(x,k) = \int \mathcal{L}_t\varphi(x,k)\,\theta_\epsilon(t)\,dt = \sum_{j\geq 0}\int_{B_j} \frac{1}{w_\epsilon(x,k)}\varphi(y,j)w_\epsilon(y,j)\hat{\theta}_\epsilon((x,k),\hat{f}(y,j))\,dy$$

if $k=0$ or $k\geq 1$ with $\tilde{a}_k^\epsilon < x < \tilde{b}_k^\epsilon$. For $k\geq 1$ and $x\notin(\tilde{a}_k^\epsilon, \tilde{b}_k^\epsilon)$ we define $\mathcal{L}_\epsilon\varphi(x,k)$ using limits in the same way as before.

We consider the Banach space $BV(\hat{I})$ of functions $\varphi : \hat{I}\to\mathbb{C}$ such that

$$\|\varphi\|_{BV} = \|\varphi\| = \sup_{\hat{I}}|\varphi| + \operatorname*{var}_{\hat{I}}\varphi + \int_{\hat{I}}|\varphi(x)|w_0(x)\,dx$$

is finite. It will follow from the results in Sections 3 and 4 that $\mathcal{L}$ and $\mathcal{L}_\epsilon$ are bounded operators on $BV(\hat{I})$ (in particular that $BV(\hat{I})$ is invariant) and, in fact, that they are quasicompact.

Intervals of monotonicity.

An interval $\eta\subset E_k$ for some $k\geq 0$ is called an *interval of monotonicity* for a map $\hat{F}:\hat{I}\to\hat{I}$ if the map $F=\pi\circ\hat{F}$ is monotone on η and if there is a j such that $\hat{F}(\eta)\subset E_j$. Let $\mathcal{Z}_0^n$ be the set of intervals of monotonicity of $\hat{f}_0^n$, i.e.,

$$\mathcal{Z}_0^n = \{\eta_1\cap\hat{f}_0^{-1}\eta_2\cap\cdots\cap\hat{f}_0^{-n+1}\eta_n \mid \eta_1,\ldots,\eta_n \text{ intervals of monotonicity of } \hat{f}_0\}\,.$$

Observe that property (2.6) from the definition of the tower implies that no element η_0 of $\mathcal{Z}_0^n$ is reduced to a point, and that η_0 is either disjoint from the support of the measure μ_0 or meets this support on an interval with nonempty interior (in the second case, $\mu_0(\eta_0)>0$).

Note that each level E_k contains at most three intervals of monotonicity of $\hat{f}_0$ for $k\geq H_0$, and at most four such intervals for $0\leq k<H_0$. Since, by definition, the image of an interval of monotonicity of $\hat{f}_0$ is always contained in some level E_j, we conclude that $\#\{\eta\in\mathcal{Z}_0^n \mid \eta\subset E_k\}\leq 4^n$ for all $k\geq 0$. For fixed values of n, we will need to consider monotonicity intervals corresponding to orbit pieces lying in a bounded part of the tower. Fixing $N\geq n$ we denote

$$\mathcal{Z}_0^{n,N} = \{\eta_1\cap\hat{f}_0^{-1}\eta_2\cap\cdots\cap\hat{f}_0^{-n+1}\eta_n\in\mathcal{Z}_0^n \mid \eta_i\subset\left(\cup_{k\leq N}E_k\right), 1\leq i\leq n\}\,.$$

The considerations above imply that $\#\mathcal{Z}_0^{n,N}\leq(N+1)4^n<\infty$ and that there is a constant $C_n(N)>0$ so that $|\eta|>1/C_n(N)$ for each nonempty $\eta\subset\mathcal{Z}_0^{n,N}$.

For $\vec{t} = (t_1, \dots, t_n) \in J^n_\epsilon$, let $\mathcal{Z}^n_{\vec{t}}$ be the set of intervals of monotonicity of $\hat{f}^n_{\vec{t}}$:

$$\mathcal{Z}^n_{\vec{t}} = \{\eta_1 \cap \hat{f}^{-1}_{t_1}\eta_2 \cap \cdots \cap (\hat{f}^{n-1}_{t_{n-1}\cdots t_1})^{-1}\eta_n \mid \eta_i \text{ monotonicity interval of } \hat{f}_{t_i}, 1 \le i \le n\},$$

and for $N \ge n$

$$\mathcal{Z}^{n,N}_{\vec{t}} = \{\eta_1 \cap \hat{f}^{-1}_{t_1}\eta_2 \cap \cdots \cap (\hat{f}^{n-1}_{t_{n-1}\cdots t_1})^{-1}\eta_n \in \mathcal{Z}^n_{\vec{t}} \mid \eta_i \subset \Big(\cup_{k\le N} E_k\Big), 1 \le i \le n\}.$$

Clearly, endpoints of nontrivial intervals in $\mathcal{Z}^n_{\vec{t}}$ and $\mathcal{Z}^{n,N}_{\vec{t}}$ vary continuously with $\vec{t}$. It follows that given any η_0 in $\mathcal{Z}^n_0$, for each $\vec{t}$ close enough to $\vec{0}$ there is $\eta(\vec{t}, \eta_0) \in \mathcal{Z}^n_{\vec{t}}$ with endpoints depending continuously of $\vec{t}$ and such that $\eta(\vec{0}, \eta_0) = \eta_0$. Moreover, there is $\epsilon(n, N) > 0$ such that for $\epsilon \le \epsilon(n, N)$ and any $\vec{t} \in J^n_\epsilon$, the map $\eta(\vec{t}, \cdot)$ sends $\mathcal{Z}^{n,N}_0$ bijectively to $\mathcal{Z}^{n,N}_{\vec{t}}$. For $\eta_0 \in \mathcal{Z}^{n,N}_0$ and $\epsilon < \epsilon(n, N)$ we define

$$\eta^+(\epsilon, \eta_0) = \bigcup_{\vec{t} \in J^n_\epsilon} \eta(\vec{t}, \eta_0) \text{ and } \eta^-(\epsilon, \eta_0) = \bigcap_{\vec{t} \in J^n_\epsilon} \eta(\vec{t}, \eta_0).$$

Then we have the uniform bounds $|\eta^+(\epsilon, \eta_0) \setminus \eta_0| \le c(\epsilon)$ and $|\eta_0 \setminus \eta^-(\epsilon, \eta_0)| \le c(\epsilon)$. Therefore for all $0 < \epsilon < \epsilon(n, N)$ and $\eta_0 \in \mathcal{Z}^{n,N}_0$ (reducing $\epsilon(n, N)$ if necessary) we have $|\eta^-(\epsilon, \eta_0)| \ge 1/C_n(N)$. It follows also from the above considerations that for any $0 < \epsilon < \epsilon(n, N)$, each point $z \in (\cup_{0 \le k \le N} E_k)$ is contained in no more than two $\eta^+(\epsilon, \eta_0)$ (we call this the *bounded overlap property*). Finally, for fixed $N \ge n$, the consequence of (2.6) mentioned above, the pointwise convergence of w_0 to w_ϵ, and the properties of the support of μ_0 imply that for all $\eta_0 \in \mathcal{Z}^{n,N}_0$ and all $\epsilon < \epsilon(n, N)$ (reducing $\epsilon(n, N)$ if necessary), either $\mu_\epsilon(\eta^+(\epsilon, \eta_0)) = 0$, or $\mu_\epsilon(\eta^-(\epsilon, \eta_0)) > 1/C_n(N)$.

3. Preliminary lemmas

In this section, we derive some preliminary lemmas on the objects introduced in Section 2. These lemmas will be used to prove our main bounds in the next section. Sometimes one may omit the hats (e.g. write f for $\hat{f}$) without ambiguity, and we do so.

The expansion constant σ.

Lemma 1. *There exist $\sigma > 1$, $b > 0$ and $\delta_0 > 0$ such that for any $0 < \delta < \delta_0$ there are $c(\delta) > 0$ and $\epsilon_0(\delta) > 0$ such that for any $n \ge 1$, $|t_1|, \dots, |t_n| < \epsilon_0(\delta)$ and $x \in I$:*

1. *if $x, f_{t_1}(x), \dots, f^{n-1}_{t_{n-1}\cdots t_1}(x) \notin (-\delta, \delta)$ then $\left|\left(f^n_{t_n\cdots t_1}\right)'(x)\right| \ge c(\delta)\sigma^n$;*
2. *if, in addition, $f^n_{t_n \dots t_1}(x) \in (-\delta, \delta)$ then $\left|\left(f^n_{t_n\cdots t_1}\right)'(x)\right| \ge b\sigma^n$.*

Proof of Lemma 1. We begin by noting that given $\delta_1 > 0$ there are $m \ge 1$, $\sigma_1 > 1$ and $\epsilon_1 > 0$ such that, for all $|t_1|, \dots, |t_m| < \epsilon_1$,

$$|(f^m_{t_m\cdots t_1})'(y)| \ge \sigma_1^m \quad \text{whenever} \quad y, f_{t_1}(y), \dots, f^{m-1}_{t_{m-1}\cdots t_1}(y) \notin (-\delta_1, \delta_1). \tag{3.1}$$

Indeed, by (A2) and [Si], all periodic points of f are repelling. Then e.g. [MS, Section III.3] implies (3.1) restricted to $t_i \equiv 0$. The full statement follows by choosing ϵ_1 small enough. In the sequel we fix $\delta_1 > 0$ small, depending only on H_0, α, and λ_c, see (3.4), (3.7), and (3.8). Now, there are $\sigma_2 > 1$, $\delta_2 > 0$, $K_2 \geq 1$ and $\epsilon_2 > 0$ such that, given any $1 \leq \ell < m$ and $|t_1|, \ldots, |t_\ell| < \epsilon_2$,

$$|(f^\ell_{t_\ell \cdots t_1})'(y)| \geq \frac{1}{K_2}\sigma_2^\ell \quad \text{whenever} \quad f^\ell_{t_\ell \cdots t_1}(y) \in (-\delta_2, \delta_2). \tag{3.2}$$

For $t_i \equiv 0$ this is a consequence of (A2), as proved by Nowicki [No]. The general case follows, once more, by continuity. Now we take $\sigma = \min\{\sigma_1, \sigma_2, \lambda\}$ and $\delta_0 = \min\{\delta_1, \delta_2\}$ and, for each $0 < \delta < \delta_0$, we define $c(\delta) = (\inf_{I \setminus (-\delta,\delta)} |f'|/\sigma)^m$ and $\epsilon_0(\delta) = \min\{\epsilon_1, \epsilon_2, \delta^2\}$. The constant $b > 0$ is defined below. Clearly, for all $\ell < m$ and $|t_1|, \ldots, |t_\ell| < \epsilon_0(\delta)$,

$$|(f^\ell_{t_\ell \cdots t_1})'(y)| \geq c(\delta)\sigma^\ell \quad \text{if} \quad y, f_{t_1}(y), \ldots, f^{\ell-1}_{t_{\ell-1} \cdots t_1}(y) \notin (-\delta, \delta). \tag{3.3}$$

Given n, $t_1, \ldots, t_n$, and x as in the statement, we denote $x_j = f^j_{t_j \cdots t_1}(x)$, $0 \leq j \leq n$. If $x_j \notin (-\delta_1, \delta_1)$ for all $0 \leq j < n$ then both (1) and (2) in the lemma follow immediately: Just write $n = qm + \ell$, with $0 \leq \ell < m$, and use (3.1), (3.2), (3.3). ¿From now on we suppose otherwise and define $0 \leq \nu_1 < \cdots < \nu_s < n$ as follows. Let ν_1 be the smallest $j \geq 0$ with $x_j \in (-\delta_1, \delta_1)$. For each ν_i, $i \geq 1$, define

$$p_i = \max\{k \geq 1 \colon |x_{\nu_i + j} - c_j| < e^{-\beta j} \text{ for every } 1 \leq j \leq k\},$$

where $\beta = 2\alpha$. Then let ν_{i+1} be the smallest $n > j > \nu_i + p_i$ for which $x_j \in (-\delta_1, \delta_1)$. For the time being we fix $1 \leq i \leq s$, and simply write $p = p_i$ and $\nu = \nu_i$. The previous definition and (A1) yield $|x_{\nu+j} - c_j| \leq e^{-\alpha j}|c_j|$ (we reduce δ_1 so that this holds also for $j < H_0$, c.f. (2.5)) and so $(1 - e^{-j\alpha})|f'(c_j)| \leq |f'(x_{\nu+j})| \leq (1 + e^{-j\alpha})|f'(c_j)|$. Then

$$\frac{1}{C}|(f^p)'(c_1)| \leq |(f^p_{t_{\nu+p} \cdots t_{\nu+1}})'(x_{\nu+1})| \leq C|(f^p)'(c_1)|\,. \tag{3.4}$$

In this proof $C > 0$ is some large constant depending only on H_0, α and λ_c. Moreover,

$$e^{-\beta(p+1)} \leq |x_{\nu+p+1} - c_{p+1}| \leq |x_{\nu+p} - c_p|(1 + e^{-p\alpha})|f'(c_p)| + \epsilon_0\,,$$

and so, by recurrence,

$$\begin{aligned} e^{-\beta(p+1)} &\leq \prod_{j=1}^p (1 + e^{-j\alpha})|(f^p)'(c_1)| \left[|x_{\nu+1} - c_1| + \epsilon_0 \sum_{j=1}^p \frac{|(f^j)'(c_1)|^{-1}}{\prod_{i=1}^j (1 + e^{-i\alpha})} \right] \\ &\leq C|(f^p)'(c_1)|\big[|x_\nu|^2 + \epsilon_0\big] \leq C|(f^p)'(c_1)||x_\nu|^2\,, \end{aligned} \tag{3.5}$$

where we also use $|x_\nu|^2 \geq \delta^2 \geq \epsilon_0$. Combining this with (3.4) and (A2), we conclude

$$|(f^{p+1}_{t_{\nu+p+1}\cdots t_{\nu+1}})'(x_\nu)|^2 \geq \frac{1}{C}|(f^p)'(c_1)|^2|x_\nu|^2 \geq \frac{1}{C}(\lambda_c e^{-\beta})^{p+1}. \tag{3.6}$$

Up to taking δ_1 small enough with respect to α and λ_c, we may suppose the p_i (uniformly) sufficiently large so that (3.6) implies

$$|(f^{p_i+1}_{t_{\nu_i+p_i+1}\cdots t_{\nu_i+1}})'(x_{\nu_i})| \geq \frac{1}{C}(\lambda_c e^{-\beta})^{(p_i+1)/2} \geq \frac{1}{C}(\lambda\rho)^{p_{i+1}} \geq K_2\lambda^{p_i+1}, \tag{3.7}$$

for each $1 \leq i \leq s$. At this point we write $|(f^n_{t_n\cdots t_1})'(x)| = \prod_{j=0}^{n-1}|f'(x_j)|$ and partition the range $[0,n)$ of this product into subintervals $J \subset [0,n)$ as follows. Let $|J|$ denote the number of elements of J. First, we suppose $\nu_s + p_s < n$. For $J = [0, \nu_1)$ and for each $J = (\nu_i + p_i, \nu_{i+1})$, $1 \leq i < s$, we have $\prod_{j\in J}|f'(x_j)| \geq K_2^{-1}\sigma^{|J|}$, as a consequence of (3.1) and (3.2). The same holds for $J = (\nu_s + p_s, n)$ if $|x_n| < \delta$. In general, $J = (\nu_s + p_s, n)$ has $\prod_{j\in J}|f'(x_j)| \geq c(\delta)\sigma^{|J|}$, by (3.1) and (3.3). Moreover, $\prod_{j\in J}|f'(x_j)| \geq K_2\sigma^{|J|}$ for each $J = [\nu_i, \nu_i + p_i]$, $1 \leq i \leq s$, recall (3.7). Altogether, this proves both parts of the lemma when $\nu_s + p_s < n$ (we shall take $b \leq (1/K_2)$). Now we treat the case $\nu_s + p_s \geq n$. We only have to consider $J = [\nu_s, n)$, as the previous estimates remain valid for all other subintervals involved. In general, (3.1) and (3.3) give $\prod_{j\in J}|f'(x_j)| \geq c(\delta)\sigma^{|J|}$. Part (1) follows, in the same way as before. In order to prove (2), we let $q = n - \nu_s - 1$. Then $0 \leq q < p_s$ and so, recall also (A1),

$$|x_{\nu_s}| \geq \delta > |x_n| \geq |c_{q+1}| - |x_n - c_{q+1}| \geq (1/C)e^{-\alpha(q+1)} \tag{3.8}$$

(when $q < H_0$ just reduce δ_1 to ensure $|x_n - c_{q+1}| < |c_{q+1}|/2$, then take C large with respect to $|c_{q+1}|$; similarly in the next equation). Moreover, (3.4) holds for $\nu = \nu_s$ and $p = q$. Hence,

$$|(f^{q+1}_{t_n\cdots t_{\nu_s+1}})'(x_{\nu_s})| \geq \frac{1}{C}|(f^q)'(c_1)||x_{\nu_s}| \geq \frac{1}{C}(\lambda_c e^{-\alpha})^{q+1} \geq \frac{1}{C}\lambda^{q+1}.$$

We take $b = (CK_2)^{-1}$, for $C > 0$ as in the last term. □

Remark. While the previous general argument gives $\sigma \leq \lambda$, better estimates are possible in some special cases. For instance, it is well-known that for quadratic maps with parameter $a \approx 2$ one may take σ close to 2. Note that σ^{-1} will be our uper bound for the essential spectral radius of $\mathcal{L}_0$ (Corollary 2), and that the constant $\bar{\tau}$ in our main theorem can be taken to be any number larger than $\sigma^{-1/2}$.

Lemma 2. *Let σ, δ_0, $c(\delta)$ and $\epsilon_0(\delta)$ be the objects from Lemma 1. Up to reducing δ_0 and $c(\delta)$ if necessary, the following holds for $\hat{f}_t$ as long as $0 < \delta < \delta_0$: Given any $n \geq 1$, there is $\epsilon(n) > 0$ such that, for all $\vec{t} = (t_1, \ldots, t_n)$ with $|t_1|, \ldots, |t_n| < \min(\epsilon(n), \epsilon_0(\delta))$*

and any $(x,0) \in E_0$ with $\hat f^n_{\vec t}(x,0) \in E_0$, we have $|(f^n_{\vec t})'(x)| \ge c(\delta)\sigma^{n-\ell}\lambda^\ell$ where ℓ is the maximum integer such that $\hat f^j_{t_j\dots t_1}(x,0) \in E_j$ for all $0 \le j \le \ell$.

Proof of Lemma 2. Here $C = C(H_0, \alpha, \beta_1, \beta_2, \lambda_c) > 0$. We take $|t_1|, \dots, |t_n|$ bounded by some $\epsilon > 0$. The case where all concerned iterates of $(x,0)$ are in E_0 is treated in Lemma 1 (1). Otherwise, the orbit $(x,0), \dots, \hat f^n_{\vec t}(x,0)$ consists of $q \ge 1$ loops of the form: $m \ge 0$ iterations in level E_0, climbing the tower up to some level $k \ge H(\delta)$ then falling down to level 0; finally, there may be an additional $s \ge 0$ iterations in level 0. By Lemma 1, it suffices to consider the case $q = 1$, $x \in (-\delta, \delta)$ (that is $m = 0$), and $s = 0$ and to prove that $|(f^n_{\vec t})'(x)| \ge \lambda^n/b$. As in the proof of Lemma 1, assumptions (A1) and (2.5) and the definition of E_j yield, for all $(y,j) \in E_j$ and $j \ge 1$,

$$(1 - e^{j(\alpha-\beta_1)})|f'(c_j)| \le |f'(y)| \le (1 + e^{j(\alpha-\beta_1)})|f'(c_j)| \,. \tag{3.9}$$

Then, using $\hat f^j_{t_j\dots t_1}(x,0) \in E_j$ for $1 \le j \le n-1$ and $\hat f^n_{\vec t}(x,0) \in E_0$ (i.e., $f^n_{\vec t}(x) \notin B_n$), we obtain, in just the same way as in the deduction of (3.4), (3.5),

$$e^{-\beta_2 n} \le C|(f^{n-1})'(c_1)|[|c_1 - f_{t_1}(x)| + C\epsilon] \le C|(f^{n-1})'(c_1)|[|x|^2 + \epsilon] \,. \tag{3.10}$$

We take $\epsilon < \epsilon(n) = e^{-\beta_2 n}/(2C|(f^{n-1})'(c_1)|)$, where $C > 0$ is as in the last term. Then (3.10) implies $|f'(x)|^2 \ge \frac{1}{C}|x|^2 \ge \frac{1}{C}e^{-\beta_2 n}|(f^{n-1})'(c_1)|^{-1}$. Hence, using (3.9),

$$|(f^n_{\vec t})'(x)| \ge \frac{1}{C}e^{-\beta_2 n/2}|(f^{n-1})'(c_1)|^{-1/2}|(f^{n-1})'(c_1)| \ge \frac{1}{C}\rho^{n-1}\lambda^{n-1} \,. \tag{3.11}$$

By the definition of $\hat f_t$ we must have $n - 1 \ge H(\delta)$, and we assume that δ_0 is small enough to ensure $\frac{1}{C}\rho^{H(\delta)} \ge \lambda/b$ for all $0 < \delta < \delta_0$. $\square$

Falling down from the tower.

Our next bounds concern the weight $g_t(y,k)$ evaluated at points in the support of μ_ϵ which “fall down” from the tower, i.e., such that $k \ge H(\delta)$ and $\hat f_t(y,k) \in E_0$.

Lemma 3. *There is $C > 0$ so that $w_\epsilon(y,k)|f'(y)|^{-1} \le C\rho^{-k}$ for all $\epsilon \ge 0$, all $k \ge 1$, and all $(y,k) \in E_k$ having $\hat f_t(y,k) \in E_0$ for some $|t| \le \epsilon$,*

Proof of Lemma 3. Suppose first that $\epsilon = 0$. By definition, if $w_0(y,k) \ne 0$ then

$$w_0(y,k)|f'(y)|^{-1} = \frac{\lambda^k}{|(f^{k+1})'(x)|}, \text{ where } x = \hat f_+^{-k}(y,k) \in (0,\delta) \,.$$

Since $\hat f^{k+1}(x,0) \in E_0$, (3.11) applies and yields $w_0(y,k)|f'(y)|^{-1} \le (C\lambda^k/\lambda^k\rho^k)$.

Assuming now that $\epsilon > 0$, we derive a preliminary estimate for w_ϵ on E_1. We continue to denote $C = C(H_0, \alpha, \beta_1, \beta_2, \lambda_c) > 0$. If $z \ge a + \epsilon$ then $w_\epsilon(z,1) = 0$. Otherwise, we use

$$w_\epsilon(z,1) = \lambda \int \frac{\theta_\epsilon(t)}{|f'(z_t)|}\, dt \le \lambda \sup \theta_\epsilon \cdot \int \frac{1}{|f'(z_t)|}\, dt \,,$$

where $z_t = \hat{f}_{t,+}^{-1}(z,1)$, the first integral is taken over $\{t \geq z-a, |z_t| \leq \delta\}$ and the second one over $\{t \geq z-a,\ |t| \leq \epsilon\}$. Hence, if $a-\epsilon < z < a+\epsilon$ then

$$w_\epsilon(z,1) \leq \lambda \sup\theta_\epsilon \cdot \int_{z-a}^{\epsilon} \frac{dt}{|f'(z_t)|} = \lambda \sup\theta_\epsilon \int_0^{z_\epsilon} dx = \lambda(\epsilon \sup\theta_\epsilon)\frac{z_\epsilon}{\epsilon} \leq \frac{C\lambda M}{\sqrt{\epsilon}}\,, \tag{3.12}$$

because $|z_\epsilon| < C\sqrt{\epsilon}$. On the other hand, for $z \leq a-\epsilon$ we have

$$w_\epsilon(z,1) \leq \lambda\ \sup\theta_\epsilon \cdot \int_{-\epsilon}^{\epsilon} \frac{dt}{|f'(z_t)|} = \lambda\ \sup\theta_\epsilon \frac{(z_\epsilon)^2-(z_{-\epsilon})^2}{z_\epsilon+z_{-\epsilon}} \leq \frac{C\lambda M}{z_0}\,, \tag{3.13}$$

since $(z_t)^2$ is a smooth function of t and $z_\epsilon + z_{-\epsilon} \geq z_0$. Now we consider a general $k \geq H(\delta)$. From the definition in Section 2,

$$w_\epsilon(y,k)|f'(y)|^{-1} = \lambda^{k-1}\int w_\epsilon(y_{t_{k-1}\cdots t_1},1)|(f^k_{0,t_{k-1}\cdots t_1})'(y_{t_{k-1}\cdots t_1})|^{-1} d\vec{\theta}_\epsilon(\vec{t})\,. \tag{3.14}$$

We split this into a sum W_1+W_2, where the two terms correspond to restricting the domain of integration, respectively, to $\{|a-y_{t_{k-1}\cdots t_1}| \geq \epsilon\}$ and to $\{|a-y_{t_{k-1}\cdots t_1}| < \epsilon\}$. In order to bound W_1 and W_2, we note that

$$e^{-\beta_2(k+1)} \leq C|(f^k)'(c_1)|\big[|a-y_{t_{k-1}\cdots t_1}|+\epsilon\big]\,. \tag{3.15}$$

This is deduced in just the same way as (3.10), using $\hat{f}^{j-1}_{t_{j-1}\cdots t_1}(y_{t_{k-1}\cdots t_1},1) \in E_j$ for $1 \leq j \leq k$ and $\hat{f}^k_{t,t_{k-1}\cdots t_1}(y_{t_{k-1}\cdots t_1},1) \in E_0$ for some $|t| \leq \epsilon$.

Let first $|a-y_{t_{k-1}\cdots t_1}| \geq \epsilon$. Then (3.15) gives $|a-y_{t_{k-1}\cdots t_1}| \geq (e^{-\beta_2 k}/C|(f^k)'(c_1)|)$ and, since $|z_0| \geq \sqrt{|a-z|}/C$, (3.13) then yields

$$w_\epsilon(y_{t_{k-1}\cdots t_1},1) \leq C\lambda M\sqrt{\frac{|(f^k)'(c_1)|}{e^{-\beta_2 k}}} \leq Ce^{\beta_2 k/2}|(f^k)'(c_1)|^{1/2}\,.$$

Replacing in W_1 and using again the distorsion inequality (3.9),

$$W_1 \leq \lambda^{k-1}\int Ce^{\beta_2 k/2}\frac{|(f^k)'(c_1)|^{1/2}}{|(f^k)'(c_1)|}d\vec{\theta}_\epsilon(\vec{t}) \leq C(\lambda e^{\beta_2/2}\lambda_c^{-1/2})^k \leq C\rho^{-k}\,.$$

For W_2, we use (3.15) to get that $\epsilon \geq (e^{-\beta_2 k}/C|(f^k)'(c_1)|)$ if $|a-y_{t_{k-1}\cdots t_1}| \leq \epsilon$. Then we use (3.12) to conclude that $w_\epsilon(y_{t_{k-1}\cdots t_1},1) \leq Ce^{\beta_2 k/2}|(f^k)'(c_1)|^{1/2}$. The same calculation as before gives $W_2 \leq C\rho^{-k}$, ending the proof of Lemma 3. $\square$

For $k \geq 1$, we introduce the subintervals of E_k

$$\beta_k^+ = \{(y,k) \mid f(y) > b_{k+1}-\epsilon\} \quad \text{and} \quad \beta_k^- = \{(y,k) \mid f(y) < a_{k+1}+\epsilon\}\,.$$

Note that $(y,k) \in \beta_k^+ \cup \beta_k^-$ if and only if $\hat{f}_t(y,k) \in E_0$ for some $|t| \leq \epsilon$.

Lemma 4. *There is a constant $C > 0$ such that for all $\epsilon \geq 0$ and $k \geq 1$*

$$\operatorname*{var}_{\beta_k^\pm}(w_\epsilon(y,k)|f'(y)|^{-1}) \leq C(e^\alpha \rho^{-1})^k .$$

Proof of Lemma 4. Recall that, for each fixed $\epsilon \geq 0$ and $k \geq 1$, $\{w_\epsilon(y,k) \neq 0\}$ is an interval. Denote by $\gamma_k^\pm$ its intersection with $\beta_k^\pm$. We suppose $k \geq H(\delta)$ for otherwise $\gamma_k^\pm$ is empty. First suppose that $\epsilon = 0$. For $(y,k) \in \gamma_k^\pm$ we have $w_0(y,k)|f'(y)|^{-1} = (\lambda^k/|(f^{k+1})'(\hat{f}_+^{-k}(y,k))|)$. Note that f^{k+1} has negative Schwarzian derivative, because f has. Moreover, f^{k+1} does not have critical points on $\hat{f}_+^{-k}(\gamma_k^\pm)$, because this last set does not contain $c = 0$, neither does $\pi(E_j \cap \operatorname{supp} w_0)$ for $j \geq 1$, see Section 2. This implies that $|(f^{k+1})'(\hat{f}_+^{-k}(y,k))|$ has a unique maximum and so $w_0(y,k)|f'(y)|^{-1}$ has a unique minimum, restricted to $\gamma_k^\pm$. Hence

$$\operatorname*{var}_{\beta_k^\pm}(w_0(y,k)|f'(y)|^{-1}) \leq 2 \sup_{\gamma_k^\pm}(w_0(y,k)|f'(y)|^{-1}) ,$$

and the claim corresponding to $\epsilon = 0$ follows from Lemma 3.

Assume now that $\epsilon > 0$. The main step is to prove that w_ϵ is at most two-to-one on each E_k. For this we use the assumption that $\phi_\epsilon = \log(\theta_\epsilon|_{J_\epsilon})$ is concave. Observe that a function ψ is concave if and only if

$$\psi(x_1) + \psi(x_4) \leq \psi(x_2) + \psi(x_3) \text{ for every } x_1 < x_2 \leq x_3 < x_4 \text{ with } x_1 + x_4 = x_2 + x_3 .$$

Given $j \geq 0$ (if $j = 0$ replace B_j by $\tilde{B}_0^+$)

$$\begin{aligned} & w_\epsilon(x_1, j+1) w_\epsilon(x_4, j+1) - w_\epsilon(x_2, j+1) w_\epsilon(x_3, j+1) = \\ & = \lambda^2 \int_{B_j} \int_{B_j} w_\epsilon(y,j) w_\epsilon(z,j)(\theta_\epsilon(x_1 - fy)\theta_\epsilon(x_4 - fz) - \theta_\epsilon(x_2 - fy)\theta_\epsilon(x_3 - fz))\, dy\, dz \\ & \leq 0 , \qquad \text{for all } x_1 < x_2 \leq x_3 < x_4 \text{ with } x_1 + x_4 = x_2 + x_3 . \end{aligned}$$

For the last inequality observe that the integrand is always nonpositive since we have $(x_1 - fy) + (x_4 - fz) = (x_2 - fy) + (x_3 - fz)$ and $\log(\theta_\epsilon|_{J_\epsilon})$ is concave. This proves that $\log w_\epsilon$ is concave and so w_ϵ is at most two-to-one on E_{j+1}, see Section 2. As a consequence

$$\operatorname*{var}_{\beta_k^\pm} w_\epsilon(y,k) \leq 2 \sup_{\gamma_k^\pm} w_\epsilon(y,k) \leq C\rho^{-k} ,$$

by Lemma 3. Therefore, since $|c_k| \geq e^{-\alpha k}$ and $|f'|$ has a unique maximum on each B_k for $k \geq H(\delta)$ (use $Sf < 0$ once more) we get

$$\begin{aligned} \operatorname*{var}_{\beta_k^\pm}(w_\epsilon(y,k)|f'(y)|^{-1}) & \leq \operatorname*{var}_{\beta_k^\pm} w_\epsilon(y,k) \sup_{\beta_k^\pm} |f'(y)|^{-1} + \sup_{\beta_k^\pm} w_\epsilon(y,k) \operatorname*{var}_{\beta_k^\pm} |f'(y)|^{-1} \\ & \leq C\rho^{-k} \cdot Ce^{\alpha k} + C\rho^{-k} \cdot Ce^{\alpha k} . \quad \square \end{aligned}$$

Climbing the tower.

We now proceed with some preliminary bounds on $\mathcal{L}_\epsilon$ concerning points which are "climbing the tower" : $(y,k) \in E_k$ and $\hat{f}_t(y,k) \in E_{k+1}$. Given $x \neq c$ we let x_- be the unique point with $x_- \neq x$ and $f(x_-) = f(x)$ and write $K(x) = |f'(x_-)|/|f'(x)|$. Then we set $K = \sup_{x \neq c} K(x)$, and $\tilde{K} = \mathrm{var}_{x \neq c} K(x)$. Note that under our assumptions K and $\tilde{K}$ are finite because $K(x)$ is C^1 (apply Morse's lemma; this is the only place where we use $f \in C^4$, in particular, C^3 suffices for all our purposes if f is symmetric) .

Lemma 5. *Let $\varphi \in BV(\hat{I})$ and $\epsilon \geq 0$.*

(1) *For $k \geq 1$ and each $\beta \subset E_{k+1} \cap \mathit{supp}\, \mu_\epsilon$, we have $\sup_\beta |\mathcal{L}_\epsilon \varphi| \leq \frac{1}{\lambda} \sup_\gamma |\varphi|$, where $\gamma = \cup_{t \in J_\epsilon} (\hat{f}_t|_{E_k})^{-1}(\beta) \cap \mathit{supp}\, \mu_\epsilon$.*

(2) *For each $\beta \subset E_1 \cap \mathit{supp}\, \mu_\epsilon$, we have $\sup_\beta |\mathcal{L}_\epsilon \varphi| \leq \frac{K}{\lambda}(\sup_{\gamma^+} |\varphi| + \sup_{\gamma^-} |\varphi|)$, where $\gamma^\pm = \cup_{t \in J_\epsilon} (\hat{f}_t|_{\tilde{E}_0^\pm})^{-1}(\beta)$.*

Proof of Lemma 5. The case $\epsilon = 0$ being easy, we assume $\epsilon > 0$. We first consider (1). By definition, for $k \geq 1$ and $x \in B_{k+1}$ such that $w_\epsilon(x, k+1) \neq 0$,

$$\mathcal{L}_\epsilon \varphi(x, k+1) = \frac{\int_{B_k} w_\epsilon(z,k) \varphi(z,k) \theta_\epsilon(x - fz)\, dz}{\lambda \int_{B_k} w_\epsilon(y,k) \theta_\epsilon(x - fy)\, dy} ,$$

and the claim follows. To show (2), we note that if $w_\epsilon(x,1) \neq 0$ then

$$\mathcal{L}_\epsilon \varphi(x, 1) = \frac{\int_0^\delta \varphi(z) \theta_\epsilon(x - fz)\, dz}{\lambda \int_0^\delta \theta_\epsilon(x - fy)\, dy} + \frac{\int_{-\delta}^0 \varphi(z) \theta_\epsilon(x - fz)\, dz}{\lambda \int_0^\delta \theta_\epsilon(x - fy)\, dy} ,$$

and use the change of variable $z = w_-$, with Jacobian bounded above by K, in the numerator of the second term (recalling that we take $(-\delta)_- \leq \delta$). $\square$

Lemma 6. *Let $\varphi \in BV(\hat{I})$ and $\epsilon \geq 0$.*

(1) *For all $k \geq 1$ and each interval $\beta \subset E_{k+1}$, we have $\mathrm{var}_\beta \mathcal{L}_\epsilon \varphi \leq \frac{1}{\lambda} \mathrm{var}_\gamma \varphi$, where $\gamma = \cup_{t \in J_\epsilon} (\hat{f}_t|_{E_k})^{-1}(\beta) \cap \mathit{supp}\, \mu_\epsilon$.*

(2) *For each interval $\beta \subset E_1$, we have $\mathrm{var}_\beta \mathcal{L}_\epsilon \varphi \leq \frac{K}{\lambda} \mathrm{var}_{\gamma^+ \cup \gamma^-} \varphi + \frac{\tilde{K}}{\lambda} \sup_{\gamma^-} |\varphi|$, where we write $\gamma^\pm = \cup_{t \in J_\epsilon} (\hat{f}_t|_{\tilde{E}_0^\pm})^{-1}(\beta)$.*

Proof of Lemma 6. Again, the easier case $\epsilon = 0$ is left to the reader. We start with $k \geq 1$. Consider first $\varphi|_{E_k} = H_u = \chi|_{[u,b_k] \times \{k\}}$ for some point $u \in B_k$. We shall prove that $\mathcal{L}_\epsilon \varphi$ is monotone on E_{k+1}. Obviously, we may disregard the points $(x, k+1)$ where $\mathcal{L}_\epsilon \varphi$ is defined by a limit (recall Section 2). At all other points,

$$\mathcal{L}_\epsilon \varphi(x, k+1) = \frac{\int_u^{b_k} w_\epsilon(z,k) \theta_\epsilon(x - fz)\, dz}{\lambda \int_{a_k}^{b_k} w_\epsilon(y,k) \theta_\epsilon(x - fy)\, dy} .$$

Fix $x_1 > x_2$ in $\pi(\beta)$ with $w_\epsilon(x_i, k+1) \neq 0$, $i = 1, 2$. Then, up to a positive factor, the difference $\mathcal{L}_\epsilon\varphi(x_1, k+1) - \mathcal{L}_\epsilon\varphi(x_2, k+1)$ is equal to

$$\int_{a_k}^{u} dy \int_{u}^{b_k} dz\, w_\epsilon(y,k) w_\epsilon(z,k) \Big(\theta_\epsilon(x_1 - fz)\theta_\epsilon(x_2 - fy) - \theta_\epsilon(x_2 - fz)\theta_\epsilon(x_1 - fy) \Big) . \quad (3.16)$$

Assume that $f|_{B_k \cap \operatorname{supp} \mu_\epsilon}$ is increasing (the other case is similar). Then $f(y) \le f(z)$ in (3.16). Thus $x_1 - fy \ge \max(x_1 - fz, x_2 - fy)$, and $x_2 - fz \le \min(x_1 - fz, x_2 - fy)$, so that, using $(x_1 - fy) + (x_2 - fz) = (x_1 - fz) + (x_2 - fy)$ together with the concavity of $\log(\theta_\epsilon|_{J_\epsilon})$, we get $\theta_\epsilon(x_1 - fz)\theta_\epsilon(x_2 - fy) \ge \theta_\epsilon(x_2 - fz)\theta_\epsilon(x_1 - fy)$. Hence $\mathcal{L}_\epsilon\varphi(x_1) \ge \mathcal{L}_\epsilon\varphi(x_2)$, i.e., $\mathcal{L}_\epsilon\varphi$ is nondecreasing on β. This proves, using Lemma 5, that $\operatorname{var}_\beta \mathcal{L}_\epsilon\varphi = \sup_\beta \mathcal{L}_\epsilon\varphi - \inf_\beta \mathcal{L}_\epsilon\varphi \le \frac{1}{\lambda}(1 - 0)$.

Consider now the case where

$$\varphi|_{E_k} = \sum_{j=1}^{m} d_j H_{u_j} , \quad (3.17)$$

for some $u_j \in B_k$ and $d_j > 0$. Then $\operatorname{var}_\beta \mathcal{L}_\epsilon\varphi = \operatorname{var}_\beta \mathcal{L}_\epsilon(d_0\chi_\gamma + \sum_{u_j\in\gamma} d_j H_{u_j})$ for some constant $d_0 \ge 0$. Observe that $\mathcal{L}_\epsilon(d_0\chi_\gamma)$ is constant on β. Therefore, by linearity, $\operatorname{var}_\beta \mathcal{L}_\epsilon\varphi \le \frac{1}{\lambda}\sum_{u_j\in\gamma} d_j = \frac{1}{\lambda}\operatorname{var}_\gamma \varphi$.

If $\varphi|_{E_k}$ is nonnegative and nondecreasing, we take a sequence of φ_n of the form (3.17) with $\varphi_n|_{E_k} \le \varphi|_{E_k}$ and converging uniformly to $\varphi|_{E_k}$. Since $\mathcal{L}_\epsilon\varphi_n$ converges pointwise to $\mathcal{L}_\epsilon\varphi$ on E_{k+1}, we get

$$\operatorname*{var}_\beta \mathcal{L}_\epsilon\varphi \le \liminf_n \operatorname*{var}_\beta \mathcal{L}_\epsilon\varphi_n \le \frac{1}{\lambda}\limsup_n \operatorname*{var}_\gamma \varphi_n \le \frac{1}{\lambda}\operatorname*{var}_\gamma \varphi .$$

Finally, if $\varphi|_{E_k}$ is any function with bounded variation, we may write $\varphi|_{E_k} = (\varphi_1 - \varphi_2) + i(\varphi_3 - \varphi_4)$ with the φ_j nonnegative, nondecreasing, and such that $\operatorname{var}_\gamma \varphi = \sum_{j=1}^4 \operatorname{var}_\gamma \varphi_j$. Then

$$\operatorname*{var}_\beta \mathcal{L}_\epsilon\varphi \le \sum_j \operatorname*{var}_\beta \mathcal{L}_\epsilon\varphi_j \le \sum_j \frac{1}{\lambda}\operatorname*{var}_\gamma \varphi_j = \frac{1}{\lambda}\operatorname*{var}_\gamma \varphi .$$

Consider now $\beta \subset E_1$. For a function φ which vanishes on $[-\delta, 0)$, the argument above may be reproduced and yields $\operatorname{var}_\beta \mathcal{L}_\epsilon\varphi \le \operatorname{var}_{\gamma^+} \varphi/\lambda$. It therefore suffices to consider functions with $\varphi(x, 0) = \text{const}$ for $x \ge 0$. Observe that $\mathcal{L}_\epsilon\varphi(x, 1)$ may be rewritten as

$$\mathcal{L}_\epsilon(\varphi\chi_{E_0^+})(x,1) + \frac{\int_0^{(-\delta)_-} \varphi(w_-, 0)\theta_\epsilon(x - fw)K(w_-)dw}{\lambda\int_0^\delta \theta_\epsilon(x - fy)dy} = \mathcal{L}_\epsilon(\varphi\chi_{E_0^+})(x,1) + \mathcal{L}_\epsilon(\psi)(x,1) ,$$

where we used the change of variable $w = z_-$, with Jacobian $K(w_-) = |f'(w)/f'(w_-)|$ and $\psi(z, 0) = \varphi(z_-, 0)K(z_-)\chi_{[-\delta,0)}(z_-)$. This yields $\operatorname{var}_\beta \mathcal{L}_\epsilon(\varphi) = \operatorname{var}_\beta \mathcal{L}_\epsilon(\psi)$ and, since ψ vanishes on $[-\delta, 0]$, we are reduced to the previous case. An application of properties (b) and (c) from Section 2 ends the proof of the lemma. $\square$

The measures μ_0 and μ_ϵ.

Lemma 7. $\lim_{\epsilon\to 0}\sum_{k\ge 0}\int_{B_k}|w_\epsilon(x,k)-w_0(x,k)|\,dx=0$.

Proof of Lemma 7. The term for $k=0$ vanishes. For $k=1$ we have

$$w_0(x,1)=\begin{cases}0 & x\notin f(-\delta,\delta)\\ \frac{\lambda}{|f'(x_0)|} & \text{otherwise,}\end{cases}\qquad w_\epsilon(x,1)=\begin{cases}0 & x\notin \cup_{t\in J_\epsilon} f_t(-\delta,\delta)\\ \lambda\int_{J_\epsilon}\frac{\theta_\epsilon(t)\,dt}{|f'(x_t)|} & \text{otherwise,}\end{cases}$$

with $x_t=\hat f_{t,+}^{-1}(x,1)$ for $t\ge 0$. Therefore, for small fixed $\zeta>0$ we have, by a computation similar to (3.12), $\int_{|a-x|\le\zeta} w_0(x,1)\,dx=\lambda(x_0(a-\zeta,1)-x_0(a,1))\le\lambda C\sqrt{\zeta}$. Since $w_\epsilon(x)$ converges uniformly to $w_0(x)$ on $|a-x|\ge\zeta$, the integral $\int_{|a-x|\ge\zeta}|w_\epsilon-w_0|\,dx$ can be made arbitrarily small by taking ϵ small. We split $\int_{|a-x|\le\zeta} w_\epsilon(x,1)\,dx$ into a sum W_1+W_2 where W_1, W_2 correspond to restricting the domain of integration to $\zeta\ge|a-x|\ge 2\epsilon$, respectively $|a-x|\le\min(2\epsilon,\zeta)$. The first term vanishes if $\zeta<2\epsilon$, otherwise it satisfies

$$W_1\le\int_{|a-x|\le\zeta}\lambda\frac{C}{|f'(x_0)|}\,dx\le C\sqrt{\zeta}\,,$$

since $|a-x+t|\ge|a-x|/2$. For the second summand, we have (recall (3.12))

$$W_2\le\int_{|a-x|<\min(2\epsilon,\zeta)}\frac{C\lambda M}{\sqrt{\epsilon}}\,dx\le\frac{C}{\sqrt{\epsilon}}\min(2\epsilon,\zeta)\le C\min(\sqrt{\epsilon},\sqrt{\zeta})\,.$$

We have thus proved:

$$\lim_{\epsilon\to 0}\int_{E_1}|w_0-w_\epsilon|\,dx=0\,,\tag{3.18}$$

$$\int_{|a-x|\le\zeta} w_\epsilon(x,1)\,dx\le C\sqrt{\zeta}\,.\tag{3.19}$$

For the levels $k\ge 2$, we get by the definition of $w_\epsilon(x,k)$ and a change of variable

$$\int_{B_k} w_\epsilon(x,k)\,dx=\lambda^{k-1}\int_{\vec t\in J_\epsilon^{k-1}}\int_{\gamma_k(\vec t)} w_\epsilon(y,1)dy\,d\vec\theta_\epsilon(\vec t)\tag{3.20}$$

where $\gamma_k(\vec t)\subset B_1$ is defined by $\gamma_k(\vec t)=\pi\big((\hat f_{\vec t}^{k-1})^{-1}(E_k)\cap\operatorname{supp}\mu_\epsilon\big)$. To control the top levels, we shall use the fact that for all $k\ge 1$, $\vec t\in J_\epsilon^{k-1}$ and $y\in\gamma_k(\vec t)$

$$|y-a|\le Ce^{-\beta_1 k}\lambda_c^{-(k-1)}\le C(e^{\beta_1}\lambda_c)^{-k}\,.\tag{3.21}$$

(To prove (3.21), use $|B_k|\le 2e^{-\beta_1 k}$ and (3.9) to obtain a constant $C>0$ such that for all $k\ge 1$, all $\vec t\in J_\epsilon^{k-1}$, and all $y\in\gamma_k(\vec t)$ we have $|(f_{\vec t}^{k-1})'(y)|\ge(1/C)\lambda_c^{k-1}$.) It then follows from (3.19)-(3.21) that for any $N_0\ge 2$

$$\sum_{k\ge N_0}\int_{B_k} w_\epsilon(x,k)\,dx\le\sum_{k\ge N_0}C\lambda^{k-1}(e^{\beta_1}\lambda_c)^{-k/2}\le\sum_{k\ge N_0}C(e^{\alpha+\beta_1/2}\rho)^{-k}\,.\tag{3.22}$$

Therefore, by taking N_0 large enough we may assume that $\sum_{k\ge N_0}\int_{B_k} w_\epsilon(x,k)\,dx$ is arbitrarily small, uniformly in ϵ. The same argument also gives

$$\sum_{k\ge N_0}\int_{B_k} w_0(x,k)\,dx = \sum_{k\ge N_0}\lambda^{k-1}\int_{\gamma_k(\vec{0})} w_0(y,1)\,dx \le \sum_{k\ge N_0} C(e^{\alpha+\beta_1/2}\rho)^{-k} \tag{3.23}$$

which is small if N_0 is large. It remains to bound $\sum_{2\le k\le N_0}\int_{B_k}|w_\epsilon(x,k)-w_0(x,k)|\,dx$. Using (3.20), and the equality in (3.23), we find

$$\sum_{2\le k\le N_0}\int_{B_k}|w_\epsilon(x,k)-w_0(x,k)|\,dx \le N_0\lambda^{N_0}\max\Big[\int_{\gamma_k(\vec{t})\cap\gamma_k(\vec{0})}|w_\epsilon(y,1)-w_0(y,1)|\,dy$$
$$+\int_{\gamma_k(\vec{t})\setminus\gamma_k(\vec{0})}|w_\epsilon(y,1)|\,dy+\int_{\gamma_k(\vec{0})\setminus\gamma_k(\vec{t})}|w_0(y,1)|\,dy\Big]$$

(the maximum is taken over all $\vec{t}$ and $2\le k\le N_0$) and the three terms of the right-hand-side tend to zero with ϵ, by (3.18), (3.19) and the properties of the intervals of monotonicity from Section 2 (note that $\gamma_k(\vec{t})$ is contained in some element of $\mathcal{Z}^{k,k}_{\vec{t}}$). □

Lemma 8. *For all $\epsilon\ge 0$ and $\varphi\in BV(\hat{I})$ we have $\int_{\hat{I}}\mathcal{L}_\epsilon\varphi\,d\mu_\epsilon=\int_{\hat{I}}\varphi\,d\mu_\epsilon$.*

Proof of Lemma 8. The presence of the w_ϵ factor in the integrand means that we do not need to consider the points (x,k) for which $\mathcal{L}_\epsilon\varphi(x,k)$ is defined by a limit. If $\epsilon=0$, use the change of variable formula in the integral. For $\epsilon>0$, by Fubini's theorem

$$\sum_{k\ge 0}\int_{B_k}\mathcal{L}_\epsilon\varphi(x,k)w_\epsilon(x,k)\,dx=$$
$$=\sum_{j\ge 0}\int_{B_j}\varphi(y,j)w_\epsilon(y,j)\left(\sum_{k\ge 0}\int_{B_k}\hat{\theta}_\epsilon((x,k),\hat{f}(y,j))\,dx\right)dy$$

and (using $f_t(I)\subset I$, all t) $\sum_{k\ge 0}\int_{B_k}\hat{\theta}_\epsilon((x,k),\hat{f}(y,j))\,dx=1$ for any $(y,j)\in\hat{I}$ (this just corresponds to the fact the Markov transitions $\hat{P}^\epsilon$ are probability measures). □

4. Main estimates

Bounding the variation.

We prove our main estimate:

Variation Lemma. *For each $\sigma_0<\sigma$ there is a constant $C>0$ and for each $n\ge 1$ there is an $\epsilon(n)>0$ such that for all $0\le\epsilon<\epsilon(n)$ and all $\varphi\in BV(\hat{I})$*

$$\operatorname*{var}_{\hat{I}}\mathcal{L}_\epsilon^n\varphi\le C\sigma_0^{-n}(\operatorname*{var}_{\hat{I}}\varphi+\sup_{\hat{I}}|\varphi|)+C\int_{\hat{I}}|\varphi(x)|w_0(x)\,dx\,.$$

We shall need:

Sublemma. *There is $C > 0$ and given $n \geq 1$ there are $\epsilon(n) > 0$ and $C(n) > 0$ such that for every $0 \leq \epsilon < \epsilon(n)$, every $\varphi \in BV(\hat{I})$, and every interval $A \subset E_0$,*

$$\operatorname*{var}_A \mathcal{L}_\epsilon^n \varphi \leq C\sigma^{-n}(\operatorname*{var}_{\hat{I}} \varphi + \sup_{\hat{I}} |\varphi|) + C(n) \int_{\hat{I}} |\varphi| d\mu_\epsilon \,.$$

Proof of the Sublemma. We only consider $\epsilon > 0$: the case $\epsilon = 0$ is obtained by (simpler forms of) the same arguments, using the versions for $\epsilon = 0$ of the lemmas in Section 3.

Fix $A \subset E_0$ and $n \geq 1$. Our starting point is the following decomposition of backwards orbits. Let $\epsilon > 0$ and $\vec{t} = (t_1, \ldots, t_n) \in J_\epsilon^n$ be fixed. For each $0 \leq j < n$ we define $\mathcal{G}(j) = \mathcal{G}(j, \epsilon, t_n, \ldots, t_{n-j})$ to be the set of all nonempty intervals of the form $\gamma = \eta \cap \left(\hat{f}^{j+1}_{t_n \cdots t_{n-j}}\right)^{-1}(A) \cap \operatorname{supp} \mu_\epsilon$, where η is an interval of monotonicity of $\hat{f}^{j+1}_{t_n \cdots t_{n-j}}$ with $\eta \subset (\cup_{k \geq 1} E_k)$ and $\hat{f}^{\ell+1}_{t_{n-j+\ell} \cdots t_{n-j}}(\eta) \subset E_0$ for $0 \leq \ell \leq j$. Moreover, we let $\underline{\mathcal{G}} = \underline{\mathcal{G}}(\vec{t})$ be the set of all nonempty intervals of the form $\underline{\gamma} = \eta \cap \left(\hat{f}^n_{t_n \cdots t_1}\right)^{-1}(A)$, where η is an interval of monotonicity of $\hat{f}^n_{t_n \cdots t_1}$ with $\hat{f}^\ell_{t_\ell \cdots t_1}(\eta) \subset E_0$ for $0 \leq \ell \leq n$.

Using this decomposition, the definition of $\mathcal{L}^n_{\vec{t}}$, and (a), (c) from Section 2, we obtain the basic inequality:

$$\operatorname*{var}_{\hat{I}}(\chi_A \mathcal{L}^n_{\vec{t}} \varphi) \leq \sum_{0 \leq j < n} \sum_{\gamma \in \mathcal{G}(j)} \operatorname*{var}_{\hat{I}}(\chi_\gamma g^{(j+1)} \mathcal{L}^{n-j-1}_{t_1 \ldots t_{n-j-1}} \varphi) + \sum_{\underline{\gamma} \in \underline{\mathcal{G}}} \operatorname{var}(\chi_{\underline{\gamma}} g^{(n)} \varphi) \,. \tag{4.1}$$

(we write $g^{(\ell)} = g^{(\ell)}_{t_n \cdots t_{n-\ell+1}}$ for simplicity). Note that $\operatorname{var}_A(\mathcal{L}^n_\epsilon \varphi) \leq \operatorname{var}_{\hat{I}}(\chi_A \mathcal{L}^n_\epsilon \varphi)$. Hence, by the definition of $\mathcal{L}^n_\epsilon$ and using inequality (e) from Section 2 (observe also that neither $g^{(j+1)}$ nor $\mathcal{G}(j)$ depend on $t_1, \ldots, t_{n-j-1}$),

$$\operatorname*{var}_A(\mathcal{L}^n_\epsilon \varphi) \leq \sum_{0 \leq j < n} \sup_{t_{n-j}, \ldots, t_n} \sum_{\gamma \in \mathcal{G}(j)} \operatorname*{var}_{\hat{I}}(\chi_\gamma g^{(j+1)} \mathcal{L}^{n-j-1}_\epsilon \varphi) + \sup_{t_1, \ldots, t_n} \sum_{\underline{\gamma} \in \underline{\mathcal{G}}} \operatorname*{var}_{\hat{I}}(\chi_{\underline{\gamma}} g^{(n)} \varphi) \,.$$

Writing the right-hand-side of the above inequality as $S_1 + S_2$, we first bound S_2 (recall (b) and (d) from Section 2):

$$S_2 \leq \sup_{\vec{t}} \sum_{\underline{\gamma} \in \underline{\mathcal{G}}} \big((\operatorname*{var}_{\underline{\gamma}} g^{(n)} + 2 \sup_{\underline{\gamma}} g^{(n)}) \cdot \sup_{\underline{\gamma}} |\varphi| + \operatorname*{var}_{\underline{\gamma}} \varphi \cdot \sup_{\underline{\gamma}} g^{(n)} \big).$$

By Lemma 1 and the definition of $\underline{\gamma}$ and $g^{(n)}$,

$$\sup_{\underline{\gamma}} g^{(n)} = \sup_{\underline{\gamma}} |(f^n_{\vec{t}})'|^{-1} \leq \begin{cases} C\sigma^{-n} & \text{if } A \subset (-\delta, \delta) \\ C(1/\delta)\sigma^{-n} & \text{in general.} \end{cases} \tag{4.2}$$

Also, since $f^n_{\vec{t}}$ has negative Schwarzian derivative and no critical points in $\underline{\gamma}$, the function $g^{(n)}$ has at most one local minimum on $\underline{\gamma}$. Therefore

$$\operatorname*{var}_{\underline{\gamma}} g^{(n)} \leq 2 \sup_{\underline{\gamma}} g^{(n)} \,. \tag{4.3}$$

For each $\underline{\gamma} \in \underline{\mathcal{G}}$ let η be the corresponding interval of monotonicity of $\hat{f}_{\vec{t}}^n$. This is the continuation (in the sense of the last subsection of Section 2; we assume $\epsilon < \epsilon(n)$) of some interval $\eta_0 \in \mathcal{Z}_0^{n,n}$. Then $\underline{\gamma} \subset \eta \subset \eta^+(\epsilon, \eta_0)$. We write $\underline{\gamma}^+ = \underline{\gamma}^+(\underline{\gamma}) = \eta^+(\epsilon, \eta_0)$ and denote $\underline{\mathcal{G}}^+(S_2)$ the set of all $\underline{\gamma}^+$ found in this way. By the properties of the monotonicity intervals described in Section 2, we have $1/\mu_\epsilon(\underline{\gamma}^+) \le C(n)$ for all $\underline{\gamma}^+ \in \underline{\mathcal{G}}$ (observe that $\mu_\epsilon|_{E_0} = \mu_0|_{E_0}$ is Lebesgue measure on E_0). Altogether, this allows us to write

$$\sup_{\underline{\gamma}} |\varphi| \le \sup_{\underline{\gamma}^+} |\varphi| \le \operatorname{var}_{\underline{\gamma}^+} \varphi + \frac{1}{\mu_\epsilon(\underline{\gamma}^+)} \int_{\underline{\gamma}^+} |\varphi| \, d\mu_\epsilon \le \operatorname{var}_{\underline{\gamma}^+} \varphi + C(n) \int_{\hat{I}} |\varphi| \, d\mu_\epsilon \,, \tag{4.4}$$

for all $\epsilon < \epsilon(n)$, where $\underline{\gamma}^+ = \underline{\gamma}^+(\underline{\gamma})$. Combining the bounds (4.2)–(4.4) with the remark that $\# \underline{\mathcal{G}}^+(S_2) \le \mathcal{Z}_0^{n,n} \le C(n)$, we obtain

$$\begin{aligned} S_2 &\le \sum_{\underline{\gamma}^+ \in \underline{\mathcal{G}}^+(S_2)} \big(C(\frac{1}{\delta}) \sigma^{-n} \operatorname{var}_{\underline{\gamma}^+} \varphi + C(\frac{1}{\delta}) C(n) \int_{\hat{I}} |\varphi| \, d\mu_\epsilon \big) \\ &\le C(\frac{1}{\delta}) \sigma^{-n} \big(\sum_{\underline{\gamma}^+ \in \underline{\mathcal{G}}^+(S_2)} \operatorname{var}_{\underline{\gamma}^+} \varphi \big) + C(\frac{1}{\delta}) C(n) \int_{\hat{I}} |\varphi| \, d\mu_\epsilon \end{aligned} \tag{4.5}$$

where $C(1/\delta)$ may be replaced by C if $A \subset (-\delta, \delta)$.

We now move on to bound S_1, and again start with the observation that

$$\begin{aligned} S_1 \le \sum_{0 \le j < n} \sup_{t_{n-j}, \dots, t_n} \sum_{\gamma \in \mathcal{G}(j)} \\ \big((\operatorname{var}_\gamma g^{(j+1)} + 2 \sup_\gamma g^{(j+1)}) \sup_\gamma |\mathcal{L}_\epsilon^{n-j-1} \varphi| + \sup_\gamma g^{(j+1)} \operatorname{var}_\gamma |\mathcal{L}_\epsilon^{n-j-1} \varphi| \big) \,. \end{aligned} \tag{4.6}$$

Note that for $\gamma \in \mathcal{G}(j)$ with $\gamma \subset E_k$, Lemmas 3 and 4 together with the analogues of (4.2)–(4.3) obtained replacing $\underline{\gamma}$ by $\hat{f}_{t_{n-j}}(\gamma)$ and n by j, yield for all $\vec{t}$

$$\begin{aligned} &\sup_\gamma g^{(j+1)} = \sup_\gamma (g^{(j)} \circ \hat{f}_{t_{n-j}} \cdot g) \le C(1/\delta) \sigma^{-j} \rho^{-k} \,, \\ &\operatorname{var}_\gamma g^{(j+1)} \le \sup_{\hat{f}_{t_{n-j}}(\gamma)} g^{(j)} \operatorname{var}_\gamma g + \operatorname{var}_{\hat{f}_{t_{n-j}}(\gamma)} g^{(j)} \sup_\gamma g \le C(1/\delta) \sigma^{-j} e^{\alpha k} \rho^{-k} \end{aligned} \tag{4.7}$$

(we used (b), (c) from Section 2), where $C(1/\delta)$ may be replaced by C if $A \subset (-\delta, \delta)$.

For $\gamma \in \mathcal{G}(j)$ we write $j(\gamma) = j$, and also $k(\gamma) = k$ if $\gamma \subset E_k$. Observe that we always have $k(\gamma) \ge H(\delta)$, because $\gamma \subset (\cup_{k \ge 1} E_k) \cap \operatorname{supp} \mu_\epsilon$ and $\hat{f}_{t_{n-j}}(\gamma) \subset E_0$. Fixing a large value of $N \ge n$, to be determined below as a function of n only, we split the sums in S_1 into

$$\sum_{0 \le j < n} \sup_{t_{n-j}, \dots, t_n} \Big(\sum_{\substack{\gamma \in \mathcal{G}(j): \\ k(\gamma) > N}} + \sum_{\substack{\gamma \in \mathcal{G}(j): k(\gamma) \le N \\ k(\gamma) \ge n-j-1}} + \sum_{\substack{\gamma \in \mathcal{G}(j): \\ k(\gamma) < n-j-1}} \Big) = s_1 + s_2 + s_3 \,.$$

We now proceed to bound s_1, s_2, and s_3, using each time a decomposition such as in (4.6) as a starting point.

First, using Lemmas 5 and 6 ($n-j-1$ times) together with (4.7), we get

$$s_1 \le \sum_{0\le j<n} \sup_{t_{n-j},\dots,t_n} \sum_{\gamma\in\mathcal{G}(j):k(\gamma)>N} C(\frac{1}{\delta})\rho^{-k}\sigma^{-j}\lambda^{-(n-j-1)}(\operatorname*{var}_{\underline{\gamma}}\varphi + e^{\alpha k}\sup_{\underline{\gamma}}|\varphi|)\,, \tag{4.8}$$

where $\underline{\gamma}=\underline{\gamma}(\gamma)=\left(\cup_{\vec{t}}(\hat{f}^{n-j-1}_{t_{n-j-1}\dots t_1})^{-1}(\gamma)\right)\cap\operatorname{supp}\mu_\epsilon\subset E_{k(\gamma)-(n-j-1)}$ and $C(1/\delta)$ may be replaced by C if $A\subset(-\delta,\delta)$. Since $\hat{f}^j_{t_n\dots t_{n-j+1}}|_{E_0}$ is at most 2^j-to-1, and E_k contains at most two intervals of monotonicity mapped to E_0 by $\hat{f}_{t_{n-j}}$, each sum on $\gamma\in\mathcal{G}(j)$ in s_1 (or s_2) ranges over at most 2^{j+1} elements for each given value of $k=k(\gamma)$. Therefore, (4.8) yields

$$\begin{aligned} s_1 &\le \sum_{0\le j<n}\sum_{k>N} 2^{j+1}C(\frac{1}{\delta})(e^\alpha/\rho)^k(\lambda/\sigma)^j\lambda^{-n}(\operatorname*{var}_{\hat{I}}\varphi+\sup_{\hat{I}}|\varphi|) \\ &\le C(\frac{1}{\delta})\left(\frac{e^\alpha}{\rho}\right)^N\left(\frac{2\lambda}{\sigma}\right)^n\lambda^{-n}(\operatorname*{var}_{\hat{I}}\varphi+\sup_{\hat{I}}|\varphi|)\le\frac{C(1/\delta)}{C_n(N)}\sigma^{-n}(\operatorname*{var}_{\hat{I}}\varphi+\sup_{\hat{I}}|\varphi|) \end{aligned} \tag{4.9}$$

(recall that $\rho>e^\alpha$ and $\sigma<2\lambda$). As usual, $C(1/\delta)$ may be replaced by C if $A\subset(-\delta,\delta)$.

Analogously

$$s_2 \le \sum_{0\le j<n} \sup_{t_{n-j},\dots,t_n} \sum_{\substack{\gamma\in\mathcal{G}(j):k(\gamma)\le N\\ k(\gamma)\ge n-j-1}} C(\frac{1}{\delta})\rho^{-k}e^{\alpha k}\sigma^{-j}\lambda^{-(n-j-1)}(\operatorname*{var}_{\underline{\gamma}}\varphi+\sup_{\underline{\gamma}}|\varphi|) \tag{4.10}$$

where $\underline{\gamma}=\underline{\gamma}(\gamma)\subset E_{k(\gamma)-(n-j-1)}$ is as defined after (4.8). Note that $\underline{\gamma}$ is an interval if $k(\gamma)>n-j-1$ and $\underline{\gamma}$ is a union of two subintervals of $\tilde{E}_0^+$, respectively $\tilde{E}_0^-$, if $k(\gamma)=n-j-1$. In what follows we consider the first case, the second one being entirely analogous (just treat the two subintervals separately). Proceeding as before in the case of S_2, we find $\eta_0\in\mathcal{Z}_0^{j+1,N}$ such that $\gamma\subset\eta^+(\epsilon,\eta_0)$ and also $\xi_0\in\mathcal{Z}_0^{n,N}$ such that $\underline{\gamma}\subset\eta^+(\epsilon,\xi_0)$. Observe that

$$\hat{f}^{n-j-1}(\xi_0)\subset\eta_0 \quad\text{and}\quad \hat{f}^{n-j-1}(\xi_0\cap\operatorname{supp}\mu_0)=\eta_0\cap\operatorname{supp}\mu_0. \tag{4.11}$$

Indeed, the first statement is a direct consequence of the properties of the partitions into intervals of monotonicity studied in Section 2. For the second one, recall also that $E_\ell\cap\operatorname{supp}\mu_0=E_\ell\cap\operatorname{Im}\hat{f}^\ell$ is an interval for each $\ell\ge1$ and that $\hat{f}$ is monotone on each of these intervals. Now we write $\underline{\gamma}^+=\underline{\gamma}^+(\gamma)=\eta^+(\epsilon,\xi_0)$ and denote $\underline{\mathcal{G}}^+(s_2)$ the set of $\underline{\gamma}^+$ obtained by varying j and $\gamma\in\mathcal{G}(j)$ in (4.10). Clearly, $\#\underline{\mathcal{G}}^+(s_2)\le\#\mathcal{Z}_0^{n,N}\le C_n(N)$. We also claim that $1/\mu_\epsilon(\underline{\gamma}^+)\le C_n(N)$ for all $\underline{\gamma}^+\in\underline{\mathcal{G}}^+(s_2)$, as long as $\epsilon<\epsilon(n,N)$. To prove this, combine the properties of the monotonicity intervals with the remark that $\mu_0(\xi_0)>0$, which is a consequence of $\mu_\epsilon(\gamma)>0$ and (4.11).

Hence, we are in a position to apply the same kind of calculations as in (4.4)–(4.5) and (4.9), to get

$$\begin{aligned} s_2 &\le \sum_{0\le j<n} \sum_{\substack{\gamma\in\mathcal{G}(j):k(\gamma)\le N\\ k(\gamma)\ge n-j-1}} C(\frac{1}{\delta})\left(\frac{e^\alpha}{\rho}\right)^k \lambda^{-n+j}\sigma^{-j}(\operatorname*{var}_{\underline{\gamma}^+(\gamma)}\varphi + C_n(N)\int_{\hat I}|\varphi|\,d\mu_\epsilon) \\ &\le C(\frac{1}{\delta})\sigma^{-n}(\sum_{\underline{\gamma}^+\in\underline{\mathcal{G}}^+(s_2)}\operatorname*{var}_{\underline{\gamma}^+}\varphi) + C(\frac{1}{\delta})C_n(N)\int_{\hat I}|\varphi|\,d\mu_\epsilon\,, \end{aligned} \tag{4.12}$$

(we use $\rho\ge e^\alpha$ and $\sigma\le\lambda$) where, again, $C(1/\delta)$ may be replaced by C if $A\subset(-\delta,\delta)$.

Using similar arguments and similar sets $\underline{\gamma}^+(\gamma)\in\underline{\mathcal{G}}^+(s_3)$, we get

$$s_3 \le (\sum_{\underline{\gamma}^+\in\underline{\mathcal{G}}^+(s_3)} C(\frac{1}{\delta})\left(\frac{e^\alpha}{\rho}\right)^k\sigma^{-\ell}\operatorname*{var}_{\underline{\gamma}^+}(\mathcal{L}_\epsilon^{n-\ell}\varphi)) + C(\frac{1}{\delta})C(n)\int_{\hat I}|\varphi|d\mu_\epsilon\,, \tag{4.13}$$

where, for simplicity, we write $\ell=j+1+k$. Note that we used Lemma 8 in the integral term. As before, $C(1/\delta)$ can be replaced by C if $A\subset(-\delta,\delta)$.

Putting together (4.5), (4.9), (4.12), and (4.13), we obtain for general $A\subset E_0$

$$\begin{aligned} \operatorname*{var}_A(\mathcal{L}_\epsilon^n\varphi) \le{}& C(\frac{1}{\delta})\sigma^{-n}\sum_{\underline{\gamma}^+\in\underline{\mathcal{G}}^+(S_2)\cup\underline{\mathcal{G}}^+(s_2)}\operatorname*{var}_{\underline{\gamma}^+}\varphi + \frac{C(1/\delta)}{C_n(N)}\sigma^{-n}(\operatorname*{var}_{\hat I}\varphi+\sup_{\hat I}|\varphi|) \\ &+ C(\frac{1}{\delta})C_n(N)\int_{\hat I}|\varphi|\,d\mu_\epsilon + \sum_{\underline{\gamma}^+\in\underline{\mathcal{G}}^+(s_3)} C(\frac{1}{\delta})\sigma^{-\ell}\operatorname*{var}_{\underline{\gamma}^+}(\mathcal{L}_\epsilon^{n-\ell}\varphi)\,, \end{aligned} \tag{4.14}$$

and for $A\subset(-\delta,\delta)$

$$\begin{aligned} \operatorname*{var}_A(\mathcal{L}_\epsilon^n\varphi) \le{}& C\sigma^{-n}\sum_{\underline{\gamma}^+\in\underline{\mathcal{G}}^+(S_2)\cup\underline{\mathcal{G}}^+(s_2)}\operatorname*{var}_{\underline{\gamma}^+}\varphi + \frac{1}{C_n(N)}\sigma^{-n}(\operatorname*{var}_{\hat I}\varphi+\sup_{\hat I}|\varphi|) \\ &+ C_n(N)\int_{\hat I}|\varphi|\,d\mu_\epsilon + \sum_{\underline{\gamma}^+\in\underline{\mathcal{G}}^+(s_3)} 1\cdot\sigma^{-\ell}\operatorname*{var}_{\underline{\gamma}^+}(\mathcal{L}_\epsilon^{n-\ell}\varphi)\,. \end{aligned} \tag{4.15}$$

Note that the previous calculations, for $A\subset(-\delta,\delta)$, yield a factor $C(e^\alpha/\rho)^k$ in the last term of (4.15), see (4.13). On the other hand, since $\rho>e^\alpha$ and $k\ge H(\delta)$, we have $C(e^\alpha/\rho)^k\le 1$ has long as δ has been fixed small enough. This allows us to replace that factor by 1, as we did, which is necessary for the sequel of our argument.

If the sum over $\underline{\gamma}^+\in\underline{\mathcal{G}}^+(s_3)$ in (4.14) is not void, we need to proceed by recurrence. We change the name of the $\underline{\mathcal{G}}^+$, the $\underline{\gamma}^+$, and their indices $k(\underline{\gamma}^+)$, $j(\underline{\gamma}^+)$ in (4.14)-(4.15), to $\underline{\mathcal{G}}_1^+$, $\underline{\gamma}_1^+$, k_1, and j_1. The corresponding objects appearing at the i^{th} step (for each fixed $\underline{\gamma}_{i-1}^+$) will be denoted $\underline{\mathcal{G}}_i^+$, $\underline{\gamma}_i^+$, k_i, j_i and we also let $\ell_i=j_i+1+k_i$. Since every such $\underline{\gamma}_i^+$ is, by construction, a subset of $(-\delta, \delta)$, we may apply (4.15) to it. After one induction step, we get

$$\begin{aligned}
\operatorname*{var}_A \mathcal{L}_\epsilon^n \varphi \le C(\frac{1}{\delta})\sigma^{-n} \sum_{\underline{\gamma}_1^+} \operatorname*{var}_{\underline{\gamma}_1^+} \varphi + \frac{C(1/\delta)}{C_n(N)} \sigma^{-n} (\operatorname*{var}_{\hat{I}} \varphi + \sup_{\hat{I}} |\varphi|) + C(\frac{1}{\delta}) C_n(N) \int_{\hat{I}} |\varphi| \, d\mu_\epsilon \\
+ \sum_{\underline{\gamma}_1^+ \in \underline{\mathcal{G}}_1^+(s_3)} C(\frac{1}{\delta})\sigma^{-\ell_1} \cdot \Big[C\sigma^{-n+\ell_1} \sum_{\underline{\gamma}_2^+} \operatorname*{var}_{\underline{\gamma}_2^+} \varphi + \frac{1}{C_n(N)} \sigma^{-n+\ell_1} (\operatorname*{var}_{\hat{I}} \varphi + \sup_{\hat{I}} |\varphi|) \\
+ C_n(N) \int_{\hat{I}} |\varphi| \, d\mu_\epsilon + \sum_{\underline{\gamma}_2^+ \in \underline{\mathcal{G}}_2^+(s_3)} \sigma^{-\ell_2} \operatorname*{var}_{\underline{\gamma}_2^+} (\mathcal{L}_\epsilon^{n-\ell_1-\ell_2} \varphi) \Big] ,
\end{aligned}$$

the first sum being over $\underline{\gamma}_1^+ \in \underline{\mathcal{G}}_1^+(S_2) \cup \underline{\mathcal{G}}_1^+(s_2)$, the third one over $\underline{\gamma}_2^+ \in \underline{\mathcal{G}}_2^+(S_2) \cup \underline{\mathcal{G}}_2^+(s_2)$. This gives,

$$\begin{aligned}
\operatorname*{var}_A \mathcal{L}_\epsilon^n \varphi \le C(\frac{1}{\delta})\sigma^{-n} \big(\sum_{\underline{\gamma}_1^+} \operatorname*{var}_{\underline{\gamma}_1^+} \varphi + C \sum_{\underline{\gamma}_1^+, \underline{\gamma}_2^+} \operatorname*{var}_{\underline{\gamma}_2^+} \varphi \big) + \sum_{\underline{\gamma}_1^+, \underline{\gamma}_2^+} C(\frac{1}{\delta}) \sigma^{-\ell_1-\ell_2} \operatorname*{var}_{\underline{\gamma}_2^+} (\mathcal{L}_\epsilon^{n-\ell_1-\ell_2} \varphi) \\
+ (1 + \#\underline{\mathcal{G}}_1^+(s_3)) C(\frac{1}{\delta}) \left[\frac{1}{C_n(N)} \sigma^{-n} (\operatorname*{var}_{\hat{I}} \varphi + \sup_{\hat{I}} |\varphi|) + C_n(N) \int_{\hat{I}} |\varphi| \, d\mu_\epsilon \right] ,
\end{aligned}$$

the sums running over $\underline{\gamma}_1^+ \in \underline{\mathcal{G}}_1^+(S_2) \cup \underline{\mathcal{G}}_1^+(s_2)$, over $\underline{\gamma}_1^+ \in \underline{\mathcal{G}}_1^+(s_3)$, $\underline{\gamma}_2^+ \in \underline{\mathcal{G}}_2^+(S_2) \cup \underline{\mathcal{G}}_2^+(s_2)$, and over $\underline{\gamma}_1^+ \in \underline{\mathcal{G}}_1^+(s_3)$, $\underline{\gamma}_2^+ \in \underline{\mathcal{G}}_2^+(s_3)$, respectively. Now, $\#\underline{\mathcal{G}}_1^+(s_3) \le \#\mathcal{Z}_0^{\ell_1,k_1} \le C(n)$ because $k_1 < \ell_1 \le n$. In fact, $\#\underline{\mathcal{G}}_i^+(s_3) \le C(n)$ for all $1 \le i \le n$, for a similar reason. Hence, after at most n steps we get

$$\begin{aligned}
\operatorname*{var}_A \mathcal{L}_\epsilon^n \varphi \le C(\frac{1}{\delta})\sigma^{-n} \sum_{i=1}^n \Big(\sum_{\underline{\gamma}_1^+, \dots, \underline{\gamma}_{i-1}^+, \underline{\gamma}_i^+} \operatorname*{var}_{\underline{\gamma}_i^+} \varphi \Big) \\
+ C(n) C(\frac{1}{\delta}) \Big(\frac{1}{C_n(N)} \sigma^{-n} (\operatorname*{var}_{\hat{I}} \varphi + \sup_{\hat{I}} |\varphi|) + C_n(N) \int_{\hat{I}} |\varphi| \, d\mu_\epsilon \Big) ,
\end{aligned} \tag{4.16}$$

the second sum being over $\underline{\gamma}_1^+ \in \underline{\mathcal{G}}_1^+(s_3), \dots, \underline{\gamma}_{i-1}^+ \in \underline{\mathcal{G}}_{i-1}^+(s_3)$, $\underline{\gamma}_i^+ \in \underline{\mathcal{G}}_i^+(S_2) \cup \underline{\mathcal{G}}_i^+(s_2)$. Now observe that the intervals $\underline{\gamma}_i^+$ occurring in (4.16) are all distinct elements of $\{\eta^+(\epsilon, \eta_0) \mid \eta_0 \in \mathcal{Z}_0^{n,N}\}$. Therefore, they have the bounded overlap property (overlap bounded by 2 in fact, recall Section 2) and so the first term on the right-hand-side of (4.16) is bounded by $2C(1/\delta)\sigma^{-n} \operatorname{var}_{\hat{I}} \varphi$. Also, fixing $N \gg n$ depending only on n, we may ensure that $C(n)/C_n(N) \le 1$, so that the second term is bounded by $C(1/\delta)\sigma^{-n}(\operatorname{var}_{\hat{I}} \varphi + \sup_{\hat{I}} |\varphi|)$. Finally, since δ is fixed at this point, we may omit the dependence of the constants in (4.16) on it (i.e., we just write C instead of $C(1/\delta)$), thus ending the proof of the Sublemma. $\square$

Proof of the Variation Lemma. We start by fixing some $n = n_0$ and decomposing $\mathrm{var}_{\hat I}\, \mathcal{L}_\epsilon^{n_0}\varphi = \sum_k \mathrm{var}_{E_k}\, \mathcal{L}_\epsilon^{n_0}\varphi$, for $\epsilon \geq 0$. Note that for $k > n_0$, Lemma 6 yields $\mathrm{var}_{E_k}\, \mathcal{L}_\epsilon^{n_0}\varphi \leq \lambda^{-n_0}\, \mathrm{var}_{E_{k-n_0}}\, \varphi$. For $k \leq n_0$, we first use Lemma 6 (k times), then invoke $\sup_{E_0} |\mathcal{L}_\epsilon^\ell \varphi| \leq \mathrm{var}_{E_0}\, \mathcal{L}_\epsilon^\ell \varphi + C \int_{E_0} |\mathcal{L}_\epsilon^\ell \varphi|\, d\mu_\epsilon$ (for $\ell = n_0 - k$), finally apply the Sublemma (for $A = E_0$ and $n = n_0 - k \leq n_0$) and Lemma 8. In this way we get

$$\operatorname*{var}_{E_k} \mathcal{L}_\epsilon^{n_0}\varphi \leq C\lambda^{-k}\big[C\sigma^{-(n_0-k)}(\operatorname*{var}_{\hat I}\varphi + \sup_{\hat I}|\varphi|) + C(n_0)\int_{\hat I}|\varphi|\, d\mu_\epsilon\big]\,,$$

for all $\epsilon < \epsilon(n_0)$. Therefore (recall that $\sigma \leq \lambda$),

$$\operatorname*{var}_{\hat I} \mathcal{L}_\epsilon^{n_0}\varphi \leq \lambda^{-n_0}\operatorname*{var}_{\hat I}\varphi + n_0 C\sigma^{-n_0}(\operatorname*{var}_{\hat I}\varphi + \sup_{\hat I}|\varphi|) + n_0 C(n_0)\int_{\hat I}|\varphi|\, d\mu_\epsilon\,.$$

Hence, for each fixed $\sigma_0 < \bar\sigma_0 < \sigma$ there is $C > 0$ and there are $C(n_0) > 0$ and $\epsilon(n_0) > 0$ such that

$$\operatorname*{var}_{\hat I} \mathcal{L}_\epsilon^{n_0}\varphi \leq C\bar\sigma_0^{-n_0}(\operatorname*{var}_{\hat I}\varphi + \sup_{\hat I}|\varphi|) + C(n_0)\int_{\hat I}|\varphi| d\mu_\epsilon\,, \tag{4.17}$$

for all $\epsilon < \epsilon(n_0)$. We also need an analogue of this inequality for the supremum:

$$\sup_{\hat I} |\mathcal{L}_\epsilon^{n_0}\varphi| \leq C\bar\sigma_0^{-n_0}(\operatorname*{var}_{\hat I}\varphi + \sup_{\hat I}|\varphi|) + C(n_0)\int_{\hat I}|\varphi| d\mu_\epsilon\,. \tag{4.18}$$

To prove (4.18) distinguish two cases. If $\sup_{\hat I} |\mathcal{L}_\epsilon^{n_0}\varphi| = \sup_{\cup_{k>n_0} E_k} |\mathcal{L}_\epsilon^{n_0}\varphi|$, simply apply Lemma 5 repeatedly. Otherwise, use

$$\sup_{\hat I} |\mathcal{L}_\epsilon^{n_0}\varphi| \leq \sup_{k\leq n_0}\big(\operatorname*{var}_{E_k}\mathcal{L}_\epsilon^{n_0}\varphi + \frac{1}{\mu_\epsilon(E_k)}\int_{E_k} |\mathcal{L}_\epsilon^{n_0}\varphi|\, d\mu_\epsilon\big) \leq \operatorname*{var}_{\hat I}\mathcal{L}_\epsilon^{n_0}\varphi + C(n_0)\int_{\hat I}|\varphi|\, d\mu_\epsilon\,,$$

where we have invoked Lemma 8 and the fact that $\sup_{k\leq n_0}(1/\mu_\epsilon(E_k)) \leq C(n_0)$ if $\epsilon < \epsilon(n_0)$.

Now the lemma follows easily. Fix $q \geq 1$ large enough so that $2C\bar\sigma_0^{-q} < \sigma_0^{-q} < 1$ and then, for arbitrary $n \geq 1$, write $n = pq + r$ with $0 \leq r < q$. Using (4.17)-(4.18) recursively, p times with $n_0 = q$ and then once with $n_0 = r$ (together with Lemma 8),

$$\operatorname*{var}_{\hat I}(\mathcal{L}_\epsilon^{n}\varphi) \leq C\sigma_0^{-n}(\operatorname*{var}_{\hat I}\varphi + \sup_{\hat I}|\varphi|) + C\int_{\hat I}|\varphi|\, d\mu_\epsilon\,,$$

for all $n \geq 1$ and $0 \leq \epsilon < \epsilon(q)$ (the constants C depend only on q). The lemma is thus proved for $\epsilon = 0$. If $\epsilon > 0$, we also use $\int_{\hat I} |\varphi| d\mu_\epsilon \leq \int_{\hat I} |\varphi|\, d\mu_0 + \sup_{\hat I} |\varphi| \int_{\hat I} |w_0 - w_\epsilon|\, dx$, restricting if necessary to $\epsilon < \epsilon(n)$, some $\epsilon(n) > 0$, to ensure $\int_{\hat I} |w_0 - w_\epsilon|\, dx \leq \sigma_0^{-n}$ (recall Lemma 7). $\square$

Bounding the supremum.

The following is an easy consequence of (4.17)–(4.18) (use the same argument as in the proof of the Variation Lemma):

Supremum Lemma. *For each $\sigma_0 < \sigma$ there is a constant $C > 0$ and for each $n \geq 1$ there is an $\epsilon(n)$ such that for all $0 \leq \epsilon < \epsilon(n)$ and all $\varphi \in BV(\hat{I})$*

$$\sup_{\hat{I}} |\mathcal{L}_\epsilon^n \varphi| \leq C\sigma_0^{-n}(\operatorname*{var}_{\hat{I}} \varphi + \sup_{\hat{I}} |\varphi|) + C \int_{\hat{I}} |\varphi(x)| w_0(x)\, dx\,,$$

Remark. It follows from Lemma 8, the Variation Lemma and the Supremum Lemma that the operators $\mathcal{L}_\epsilon$ preserve the Banach space $BV(\hat{I})$ and are bounded for all $\epsilon \geq 0$.

Bounding the integrals.

We now deduce our final estimate:

Integral Lemma. *There is a constant $C > 0$ and for each $n \geq 1$ there is an $\epsilon(n) > 0$ such that for all $0 < \epsilon < \epsilon(n)$ and all $\varphi \in BV(\hat{I})$*

$$\int_{\hat{I}} |(\mathcal{L}_\epsilon^n \varphi - \mathcal{L}_0^n \varphi) w_0|\, dx \leq C\sigma^{-n}(\operatorname*{var}_{\hat{I}} \varphi + \sup_{\hat{I}} |\varphi|)\,.$$

Proof of the Integral Lemma. This time, we use the decomposition $\int_{\hat{I}} \mathcal{L}_\epsilon^n \varphi\, w_0\, dx = \int_{\hat{I}} \sum_{k\geq 0} \mathcal{L}_\epsilon^n(\varphi \chi_{E_k}) w_0\, dx$. We first bound the integral on the unbounded top part of the tower. Fix some large $N \geq n$ and write $T = \cup_{k\geq N-n} E_k$. Then, using the positivity of $\mathcal{L}_\epsilon$, Lemmas 7–8, and the Supremum Lemma (recall also from the proof of Lemma 7, e.g. (3.22), that w_ϵ is integrable over $\hat{I}$ for $\epsilon \geq 0$), we have

$$\begin{aligned}
\int_{\hat{I}} |\mathcal{L}_\epsilon^n(\varphi\chi_T)| w_0\, dx &\leq \int_{\hat{I}} |\varphi| \chi_T\, w_\epsilon\, dx + \sup_{\hat{I}} \mathcal{L}_\epsilon^n(|\varphi|\chi_T) \int_{\hat{I}} |w_0 - w_\epsilon|\, dx \\
&\leq \sup_{\hat{I}} |\varphi| \mu_\epsilon(T) + \big[C\sigma_0^{-n}(\sup_{\hat{I}} |\varphi| + \operatorname*{var}_{\hat{I}} \varphi) + C\int_{\hat{I}} |\varphi| w_0\, dx\big] \cdot \int_{\hat{I}} |w_0 - w_\epsilon|\, dx \\
&\leq \sup_{\hat{I}} |\varphi| \frac{1}{C_n(N)} + (\sup_{\hat{I}} |\varphi| + \operatorname*{var}_{\hat{I}} \varphi) c(\epsilon) \leq C\sigma^{-n}(\operatorname*{var}_{\hat{I}} \varphi + \sup_{\hat{I}} |\varphi|)\,, \qquad (4.19)
\end{aligned}$$

for $0 \leq \epsilon < \epsilon(n)$, as long as N is fixed large enough, depending only on n.

It remains to control the bottom part of the tower $B = \cup_{k<N-n} E_k$. We shall do this by "trimming" intervals as in [BaY, Section 5]. Our notations are as in the last subsection of Section 2. First, we note that $B \subset \cup_{\eta_0} \eta(\vec{t}, \eta_0) \subset \cup_{k\leq N} E_k$ for all $\vec{t} \in J_\epsilon^n$, the union being over all $\eta_0 \in \mathcal{Z}_0^{n,N}$. Given any such η_0 let $\zeta^-(\epsilon, \eta_0) = \cap_{\vec{t}\in J_\epsilon^n} \hat{f}_{\vec{t}}^n(\eta(\vec{t}, \eta_0))$ and define $\zeta^+(\epsilon, \eta_0)$ in a similar way, replacing intersection by union. Let $\ell = \ell(\eta_0)$ be defined by $\eta_0 \subset E_\ell$. Clearly, given $(x, k) \in \zeta^-(\epsilon, \eta_0) \subset E_k$ and $\vec{t} \in J_\epsilon^n$, there is

exactly one $y_{\vec{t}} = y_{\vec{t}}(\eta_0)$ such that $(y_{\vec{t}}, \ell) \in \eta(\vec{t}, \eta_0)$ and $\hat{f}^n_{\vec{t}}(y_{\vec{t}}, \ell) = (x, k)$. Now we define $\xi(\epsilon, \eta_0) \subset \zeta^-(\epsilon, \eta_0)$ by the condition that either $y_{\vec{0}} \in \operatorname{supp} w_0$ or else $y_{\vec{t}} \notin \operatorname{supp} w_\epsilon$ for all $\vec{t} \in J^n_\epsilon$. Observe that, for each fixed η_0 and n, all three sets $\zeta^\pm(\epsilon, \eta_0)$ and $\xi(\epsilon, \eta_0)$ converge to $\hat{f}^n(\eta_0)$ as $\epsilon \to 0$. Set $Y_{\epsilon,n} = \cup_{\eta_0} (\zeta^+(\epsilon, \eta_0) \setminus \xi(\epsilon, \eta_0))$. Then, using $\#\mathcal{Z}_0^{n,N} \le C(n)$ (recall that N is fixed depending only on n), we get $\mu_0(Y_{\epsilon,n}) \le c_n(\epsilon)$ Together with the Supremum Lemma, this gives

$$\int_{Y_{\epsilon,n}} |\mathcal{L}^n_\epsilon(\varphi\chi_B)| w_0 \, dx \le (\sup_{\hat{I}} |\varphi| + \operatorname*{var}_{\hat{I}} \varphi) c_n(\epsilon) \le C\sigma^{-n} (\operatorname*{var}_{\hat{I}} \varphi + \sup_{\hat{I}} |\varphi|) \tag{4.20}$$

for all $0 \le \epsilon < \epsilon(n)$. On the other hand, clearly, $(\mathcal{L}^n_\epsilon(\varphi\chi_B) \cdot w_0)(x, k) = 0$ if (x, k) does not belong in $X_{\epsilon,n} = \cup_{\eta_0} \zeta^+(\epsilon, \eta_0)$. Hence, we are left to bound

$$\int_{(X_{\epsilon,n} \setminus Y_{\epsilon,n})} |(\mathcal{L}^n_\epsilon - \mathcal{L}^n_0)(\varphi\chi_B)| w_0 \, dx \le C \sup_{(X_{\epsilon,n} \setminus Y_{\epsilon,n})} |(\mathcal{L}^n_\epsilon - \mathcal{L}^n_0)(\varphi\chi_B)| \,. \tag{4.21}$$

For the rest of the proof we fix an arbitrary $(x, k) \in (X_{\epsilon,n} \setminus Y_{\epsilon,n}) \cap \operatorname{supp} \mu_0$. By definition,

$$\begin{aligned} (\mathcal{L}^n_\epsilon - &\mathcal{L}^n_0)(\varphi\chi_B)(x, k) = \\ &= \sum_{\eta_0} \int (\varphi g^{(n)}_{\vec{t}})(y_{\vec{t}}(\eta_0), \ell(\eta_0)) \, d\vec{\theta}_\epsilon(\vec{t}) - (\varphi g^{(n)})(y(\eta_0), \ell(\eta_0)) \,, \end{aligned} \tag{4.22}$$

where $g^{(n)}$, $g^{(n)}_{\vec{t}}$ are the iterated weight functions introduced in Section 2, $y = y_{\vec{0}}$, and the sum is over all $\eta_0 \in \mathcal{Z}_0^{n,N}$ with $\eta_0 \subset B$ and $(x, k) \in \xi(\epsilon, \eta_0)$. We fix η_0 and consider two possibilities. If $y \notin \operatorname{supp} w_0$ then $y_{\vec{t}} \notin \operatorname{supp} w_\epsilon$ for all $\vec{t}$, recall the definition of $\xi(\epsilon, \eta_0)$. This implies $g^{(n)}_{\vec{t}}(y_{\vec{t}}) = g^{(n)}(y) = 0$, hence η_0 does not contribute to (4.22). In the opposite case, the term in (4.22) corresponding to η_0 is bounded by (we omit the reference to η_0 in ℓ, y and $y_{\vec{t}}$)

$$\begin{aligned} \int |\varphi(y_{\vec{t}}, \ell) - \varphi(y, \ell)| g^{(n)}_{\vec{t}}(y_{\vec{t}}, \ell) d\vec{\theta}_\epsilon(\vec{t}) + \int |\varphi(y, \ell)| |g^{(n)}_{\vec{t}}(y_{\vec{t}}, \ell) - g^{(n)}(y, \ell)| d\vec{\theta}_\epsilon(\vec{t}) \\ \le \operatorname*{var}_{\eta_0^+} \varphi \cdot \int g^{(n)}_{\vec{t}}(y_{\vec{t}}, \ell) d\vec{\theta}_\epsilon(\vec{t}) + \sup_{\hat{I}} |\varphi| \int \big| g^{(n)}_{\vec{t}}(y_{\vec{t}}, \ell) - g^{(n)}(y, \ell) \big| d\vec{\theta}_\epsilon(\vec{t}) \,, \end{aligned} \tag{4.23}$$

with $\eta_0^+ = \eta^+(\epsilon, \eta_0)$. Note that $g^{(n)}(y, \ell) = \lambda^{\ell - k} \le \sigma^{-n}$ if $\ell + n - k = 0$. For $\ell + n - k > 0$ we claim that

$$g^{(n)}(y, \ell) = \frac{w_0(y, \ell)}{w_0(x, k)} \frac{1}{|(f^n)'(y)|} = \frac{\lambda^{\ell - k} h(x, k)}{|(f^{n+\ell-k})'(\hat{f}_+^{-\ell}(y, \ell))|} \le C\sigma^{-n} \,, \tag{4.24}$$

where $h(x, k) = |f'(\hat{f}_+^{-k}(x, k)/f'(f^{n-k}(y)|$. Indeed, we use that $\hat{f}_+^{-k}(x, k)$ is either $f^{n-k}(y)$ or $(f^{n-k}(y))_-$ (recall the notations x_-, $K(x)$ introduced before Lemma 5) for the second equality, the we obtain the last inequality by applying Lemma 2 to the $\hat f^{n+\ell-k}$-trajectory of $(\hat f_+^{-\ell}(y,\ell),0)$ (recall that δ is fixed) and noting that $|h(x,k)|\le K$. Now for $\vec t\in J_\epsilon^n$, $\vec u\in J_\epsilon^k$, and $\vec v\in J_\epsilon^\ell$, we write $x_{\vec u}=\hat f^{-k}_{\vec u,+}(x,k)$ and $y_{\vec t,\vec v}=\hat f^{-\ell}_{\vec v,+}(y_{\vec t},\ell)$, whenever these objects are defined. Then, restricting to the case $\ell+n-k>0$ (the case $\ell+n-k=0$ is simpler) and recalling the definition of the perturbed cocycles w_ϵ,

$$
\begin{aligned}
\int g^{(n)}_{\vec t}(y_{\vec t},\ell)\,d\theta_\epsilon(\vec t) &= \lambda^{\ell-k}\int \frac{\int |(f^\ell_{\vec v})'(y_{\vec t,\vec v})|^{-1}\,d\vec\theta_\epsilon(\vec v)}{\int |(f^k_{\vec u})'(x_{\vec u})|^{-1}\,d\vec\theta_\epsilon(\vec u)}|(f^n_{\vec t})'(y_{\vec t})|^{-1}\,d\vec\theta_\epsilon(\vec t)\\
&=\lambda^{\ell-k}\int\frac{\int h(x,\vec r)|(f^k_{\vec r})'(x_{\vec r})|^{-1}\,d\vec\theta_\epsilon(\vec r)}{\int |(f^k_{\vec u})'(x_{\vec u})|^{-1}\,d\vec\theta_\epsilon(\vec u)}|(f^{n+\ell-k}_{\vec s})'(y_{\vec t,\vec v})|^{-1}\,d\vec\theta_\epsilon(\vec s)\,,
\end{aligned}
\tag{4.25}
$$

where we write $\vec r=(t_{n-k+1},\dots,t_n)$ and $\vec s=(v_1,\dots,v_\ell,t_1,\dots,t_{n-k})$, and denote $h(x,\vec r)=|f'(x_{\vec r})/f'(f^{n+\ell-k}_{\vec s}(y_{\vec t,\vec v}))|$. Observe that $f^{n+\ell-k}_{\vec s}(y_{\vec t,\vec v})$ is either $x_{\vec r}$ or $(x_{\vec r})_-$, the choice between the two possibilities depending only on (x,k) and η_0 for small enough $\epsilon(n)$. Hence, Lemma 2 yields

$$
\int g^{(n)}_{\vec t}(y_{\vec t},\ell)\,d\theta_\epsilon(\vec t)\le\lambda^{\ell-k}K\sup_{\vec t,\vec v}|(f^{n+\ell-k}_{\vec s})'(y_{\vec t,\vec v})|^{-1}\le C\sigma^{-n}\,. \tag{4.26}
$$

To control the last term in (4.23), we bound $l|g^{(n)}_{\vec t}(y_{\vec t},\ell)-g^{(n)}(y,\ell)|$ by

$$
\frac{g^{(n)}_{\vec t}(y_{\vec t},\ell)}{h(x,\vec r)}|h(x,\vec r)-h(x,k)|+g^{(n)}(y,\ell)\left|\frac{g^{(n)}_{\vec t}(y_{\vec t},\ell)}{g^{(n)}(y,\ell)}\frac{h(x,k)}{h(x,\vec r)}-1\right|, \tag{4.27}
$$

and observe first that

$$
\int\left|\frac{g^{(n)}_{\vec t}(y_{\vec t},\ell)}{g^{(n)}(y,\ell)}\frac{h(x,k)}{h(x,\vec r)}-1\right|d\theta_\epsilon(\vec t)\le\sup_{\vec t,\vec v}\left|\frac{|(f^{n+\ell-k})'(\hat f_+^{-\ell}(y,\ell))|}{|(f^{n+\ell-k}_{\vec s})'(y_{\vec t,\vec v})|}-1\right|\le c_n(\epsilon)\,. \tag{4.28}
$$

Indeed, the first inequality follows from previous considerations, see (4.24)-(4.25). Moreover, by construction, $(f^{n+\ell-k}_{\vec s})'(y_{\vec t,\vec v})$ has the same sign as $(f^{n+\ell-k})'(\hat f_+^{-\ell}(y,\ell))$ and their difference is bounded by (note that $\hat f_+^{-\ell}(y,\ell)=y_{\vec 0,\vec 0}$)

$$
\sup_{z,\vec r}|(f^{n+\ell-k}_{\vec r})''(z)|\cdot|y_{\vec t,\vec v}-\hat f_+^{-\ell}(y,\ell)|\le C(n)c_n(\epsilon)\le c_n(\epsilon)\,.
$$

Combining this with Lemma 2 (to bound the denominator) we obtain the second inequality in (4.28). Observe, moreover, that the first term in (4.27) is also bounded by $c_n(\epsilon)$, because $1/h(x,\vec r)\le K$ and $|h(x,\vec r)-h(x,k)|\le c_n(\epsilon)$ (use here that $K(x)$ is Lipschitz). Therefore, we obtain from (4.23)–(4.28) that (4.22) is bounded above by

$$
\sum_{\eta_0\in\mathcal Z_0^{n,N}}\big(C\sigma^{-n}\operatorname*{var}_{\eta_0^+}\varphi+C\sigma^{-n}c_n(\epsilon)\sup_{\hat I}|\varphi|\big)\le C\sigma^{-n}\operatorname*{var}_{\hat I}\varphi+\sigma^{-n}c_n(\epsilon)\sup_{\hat I}|\varphi|
$$

for all $0 < \epsilon < \epsilon(n)$ (use the bounded overlap property to sum the variations and $\#\mathcal{Z}_0^{n,N} \le C(n)$ to sum the suprema). Replacing in (4.21),

$$\int_{(X_{\epsilon,n}\setminus Y_{\epsilon,n})} |(\mathcal{L}_\epsilon^n - \mathcal{L}_0^n)(\varphi\chi_B)| w_0 \, dx \le C\sigma^{-n}(\operatorname*{var}_{\hat I} \varphi + \sup_{\hat I} |\varphi|) \tag{4.29}$$

Together, (4.19), (4.20), and (4.29) prove the lemma. □

In Section 5, we shall also need the following version of the integral lemma:

Nonuniform Integral Lemma. $\int_{\hat I} |\mathcal{L}_\epsilon^n \varphi - \mathcal{L}_0^n \varphi| \, d\mu_0 \le c_n(\epsilon)$ *for each fixed* $\varphi \in BV(\hat I)$.

Proof of the Nonuniform Integral Lemma. Fix $\varphi \in BV(\hat I)$ and $n \ge 1$ and let $\tau > 0$. Reading the proof of the Integral Lemma, first choose N so as to make $1/C_n(N) \le \tau$ in (4.19) (restricting to $\epsilon < \epsilon(n, N)$ if necessary). Next, since φ has bounded variation, there are $\bar\epsilon = \bar\epsilon_{n,\varphi}(\tau) > 0$ and $E = E_{n,\varphi}(\tau) \subset \hat I$ with Lebesgue measure $|E| \le \tau$ (hence $\mu_0(E) \le c(\tau)$), such that $|\varphi(y_{\vec t}, \ell) - \varphi(y, \ell)| \le \tau$ for all $\vec t \in J_{\bar\epsilon}^n$ whenever $(y, \ell) \notin E$. For $(x, k) \in X_{n,\epsilon} \setminus (\hat f^n(E) \cup Y_{n,\epsilon})$ this permits us to replace $\operatorname{var}_{\eta_0^+} \varphi$ by τ in (4.23) and so also $C\sigma^{-n} \operatorname{var}_{\hat I} \varphi$ by $C(n)\tau$ in (4.29). On the other hand, in just the same way as in (4.20), the integral of $|\mathcal{L}_\epsilon(\varphi\chi_B)|$ over $\hat f^n(E)$ is bounded by $c(\tau)$. Finally, take $\epsilon > 0$ small enough so that the factors $c(\epsilon)$, $c_n(\epsilon)$ in (4.19), (4.20), (4.29) be smaller than τ. Putting all this together, we conclude that $\int |(\mathcal{L}_\epsilon^n - \mathcal{L}_0^n)(\varphi)| w_0 dx \le c_n(\tau)$ if $\epsilon > 0$ is small. Since $\tau > 0$ is arbitrary, this proves the lemma. □

Balanced norms.

We introduce a family of equivalent norms on $BV(\hat I)$, defined for $0 < \zeta \le 1$ by

$$\|\varphi\|_\zeta = \zeta \cdot (\operatorname*{var}_{\hat I} \varphi + \sup_{\hat I} |\varphi|) + \int_{\hat I} |\varphi| \, d\mu_0 \, . \tag{4.30}$$

(Analogous "balanced" norms are used e.g. in [BaY].) We now state an immediate consequence of the Variation, Supremum and Integral Lemmas:

Dynamical Lemma. *For any* $\bar\tau$ *with* $\sigma^{-1} < \bar\tau^2 < 1$, *there is* $C > 0$, *and for any* $n \ge 0$ *there is* $\epsilon(n) > 0$ *such that for each* $0 < \epsilon < \epsilon(n)$ *we have* $\|\mathcal{L}_\epsilon^n - \mathcal{L}_0^n\|_{\bar\tau^n} \le C \cdot \bar\tau^n$.

5. Conclusion of the proof

We first reprove the result (essentially due to [BC], see [MS]) that f satisfies the conclusion of Jakobson's theorem [Ja]:

Corollary 1 (Invariant measure). *Let* f *satisfy assumptions (A1)–(A2). Then* f *has an absolutely continuous invariant probability measure* $m_0 = \rho_0 \, dx$.

Proof of Corollary 1. (The arguments are fairly standard, see e.g. [Ry].) We start by constructing an absolutely continuous $\hat f$-invariant probability measure $\hat m_0$. Consider the sequence of nonnegative functions $\varphi_n = \left(\frac{1}{n}\sum_{i=0}^{n-1}\mathcal{L}_0^i(\chi_{E_0})\right)$. It follows from the Variation Lemma (for $\epsilon = 0$) that $\mathrm{var}_{\hat{I}}\,\varphi_n \le C$, and from the Supremum Lemma that $\sup_{\hat{I}}\varphi_n \le C$, uniformly in n. By Helly's theorem, a subsequence of φ_n converges in $L^1(\hat{I}, dx)$ to some $\hat{\rho}_0 \in BV(\hat{I})$ which, by construction, is a fixed function of $\mathcal{L}_0$. By Lemma 8, $\int \hat{\rho}_0\,d\mu_0 > 0$ and $\hat{m}_0 = \hat{\rho}_0\mu_0$ is an $\hat{f}$-invariant measure. We replace $\hat{\rho}_0$ by $\hat{\rho}_0/\int\hat{\rho}_0\,d\mu_0$ in what follows. Now, let $\mathcal{P} : L^1(I, dx) \to L^1(I, dx)$ be the usual Perron-Frobenius operator for f (i.e., $\mathcal{P}(\varphi) = \sum_{fy=x, y\neq 0}\varphi(y)/|f'(y)|$), and $\hat{\mathcal{P}}$ be the usual Perron-Frobenius operator for $\hat{f}$. Then, $\hat{\mathcal{P}}(\hat{\rho}_0 w_0) = \hat{\rho}_0 w_0$, and an easy computation shows that $\mathcal{P}(\rho_0) = \rho_0$, where $\rho_0(x) = \sum_{k=0}^{\infty}(\hat{\rho}_0 w_0)\circ(\pi|_{E_k})^{-1}(x)$. Note that $\int_I \rho_0\,dx = 1$ because $\int_{\hat{I}}\hat{\rho}_0 w_0\,dx = 1$. By well-known arguments, the probability measure $m_0 = \rho_0\,dx$ is f-invariant. $\square$

Remark. Lemma 5 applied to $\hat{\rho}_0$ provides the following additional information on the measure m_0: $\sup_{E_k}\hat{\rho}_0 \le \mathrm{const}\,\lambda^{-k}$ for all $k \ge 0$, in particular $\sum_k \sup_{E_k}\hat{\rho}_0$ is finite.

Condition (A1) combined with [Si] ensures that f has no periodic attractors. Hence, by a result of Blokh-Lyubich (see e.g. [MS, Theorem V.1.2] for a statement and references), f is ergodic with respect to Lebesgue measure. It is easy to deduce that f is ergodic with respect to m_0 and, moreover, that this is the unique absolutely continuous invariant probability measure of f (see e.g. [MS, Theorem V.1.5]).

By [BL], the entropy of f with respect to this absolutely continuous invariant probability measure is strictly positive. Now, since we also assume (A3), f^n is ergodic with respect to m_0 for all $n \ge 1$. Indeed, almost every ergodic component of m_0 for f^n is absolutely continuous [Le, Corollary 4]; there are finitely many such components and their supports consist of finitely many intervals [Yo, Proposition 3.3]; the topological mixing assumption (A3) then implies that these supports must all coincide, hence m_0 is ergodic for f^n. Therefore, [Le, Theorem 1] gives that the natural extension of (f, m_0) is Bernoulli, and thus (f, m_0) is exact (that is $\cap_{n\ge 0} f^{-n}(\mathcal{B})$ contains only zero or full m_0-measure sets, where $\mathcal{B}$ is the Borel σ-algebra of I). This conclusion plays a central role in the proof of the next corollary.

Let us also note that in this context (A3) may be formulated

(A3) for any interval $J \subset I$ there is $n \ge 1$ such that $f^n(J)$ contains the interval bounded by c_1 and c_2 (which coincides with $f^k(I)$ for all $k \ge 2$),

and is equivalent to f being non-renormalisable, see [BL]. Finally, [Yo, Lemma 2.1], for quadratic maps with parameter a close to 2, conditions (A1)–(A2) imply (A3). Combined with [BC], this gives that the three conditions hold simultaneously for a positive measure set of values of a.

Corollary 2 (Quasicompacity). *Assume (A1)–(A3). The spectrum of the operator $\mathcal{L}_0$ acting on $BV(\hat{I})$ decomposes as $\Sigma(\mathcal{L}_0) = \Sigma_1 \cup \{1\}$, where $\tau_0 = \sup\{|z| \mid z \in \Sigma_1\} < 1$ and 1 is a simple eigenvalue with a positive eigenfunction $\hat{\rho}_0$ and spectral projection $\pi_0(\varphi) = \hat{\rho}_0\int_{\hat{I}}\varphi\,d\mu_0$. Moreover, the essential spectral radius of $\mathcal{L}_0$ is at most $1/\sigma < 1$.*

Proof of Corollary 2. We shall first show that the essential spectral radius of $\mathcal{L}_0$ acting on $BV(\hat{I})$ is at most $1/\sigma < 1$. Since the Variation Lemma, the Supremum Lemma, and

Lemma 8 imply that $\|\mathcal{L}_0^n\|$ is uniformly bounded, in particular the spectral radius of $\mathcal{L}_0$ is equal to 1, it will immediately follow that the spectrum of $\mathcal{L}_0$ decomposes as the union of a finite set of eigenvalues of finite multiplicity on the unit circle and a compact subset of a disc of radius $\tau_0 < 1$ (note that τ_0 is either the essential spectral radius of $\mathcal{L}_0$ or the modulus of the second largest eigenvalue). We have already observed that (A3) ensures exactness of (f, m_0). We shall then deduce that $(\hat{f}, \hat{m}_0)$ is also exact and, from this, that the only eigenvalue of $\mathcal{L}_0$ on the unit circle is 1, and is simple.

Lemmas 1 and 2 (see also (4.24)) imply that there is a constant $C > 0$ such that $\sup_{\hat{I}} g^{(n)} \leq C\sigma^{-n}$ for all $n \geq 1$. Let $N \geq n$ be such that $\mu_0(\cup_{k>N} E_k) < \sigma^{-n}$, and define projections α_n and $\alpha_{n,N} : BV(\hat{I}) \to BV(\hat{I})$ by chosing an arbitrary point x_η in each monotonicity interval $\eta \in \mathcal{Z}_0^n$ and setting

$$\alpha_n(\varphi) = \sum_{\eta \in \mathcal{Z}_0^n} \varphi(x_\eta)\, \chi_\eta \qquad \text{and} \qquad \alpha_{n,N}(\varphi) = \alpha_n\big(\varphi \cdot \chi_{(\cup_{k \leq N} E_k)}\big)\,.$$

We first bound $\|\mathcal{L}^n - \mathcal{L}^n \alpha_n\|_{BV}$, using (a)–(c) from Section 2:

$$\sup_{\hat{I}} |(\mathcal{L}^n - \mathcal{L}^n \alpha_n)\varphi| \leq \sup_y \sum_{\substack{\eta \in \mathcal{Z}_0^n \\ y \in f^n(\eta)}} g^{(n)}(y_\eta)|\varphi(y_\eta) - \varphi(x_\eta)| \leq C\sigma^{-n} \operatorname*{var}_{\hat{I}} \varphi$$

$$\begin{aligned}\operatorname*{var}_{\hat{I}}(\mathcal{L}^n - \mathcal{L}^n \alpha_n)(\varphi) \leq \sum_{\eta \in \mathcal{Z}_0^n} \Big(&\operatorname*{var}_\eta g^{(n)} \sup_\eta |\varphi - \varphi(x_\eta)| + \sup_\eta g^{(n)} \operatorname*{var}_\eta |\varphi - \varphi(x_\eta)| \\ &+ 2 \sup_\eta g^{(n)} \sup_\eta |\varphi - \varphi(x_\eta)|\Big) \leq C\sigma^{-n} \operatorname*{var}_{\hat{I}} \varphi\end{aligned}$$

$$\int_{\hat{I}} |(\mathcal{L}^n - \mathcal{L}^n \alpha_n)\varphi|\, d\mu_0 \leq C \sup_{\hat{I}} |(\mathcal{L}^n - \mathcal{L}^n \alpha_n)\varphi| \leq C\sigma^{-n} \operatorname*{var}_{\hat{I}} \varphi\,,$$

with $y_\eta = (f^{-n}|_\eta)(y)$ (we used $Sf < 0$ and the fact that $K(x)$ has bounded variation, see again (4.24), to get $\operatorname{var}_\eta g^{(n)} \leq C \sup_\eta g^{(n)}$ for all $\eta \in \mathcal{Z}_0^n$). Now, for any fixed $\sigma_0 < \sigma$, the above bounds together with the Supremum Lemma yield

$$\begin{aligned}\|(\mathcal{L}^n - \mathcal{L}^n \alpha_{n,N})\varphi\| &\leq \|(\mathcal{L}^n - \mathcal{L}^n \alpha_n)(\varphi \cdot \chi_{(\cup_{k \leq N} E_k)})\| + \|\mathcal{L}^n(\varphi \cdot \chi_{(\cup_{k>N} E_k)})\| \\ &\leq C\sigma^{-n} \operatorname*{var}_{\hat{I}} \varphi + C\sigma_0^{-n}(\operatorname*{var}_{\hat{I}} \varphi + \sup_{\hat{I}} |\varphi|) + C \int_{\hat{I}} |\varphi \cdot \chi_{(\cup_{k>N} E_k)}|\, d\mu_0 \\ &\leq C\sigma_0^{-n}(\operatorname*{var}_{\hat{I}} \varphi + \sup_{\hat{I}} |\varphi|) + C \sup_{\hat{I}} |\varphi| \mu_0(\cup_{k>N} E_k) \leq C\sigma_0^{-n} \|\varphi\|\,,\end{aligned}$$

with constants independent of N. Since each $\alpha_{n,N}$ has finite-dimensional range and is therefore compact, the essential spectral radius of $\mathcal{L}_0$ is not bigger than $1/\sigma$, as claimed.

We now go to the second part of our argument. We start by claiming that, given any $\hat{A} \in \cap_{n \geq 0} \hat{f}^{-n}(\hat{\mathcal{B}})$, where $\hat{\mathcal{B}}$ is the Borel σ-algebra of $\hat{I}$, there is a Borel set $A \subset I$ such that $\hat{A} = \pi^{-1}(A)$, up to a zero $\hat{m}_0$-measure set. In what follows we disregard zero measure sets, in particular we always restrict to $\operatorname{supp}\hat{m}_0$, respectively $\operatorname{supp} m_0$. Note that $\hat{B} \in \hat{f}^{-n}(\hat{\mathcal{B}})$ if and only if $\hat{B} \in \hat{\mathcal{B}}$ and satisfies $[\xi \in \hat{B},\ \hat{f}^n(\xi) = \hat{f}^n(\eta)] \Rightarrow \eta \in \hat{B}$. Also, there is $B \subset I$ such that $\hat{B} = \pi^{-1}(B)$ if and only if $[(z,k) \in \hat{B},\ (z,\ell) \in \hat{I}] \Rightarrow (z,\ell) \in \hat{B}$. Therefore, in order to prove our claim it suffices to show that, given ($\hat{m}_0$-almost) any (z,k), $(z,\ell) \in \hat{I}$, there is $j \geq 0$ such that $\hat{f}^j(z,k) = \hat{f}^j(z,\ell)$. For $x \in (-\delta,\delta)$ we define $p(x)$ to be the "falling time" of x, that is the smallest integer $j \geq 1$ such that $\hat{f}^{j+1}(x,0) \in E_0$. Then, $e^{-\beta_1 p(x)} \geq (1/C)|(f^{p(x)-1})'(c_1)||c_1 - f(x)| \geq (1/C)\lambda_c^{p(x)} x^2$: this is proved in the same way as (3.10) (using the first inequality of (3.9) instead). Moreover, $p(x) \geq H(\delta)$ and so, if $\delta > 0$ is small enough, the previous inequality implies

$$\lambda_c^{p(x)} x^2 \leq \gamma^2\,, \qquad \text{where we write } \gamma = (1 - e^{\alpha-\beta_1}) \min\{1, |c_1|, \ldots, |c_{H_0}|\}. \tag{5.1}$$

Now let (z,k), $(z,\ell) \in \hat{I}$. Note that $\pi(\hat{f}^j(z,k)) = \pi(\hat{f}^j(z,\ell)) = f^j(z)$ for every $j \geq 0$. We suppose that the f-orbit of z is disjoint from the critical orbit (this excludes only a countable set), so that $p(f^j(z))$ is always finite. It is not difficult to see that, either there is $j \geq 0$ such that both $\hat{f}^j(z,k)$ and $\hat{f}^j(z,\ell)$ belong in E_0, or else there are $0 < \nu_1 < \nu_2 < \cdots$ with $f^{\nu_i}(z) \in (-\delta,\delta)$ and $\nu_{i+1} \leq \nu_i + p(f^{\nu_i}(z))$ (each point starts climbing up the tower again before the other one falls down) for all $i \geq 1$. In the first case, it must be $\hat{f}^j(z,k) = \hat{f}^j(z,\ell)$, which proves our claim. In the second one, we write $p_i = p(f^{\nu_i}(z))$ and note that $\nu_{i+1} - \nu_i \leq p_i$ implies $|f^{\nu_{i+1}}(z) - c_{(\nu_{i+1}-\nu_i)}| \leq e^{-\beta_1(\nu_{i+1}-\nu_i)}$ which, together with (A1)-(2.5), yields $|f^{\nu_{i+1}}(z)| \geq \gamma e^{-\alpha(\nu_{i+1}-\nu_i)} \geq \gamma e^{-\alpha p_i}$. In view of (5.1) and our assumption $e^{2\alpha} < \sqrt{\lambda_c}$, this gives $p_{i+1} \leq (p_i/2)$ for every $i \geq 1$. Since the p_i are positive integers, we conclude that the sequence ν_i is necessarily finite. This means that one eventually gets into the first case, thence the claim is proved.

As a consequence, $(\hat{f}, \hat{m}_0)$ is exact. Indeed, given any $\hat{A} \in \cap_{n\geq 0} \hat{f}^{-n}(\hat{\mathcal{B}})$, take $A \subset I$ with $\hat{A} = \pi^{-1}(A)$. Clearly, $A \in \cap_{n\geq 0} f^{-n}(\mathcal{B})$ and so, since (f, m_0) is exact, $m_0(A) \cdot m_0(A^c) = 0$. On the other hand, recall the proof of Corollary 1, $m_0 = \pi_* \hat{m}_0$ and so $\hat{m}_0(\hat{A}) = m_0(A)$. It follows that $\hat{m}_0(\hat{A}) \cdot \hat{m}_0(\hat{A}^c) = 0$, as we wanted to prove. In particular, $\hat{f}$ is mixing with respect to $\hat{m}_0$. Combining this with the equality $\int_{\hat{I}} \psi \mathcal{L}_0^n(\varphi\hat{\rho}_0)\,d\mu_0 = \int_{\hat{I}} (\psi \circ \hat{f}^n)\varphi\hat{\rho}_0\,d\mu_0$ for ψ, $\varphi\hat{\rho}_0 \in BV(\hat{I})$ (use Lemma 8), it follows that $\mathcal{L}_0^n(\varphi\hat{\rho}_0)$ weakly converges in $L^1(\mu_0)$ to $\hat{\rho}_0 \int \varphi\hat{\rho}_0\,d\mu_0$ as $n \to \infty$, whenever $\varphi\hat{\rho}_0 \in BV(\hat{I})$.

Now let $\lambda_1 \in S^1$ be an eigenvalue of $\mathcal{L}_0$ and $\hat{\rho}_1 \in BV(\hat{I})$ be a corresponding eigenfunction. We claim that $\operatorname{supp}\hat{\rho}_1 \subset \operatorname{supp}\hat{\rho}_0$. In view of our definitions, it suffices to prove $\operatorname{supp}\hat{\rho}_1 \cap E_0 \subset \operatorname{supp}\hat{\rho}_0 \cap E_0$. Let $J(\delta)$ be the interval in E_0 bounded by $f^2(-\delta) \times \{0\}$ and $f(-\delta) \times \{0\}$ (since $(-\delta)_- \leq \delta$, we have $|f^i(\delta)| \leq |f^i(-\delta)|$, for $i = 1,2$). On the one hand, $\operatorname{supp}\hat{\rho}_1 \cap E_0 \subset J(\delta)$ because $\hat{f}^2(\hat{I}) \cap E_0 \subset J(\delta)$ (points (x,k) with $|x| < \delta$ have $\hat{f}^i(x,k) = (f^i(x), k+i)$ for $i = 1,2$). On the other hand, $\inf \hat{\rho}_0|_{J(\delta)} > 0$. Indeed (following [Yo]), let $J \subset \hat{I}$ be any interval with $\inf \hat{\rho}_0|_J > 0$. Since f is topologically mixing there is $n_1 \geq 1$ such that $\pi(\hat{f}^{n_1}(J))$ contains a neighbourhood of the (expanding) fixed point q of f. Then there is $n_2 \geq n_1$ such that $\hat{f}^{n_2}(J)$ contains a neighbourhood of $(q,0)$ in E_0. It is easy to deduce that $\hat{f}^{n_3}(J) \supset J(\delta)$ for some $n_3 \geq n_2$. This shows that $\hat{\rho}_0$ is positive on $J(\delta)$ and so proves the claim. Hence we may write $\hat{\rho}_1 = \varphi\hat{\rho}_0$

for some function φ. Applying the weak convergence statement above we get that $\lambda_1^n \hat\rho_1 = \mathcal{L}_0^n \hat\rho_1 \to \hat\rho_0 \int \hat\rho_1 \, d\mu_0$. This implies that $\lambda_1 = 1$ and $\hat\rho_1 = \hat\rho_0 \int \hat\rho_1 \, d\mu_0$. Finally, the eigenvalue 1 must also have algebraic multiplicity one because $\|\mathcal{L}_0^n\|$ is uniformly bounded. $\square$

¿From Corollary 2, we recover a result of [KN, Yo]. For a function $\varphi : I \to \mathbb{C}$ with bounded variation, we define $\|\varphi\| = \|\varphi\|_{BV} = \mathrm{var}_I \, \varphi + \sup_I |\varphi| + \int |\varphi| \, dx$.

Corollary 3 (Decay of correlations). *Let m_0 be the unique absolutely continuous invariant probability measure for f and let $\tau_0 < 1$ be as in Corollary 2. For any $\tau > \tau_0$ and $\varphi, \psi : I \to \mathbb{C}$ of bounded variation, there is $C = C(\tau, \|\varphi\|, \|\psi\|) > 0$ such that for all $n \geq 1$*

$$\Big|\int_I (\varphi \circ f^n) \psi \, dm_0 - \int_I \varphi \, dm_0 \cdot \int_I \psi \, dm_0\Big| \leq C\tau^n \, .$$

Proof of Corollary 3. The proof uses standard arguments, see e.g. [Yo]. Lifting a function ψ with bounded variation to $\hat\psi(x,k) = \psi(x)$, we have $\hat\psi \hat\rho_0 \in BV(\hat I)$, with $\|\hat\psi\hat\rho_0\| \leq \mathrm{const}\, \|\psi\|$ (recall that $\sum_k \sup_{E_k} \hat\rho_0 < \infty$). By definition of m_0, $\hat m_0$, and $\mathcal{L}_0$, we may thus write (using Lemma 8 for $\epsilon = 0$)

$$\Big|\int_I (\varphi \circ f^n)\psi \, dm_0 - \int_I \varphi \, dm_0 \int_I \psi \, dm_0\Big| = \Big|\int_{\hat I} (\hat\varphi \circ \hat f^n)\hat\psi \, d\hat m_0 - \int_{\hat I} \hat\varphi \, d\hat m_0 \int_{\hat I} \hat\psi \, d\hat m_0\Big| =$$
$$\Big|\int \hat\varphi \big[\mathcal{L}_0^n(\hat\psi\hat\rho_0) - \hat\rho_0(\int \hat\psi \hat\rho_0 \, d\mu_0)\big] d\mu_0\Big| = \Big|\int \hat\varphi \mathcal{L}_0^n(\pi_1(\hat\psi\hat\rho_0)) d\mu_0\Big| \leq \sup_{\hat I} |\hat\varphi| C(\tau, \hat\psi) \tau^{-n} \, ,$$

where π_1 is the spectral projection associated to Σ_1. $\square$

Our main result will now follow from a version of the perturbation theorems on families of linear operators in [BaY, Section 5.E and Erratum]: Let $(X, \|\cdot\|)$ be a complex Banach space and $(T_\epsilon, \epsilon \geq 0)$ be a family of bounded linear operators. Assume that the spectrum of T_0 decomposes as $\Sigma(T_0) = \Sigma_0 \cup \Sigma_1$ with $\Sigma_0 = \{1\}$ and $\kappa_1 = \sup\{|z| \mid z \in \Sigma_1\} < 1$. Let $X = X_0 \oplus X_1$ and $\pi_0 : X \to X_i$, $i = 0, 1$, be the corresponding vector space decomposition and projections. Assume further that X_0 is finite-dimensional. Now let $|\cdot|$ be a norm on X with $|x| \leq \|x\|$ for all x, and let $\|\cdot\|_\zeta$, $0 < \zeta \leq 1$ be the family of norms defined by $\|\cdot\|_\zeta = \zeta\|\cdot\| + (1-\zeta)|\cdot|$. Assume that π_0 is a bounded projection for the norm $|\cdot|$.

Perturbation Lemma ([BaY]). *Suppose that $\sup_{0 \neq x \in X_0} |T_\epsilon x - T_0 x|/|x| \to 0$ as $\epsilon \to 0$ and $|T_\epsilon|$ is uniformly bounded. Suppose also that there exists $\kappa_1 < \kappa < 1$ such that for each large enough $n \geq 1$ there exists $\epsilon(n) > 0$ with $\|T_\epsilon^n - T_0^n\|_{\kappa^n} \leq \kappa^n$ for each $0 < \epsilon < \epsilon(n)$. Then, for each small enough ϵ,*

(1) *The spectrum of T_ϵ splits as $\Sigma(T_\epsilon) = \Sigma_0^\epsilon \oplus \Sigma_1^\epsilon$ with $\kappa_1^\epsilon = \sup\{|z| \mid z \in \Sigma_1^\epsilon\} <$ $\inf\{|z| \mid z \in \Sigma_0^\epsilon\}$, and $\limsup_{\epsilon \to 0} \kappa_1^\epsilon \leq \max(\kappa_1/\kappa, \kappa)$.*
(2) *Let $X = X_0^\epsilon \oplus X_1^\epsilon$ be the associated decomposition. Then $dim\, X_0^\epsilon = dim\, X_0$, $\Sigma(T_\epsilon|_{X_0^\epsilon}) \to \Sigma(T_0|X_0)$, and $|\pi_0^\epsilon - \pi_0| \to 0$ as $\epsilon \to 0$.*

For the bound on κ_1^ϵ in (1), just note that the constant κ' in the proof of Lemma 1' in [BaY] need only satisfy $\max(\kappa_1/\kappa, \kappa) < \kappa' < 1$.

Proof of the Main Theorem. We will first check that the hypotheses of the Perturbation Lemma are satisfied, and then derive our theorem from the conclusion of this lemma.

We let $X = BV(\hat{I})$, $T_\epsilon = \mathcal{L}_\epsilon$ for $\epsilon \geq 0$, and consider the norms $\|\cdot\| = \|\cdot\|_{BV}$ and $|\cdot| = \int_{\hat{I}} |\cdot|\, d\mu_0$ (then $\|\cdot\|_\zeta$ is the balanced norm defined in (4.30)). The hypotheses on $\mathcal{L}_0$ are satisfied for $\kappa_1 = \tau_0$ because of Corollary 2, in particular $X_0 = \mathbb{C}\hat{\rho}_0$ has dimension 1. Note that $|\mathcal{L}_\epsilon| \leq 1 + c(\epsilon)$ for all ϵ, by Lemmas 7 and 8, and that the assumption on π_0 is obvious from Corollary 1. Fixing $1 > \kappa > \max(\sqrt{\tau_0}, \sqrt{1/\sigma})$, the bounds on the differences $\mathcal{L}_\epsilon^n - \mathcal{L}_0^n$ follow from the Nonuniform Integral Lemma (for $\varphi = \hat{\rho}_0$) and the Dynamical Lemma of Section 4. It follows that for all small enough ϵ, the essential spectral radius of $\mathcal{L}_\epsilon$ is smaller than $\max(\kappa_1/\kappa, \kappa)$, which can be made arbitrarily close to $\max(\sqrt{\tau_0}, \sqrt{1/\sigma}) < 1$. By the same arguments as in Corollary 1, $\mathcal{L}_\epsilon$ has a positive fixed function $\hat{\rho}_\epsilon \in BV(\hat{I})$, in particular we have $\Sigma_0^\epsilon = \{1\}$ and $X_0^\epsilon = \mathbb{C}\hat{\rho}_\epsilon$. Moreover, we normalize $\int \hat{\rho}_\epsilon\, d\mu_\epsilon = 1$ and then π_0^ϵ is given by $\pi_0^\epsilon(\varphi) = \hat{\rho}_\epsilon \int \varphi\, d\mu_\epsilon$. Note that, using the definition of $\mathcal{L}_\epsilon$ and Fubini's theorem,

$$\int \Big(\int \varphi(y,j) P_n^\epsilon((x,k), d(y,j)) \Big) (\psi \hat{\rho}_\epsilon)(x,k)\, d\mu_\epsilon(x,k) \\ = \int \varphi(y,j) \mathcal{L}_\epsilon^n (\psi \hat{\rho}_\epsilon)(y,j)\, d\mu_\epsilon(y,j)\,, \tag{5.2}$$

for all $n \geq 1$, and $\varphi, \psi \in BV(\hat{I})$. Taking $\psi = 1$, this proves that $\hat{m}_\epsilon = \hat{\rho}_\epsilon \mu_\epsilon$ is an invariant probability measure for $\hat{\chi}_\epsilon$. Now, consider the measure m_ϵ on I with density $\rho_\epsilon(x) = \sum_{k=0}^\infty (\hat{\rho}_\epsilon \cdot w_\epsilon)(\pi|_{E_k})^{-1}(x)$. Clearly, m_ϵ is absolutely continuous and, as it lifts to $\hat{m}_\epsilon$, it is the unique χ^ϵ-invariant probability measure. Moreover, the same kind of computations as in the proof of Corollary 3 (with (5.2) replacing Lemma 8) prove that the correlations of $(\chi^\epsilon, m_\epsilon)$ decay exponentially fast, with rate at most $\tau_\epsilon = \kappa_1^\epsilon$. ¿From the Perturbation Lemma (1), we get $\limsup_{\epsilon \to 0} \tau_\epsilon \leq \max(\sqrt{\tau_0}, \sqrt{1/\sigma}) < 1$, which proves the second statement in the theorem.

Finally, by Perturbation Lemma (2),

$$\int_{\hat{I}} \Big| \hat{\rho}_0 \frac{\mu_0(\hat{I})}{\mu_\epsilon(\hat{I})} - \hat{\rho}_\epsilon \Big|\, d\mu_0 = \int_{\hat{I}} \Big| (\pi_0^\epsilon - \pi_0)\Big(\frac{1}{\mu_\epsilon(\hat{I})}\Big) \Big|\, d\mu_0 \leq c(\epsilon).$$

Together with $|\mu_0(\hat{I})/\mu_\epsilon(\hat{I}) - 1| \leq c(\epsilon)$, which is a consequence of Lemma 7, this gives that $\int_{\hat{I}} |\hat{\rho}_\epsilon - \hat{\rho}_0|\, d\mu_0 \leq c(\epsilon)$. We claim that $\sup_{\hat{I}} \hat{\rho}_\epsilon$ is bounded uniformly in ϵ. Indeed,

$$\int_{\hat{I}} \hat{\rho}_\epsilon\, d\mu_0 \leq \int_{\hat{I}} \hat{\rho}_0\, d\mu_0 + \int_{\hat{I}} |\hat{\rho}_\epsilon - \hat{\rho}_0|\, d\mu_0 \leq 1 + c(\epsilon) \leq C\,.$$

Then, the Variation Lemma and the Supremum Lemma yield, for large n and $\epsilon < \epsilon(n)$, $\sup_{\hat{I}} \hat{\rho}_\epsilon = \sup_{\hat{I}} \mathcal{L}_\epsilon^n \hat{\rho}_\epsilon \leq \frac{1}{3}(\mathrm{var}_{\hat{I}}\, \hat{\rho}_\epsilon + \sup_{\hat{I}} \hat{\rho}_\epsilon) + C$, and $\mathrm{var}_{\hat{I}}\, \hat{\rho}_\epsilon \leq \frac{1}{3}(\mathrm{var}_{\hat{I}}\, \hat{\rho}_\epsilon + \sup_{\hat{I}} \hat{\rho}_\epsilon) + C$, hence $\sup_{\hat I} \hat\rho_\epsilon \le 3C$. Finally, using Lemma 7 once more,

$$\int_I |\rho_0(x) - \rho_\epsilon(x)|\, dx = \int_{\hat I} |\hat\rho_\epsilon w_\epsilon - \hat\rho_0 w_0|\, dx \le c(\epsilon) + \sup_{\hat I} \hat\rho_\epsilon \int_{\hat I} |w_0 - w_\epsilon|\, dx \le c(\epsilon)\,,$$

which completes our proof. □

References

[BaY] V. Baladi and L.-S. Young, *On the spectra of randomly perturbed expanding maps*, (see also Erratum, to appear Comm. Math. Phys.), Comm. Math. Phys. **156** (1993), 355–385.

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

[BY1] M. Benedicks and L.-S. Young, *Absolutely continuous invariant measures and random perturbations for certain one-dimensional maps*, Ergodic Theory Dynamical Systems **12** (1992), 13–37.

[BY2] M. Benedicks and L.-S. Young, *SBR-measures for certain Hénon maps*, Invent. Math. **112** (1993), 541-576.

[BL] A.M. Blokh and M.Yu. Lyubich, *Measurable dynamics of S-unimodal maps of the interval*, Ann. Scient. École Norm. Sup. **24** (1991), 545–573.

[Co] P. Collet, *Ergodic properties of some unimodal mappings of the interval*, Preprint **11** Institute Mittag Leffler (unpublished) (1984).

[Ja] M. Jakobson, *Absolutely continuous invariant measures for one-parameter families of one-dimensional maps*, Comm. Math. Phys. **81** (1981), 39–88.

[KK] A. Katok and Y. Kifer, *Random perturbations of transformations of an interval*, J. Analyse Math. **47** (1986), 193–237.

[KN] G. Keller and T. Nowicki, *Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps*, Comm. Math. Phys. **149** (1992), 31-69.

[Ki1] Y. Kifer, *Ergodic Theory of Random Perturbations*, Birkhäuser, Boston Basel, 1986.

[Ki2] Y. Kifer, *Random Perturbations of Dynamical Systems*, Birkhäuser, Boston Basel, 1988.

[LY] A. Lasota and J.A. Yorke, *On the existence of invariant measures for piecewise monotonic transformations*, Trans. Amer. Math. Soc. **186** (1973), 481–488.

[Le] F. Ledrappier, *Some properties of absolutely continuous invariant measures on an interval*, Ergodic Theory Dynamical Systems **1** (1981), 77–93.

[MS] W. de Melo and S. van Strien, *One-dimensional Dynamics*, Springer-Verlag, Berlin, 1993.

[No] T. Nowicki, *A positive Lyapunov exponent for the critical value of an S-unimodal mapping implies uniform hyperbolicity*, Ergodic Theory Dynamical Systems **8** (1988), 425-435.

[Ry] M. Rychlik, *Bounded variation and invariant measures*, Studia Math. **LXXVI** (1983), 69–80.

[Si] D. Singer, *Stable orbits and bifurcations of maps of the interval*, SIAM J. Appl. Math. **35** (1978), 260–267.

[Yo] L.-S. Young, *Decay of correlations for certain quadratic maps*, Comm. Math. Phys. **146** (1992), 123–138.

V. Baladi: ETH Zurich, CH-8092 Zurich, Switzerland
(on leave from CNRS, UMR 128, ENS Lyon, France)
E-mail address: baladi@math.ethz.ch

M. Viana: IMPA, Est. D. Castorina 110, 22460-320 Rio de Janeiro, Brazil
E-mail address: viana@impa.br