Ergod. Th. & Dynam. Sys. (2011), **31**, 1345–1361 © *Cambridge University Press, 2010*
doi:10.1017/S0143385710000337

Approximating invariant densities of metastable systems

CECILIA GONZÁLEZ-TOKMAN†, BRIAN R. HUNT‡ and PAUL WRIGHT†

† *Department of Mathematics, University of Maryland, College Park, MD 20742, USA*
(e-mail: cecilia@math.umd.edu, paulrite@math.umd.edu)

‡ *Department of Mathematics and Institute for Physical Sciences and Technology, University of Maryland, College Park, MD 20742, USA*
(e-mail: bhunt@umd.edu)

(*Received* 1 *May* 2009 *and accepted in revised form* 21 *February* 2010)

Abstract. We consider a piecewise smooth expanding map on an interval which has two invariant subsets of positive Lebesgue measure and exactly two ergodic absolutely continuous invariant probability measures (ACIMs). When this system is perturbed slightly to make the invariant sets merge, we describe how the unique ACIM of the perturbed map can be approximated by a convex combination of the two initial ergodic ACIMs. The result is generalized to the case of finitely many invariant components.

1. *Introduction*

Metastable systems are studied in relation to phenomena ranging from molecular [**MDHS06**] to oceanic [**FPET07**] dynamics. Typical trajectories of such a system remain in one of its almost invariant (metastable or quasi-stationary) components for a relatively long period of time, but eventually switch to a different component and repeat this behavior. Quantitative aspects of these phenomena have been studied through eigenvalue and eigenvector approximation techniques for Markov models [**FP08, MSF05**]. Here, we are concerned with rigorous approximation results for eigenvectors—in particular those that correspond to stationary measures of the dynamics—in a more general (non-Markov) setting.

Broadly, this paper deals with the approximation of absolutely continuous invariant probability measures (ACIMs) for certain hyperbolic maps with metastable states. These systems arise from perturbing an initial system T_0 that has two disjoint invariant sets I_l and I_r of positive Lebesgue measure. The initial map has two mutually singular ergodic ACIMs, μ_l and μ_r. When T_0 is perturbed in such a way that I_l and I_r lose their invariance and the perturbed map T_ϵ has only one ACIM μ_ϵ, we are interested in approximating μ_ϵ with μ_l and μ_r. Specifically, the systems we consider are piecewise C^2 expanding maps on an interval; see Figure 1.

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

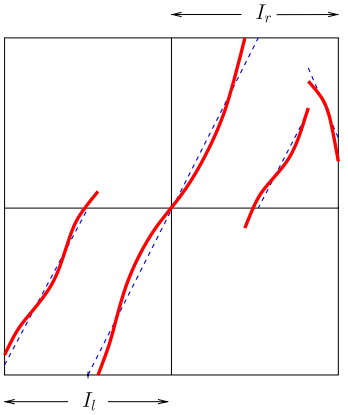


FIGURE 1. A piecewise expanding map with two invariant sets I_l and I_r, together with its perturbed version: dashed curves represent the initial system and solid curves the metastable perturbed system.

Our results can be understood in the context of dynamical systems with holes as follows. As the invariance of the two initially invariant sets is destroyed by the perturbation, we think of the small set $I_l \cap T_\epsilon^{-1} I_r$ of points that switch from I_l to I_r and, likewise, the set $I_r \cap T_\epsilon^{-1} I_l$ as being holes in the initially invariant sets. From this point of view, we expect to be able to approximate μ_ϵ, for small ϵ, by a convex combination $\alpha\mu_l + (1-\alpha)\mu_r$ of the two initially invariant measures, with the ratio $\alpha/(1-\alpha)$ depending on the relative sizes of the holes.

Before discussing our results, we present two illustrative examples. We begin with a simple random system. Consider the family of Markov chains in two states l and r, with transition matrices

$$Q_\epsilon = \begin{pmatrix} 1-\epsilon_{l\to r} & \epsilon_{l\to r} \\ \epsilon_{r\to l} & 1-\epsilon_{r\to l} \end{pmatrix}$$

where $\epsilon = (\epsilon_{l\to r}, \epsilon_{r\to l})$. We are interested in the behavior when $\epsilon \approx 0$. When $\epsilon = 0$, the two sets $I_l = \{l\}$ and $I_r = \{r\}$ are invariant, giving rise to the two ergodic stationary probability measures $\mu_l = \delta_l$ and $\mu_r = \delta_r$. When $\epsilon_{l\to r} > 0$, there is a unique stationary probability measure

$$\mu_\epsilon = \alpha_\epsilon \mu_l + (1-\alpha_\epsilon)\mu_r \quad \text{where } \frac{\alpha_\epsilon}{1-\alpha_\epsilon} = \frac{\epsilon_{r\to l}}{\epsilon_{l\to r}}.$$

Observe that the ratio $\alpha_\epsilon/(1-\alpha_\epsilon)$ of the weights, i.e. $\mu_\epsilon(I_l)/\mu_\epsilon(I_r)$, is equal to the inverse ratio of the sizes of the holes, $\epsilon_{r\to l}/\epsilon_{l\to r}$.

For our second example, consider two billiard tables $\mathcal{D}_l$ and $\mathcal{D}_r$ in the plane, as indicated in Figure 2. For $* \in \{l, r\}$, let $T_* : I_* \circlearrowleft$ be the corresponding billiard map, i.e. the Poincaré map for the first return of the billiard flow to $\partial\mathcal{D}_*$. We use $|\partial\mathcal{D}_*|$ to denote the perimeter of $\mathcal{D}_*$. A general reference for hyperbolic billiards is [**CM06**], where one can find the background for the assertions below. We use the usual coordinates (s, φ) on I_*, where s is arc length on $\partial\mathcal{D}_*$ and $\varphi \in [-\pi/2, +\pi/2]$ is the angle between the outgoing velocity vector and the inward-pointing normal vector to $\partial\mathcal{D}_*$. Then it is well known that T_* leaves (normalized) Liouville measure μ_* invariant, where μ_* has the density

$$\phi_* := d\mu_*/ds\, d\varphi = [2|\partial\mathcal{D}_*|]^{-1} \cos\varphi.$$

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

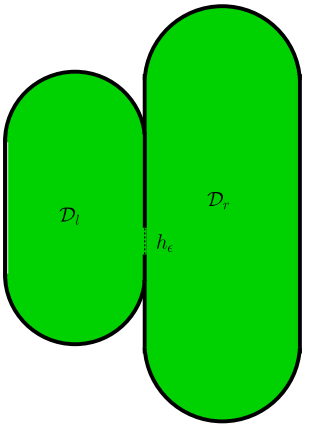


FIGURE 2. Two ergodic billiard tables connected by a hole.

Next, for $\epsilon > 0$, let h_ϵ be a subsegment of $\partial\mathcal{D}_l \cap \partial\mathcal{D}_r$ of length ϵ, and let $\mathcal{D}_\epsilon$ be the billiard table that results after h_ϵ is removed. The corresponding density for the invariant Liouville measure of the billiard map is

$$\phi_\epsilon = [2(|\partial\mathcal{D}_l| + |\partial\mathcal{D}_r| - 2\epsilon)]^{-1} \cos\varphi.$$

Thus, as $\epsilon \to 0$,

$$\phi_\epsilon \to \alpha\phi_l + (1-\alpha)\phi_r \quad \text{where} \frac{\alpha}{1-\alpha} = \frac{|\partial\mathcal{D}_l|}{|\partial\mathcal{D}_r|},$$

provided that some care is taken to define all of the density functions involved on the same space. Note that if we define the holes

$$H_{*,\epsilon} := T_*^{-1}(h_\epsilon \times [-\pi/2, +\pi/2]),$$

then we can rewrite $\alpha/(1-\alpha)$ as $\mu_r(H_{r,\epsilon})/\mu_l(H_{l,\epsilon})$, so that again the ratio of the weights equals the inverse ratio of the sizes of the holes. This example is most meaningful when T_l, T_r and T_ϵ are all ergodic, which is the case for the tables in Figure 2; see [**CM06**, §8.15].

In our main result, Theorem 1, we extend the principle behind the examples above to the deterministic setting of piecewise C^2 expanding maps, under the fairly general conditions described in §2. We show that as $\epsilon \to 0$, the invariant density ϕ_ϵ of T_ϵ converges in L^1 to a convex combination of the ergodic invariant densities of T_0, with the ratio of the weights being given by the limiting inverse ratio of the sizes of the holes. We emphasize that our results do not require any of the piecewise expanding maps involved to have a Markov partition.

The density ϕ_ϵ corresponds to an eigenvector associated with the eigenvalue 1 of the Perron–Frobenius operator acting on a suitable space of functions. Our assumptions imply that for $\epsilon > 0$, the operator has 1 as a simple eigenvalue, and that it also has another real simple eigenvalue slightly less than 1. In Theorem 2, we characterize the eigenvectors associated with this smaller eigenvalue by showing that asymptotically they lie on the line spanned by $d\mu_l/dx - d\mu_r/dx$.

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

Unlike the two examples above, in the setting of piecewise C^2 expanding maps we have no explicit formulas for the invariant densities; even their existence is non-trivial. Our methods rely on the fact that the densities of the ACIMs for T_ϵ are of bounded variation [**LY73**]. Hence, they can be decomposed into regular and singular (or saltus) parts, as in [**Bal07**]. The key technical portions of our proofs involve estimating and exploiting the locations and sizes of the jumps at the discontinuities of the invariant densities, which occur on the forward trajectories of the critical points of T_ϵ.

This work is related to other recent work involving metastable systems and piecewise expanding maps. Recently, Keller and Liverani studied in [**KL09**] metastable systems arising from piecewise smooth uniformly expanding maps with two invariant intervals. They perturbed such an initial map by a family of Markov operators close to the identity to produce a family of metastable systems for $\epsilon > 0$. The associated Perron–Frobenius operators acting on a suitable space of functions have 1 as a simple eigenvalue and another simple eigenvalue $\rho_\epsilon < 1$. As $\epsilon \to 0$, $\rho_\epsilon \to 1$, and the authors rigorously computed the derivative $\lim_{\epsilon \to 0^+}(1 - \rho_\epsilon)/\epsilon$. This provides information on the stationary exchange rate between the metastable states. Their work may be used to prove a corresponding result in our setting.

Our work is also related to current and ongoing investigations on linear response. These problems have the feature that, as Ruelle [**Rue98**] puts it, while it is possible to formulate conjectures based on intuition or formal calculations, the proofs often involve overcoming intricate technicalities. In our setting, we know that $\mu_\epsilon(I_l) \to \alpha$ as $\epsilon \to 0$. A pertinent open problem would be to characterize the higher-order terms $\mathcal{R}(\epsilon) := \mu_\epsilon(I_l) - \alpha$. We do not expect $\mathcal{R}(\epsilon)$ to be differentiable at $\epsilon = 0$ in general. As shown in [**BS08**], linear response fails precisely when the perturbations T_ϵ are transverse to the topological class of T_0, at least for certain piecewise expanding unimodal maps T_0 that are topologically mixing. In that setting, results of [**Kel82**] show that the unique ACIM ϕ_ϵ of T_ϵ satisfies $|\phi_\epsilon - \phi_0|_{L^1} = O(\epsilon \log \epsilon)$, where ϕ_0 is the unique ACIM of T_0. Baladi [**Bal07**] gives examples for which this estimate is optimal.

Theorem 1 can also be regarded as a statement of stochastic stability in the sense that as the size of the perturbation goes to zero, the invariant measures that describe the statistics of Lebesgue-almost every orbit have a computable limit. In the uniformly hyperbolic setting, this phenomenon has been studied in the ergodic case; see, for example, [**Via97**]. In recent years, various studies have been done on stochastic stability outside the setting of uniformly hyperbolic systems. For example, in [**AT05**] the authors work at the boundary of expanding maps, in [**AAV07**] the context of non-uniformly hyperbolic diffeomorphisms is considered, and in [**Vás07**] diffeomorphisms with dominated splitting are treated.

Another interesting problem for further research is to extend our results to higher-dimensional piecewise hyperbolic maps. While we use techniques specific to one-dimensional maps, we are optimistic that the main elements of our proof, found in §3.2, can be generalized.

2. *Statement of results*

In this section, we define a class of dynamical systems that have two nearly invariant (metastable) subsets. These dynamical systems are perturbations of a one-dimensional piecewise smooth expanding map with exactly two invariant subintervals I_l and I_r of

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

positive Lebesgue measure. On each of the two intervals, the unperturbed system has a unique ACIM. The perturbations break this invariance by introducing what we consider to be holes in the intervals; a hole in I_l maps to I_r, and vice versa. Each perturbed system will have only one ACIM, and we will determine an asymptotic formula for its density in terms of the invariant densities of the unperturbed system.

Let $I = [0, 1]$. In this paper, a map $T : I \circlearrowleft$ is called a piecewise C^2 map with $\mathcal{C} = \{0 = c_0 < c_1 < \cdots < c_d = 1\}$ as a critical set if for each i, $T|_{(c_i, c_{i+1})}$ extends to a C^2 function on a neighborhood of $[c_i, c_{i+1}]$. We say that T is uniformly expanding if its minimum expansion, $\inf_{x \in I \setminus \mathcal{C}_0} |T_0'(x)|$, is greater than 1. As is customary for piecewise smooth maps, we consider T to be bi-valued at points $c_i \in \mathcal{C}$ where it is discontinuous. In such cases we let $T(c_i)$ be both values obtained as x approaches c_i from either side, and we write $T(c_{i\pm})$ for the corresponding right and left limits. If $a, b \in \mathcal{C}$, $T|_{[a,b]}$ will be used to denote specifically the restriction of T with $T|_{[a,b]}(a) = T(a_+)$ and $T|_{[a,b]}(b) = T(b_-)$.

We use Leb to denote normalized Lebesgue measure on I and L^1 to denote the space of Lebesgue-integrable functions on I, with norm given by $|f|_{L^1} = \int_I |f(x)|\, dx$. Also, for $f : I \to \mathbb{C}$, we denote by $|f|_\infty$ the supremum of f over I and by $\mathrm{var}(f)$ the total variation of f over I, that is,

$$\mathrm{var}(f) = \sup\left\{\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| : n \geq 1,\ 0 \leq x_0 < x_1 < \cdots < x_n \leq 1\right\}.$$

For clarity of presentation, we do not state our results under the broadest possible assumptions. However, see §2.4 for a number of relaxations of the hypotheses below.

2.1. *The initial system and its perturbations.* We assume that the unperturbed system is a piecewise C^2 uniformly expanding map $T_0 : I \circlearrowleft$ with

$$\mathcal{C}_0 = \{0 = c_{0,0} < c_{1,0} < \cdots < c_{d,0} = 1\}$$

as a critical set. There is a boundary point $b \in (0, 1)$ such that $I_l := [0, b]$ and $I_r := [b, 1]$ are invariant under T_0, i.e. for $* \in \{l, r\}$ we have $T_0|_{I_*}(I_*) \subset I_*$. The existence of an ACIM of bounded variation for $T_0|_{I_*}$ is guaranteed by [**LY73**]. In addition, we assume the properties (I1)–(I4) below.

(I1) *Unique ACIM on the initially invariant set:*
$T_0|_{I_*}$ has only one ACIM μ_*, whose density is denoted by $\phi_* := d\mu_*/dx$.

The uniqueness of such an ACIM is guaranteed by transitivity or by the additional conditions described in [**LY78**]. From (I1), it follows that all ACIMs of T_0 are convex combinations of the ergodic ones, namely μ_l and μ_r.

We define the points in $H_0 := T_0^{-1}\{b\} \setminus \{b\}$ to be *infinitesimal holes*. These are all points that map to the boundary point b, except possibly b itself. Our reasons for excluding b from the set of infinitesimal holes will be explained in §2.4. An immediate consequence of this definition is that $H_0 \subset \mathcal{C}_0$.

(I2) *No return of the critical set to the infinitesimal holes:*
for every $k > 0$, $(T_0^k \mathcal{C}_0) \cap H_0 = \emptyset$.

This is a non-degeneracy condition that may be difficult to check for specific systems. However, one can show that any piecewise C^2 expanding map of the interval can be

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

approximated by maps satisfying (I2), by making arbitrarily small C^2 perturbations. Such perturbations can be constructed inductively, first adding to T_0 an arbitrarily small C^2 perturbation to obtain $T_{0,1}$ such that $(T_{0,1}\mathcal{C}_0) \cap H_0 = \emptyset$. Successive perturbations should be made so that after the kth perturbation the resulting map $T_{0,k}$ satisfies $(T_{0,k}^k \mathcal{C}_0) \cap H_0 = \emptyset$. Furthermore, each perturbation can be made small enough compared to previous perturbations to guarantee that the sum of the perturbations converges in C^2 and that, for each j, the distance between $T_{0,k}^j \mathcal{C}_0$ and H_0 does not decay to zero as $k \to \infty$.

In general, functions of bounded variation are only defined modulo a countable set. However, as we shall see in §4.2, condition (I2) implies that ϕ_* can be defined so that it is continuous at each of the infinitesimal holes in I_*. Thus it is meaningful to discuss the values of ϕ_* at such points.

(I3) *Positive ACIMs at infinitesimal holes:*
ϕ_l is positive at each of the points in $H_0 \cap I_l$, and ϕ_r is positive at each of the points in $H_0 \cap I_r$.

For example, this will be the case if $T_0|_{I_l}$ and $T_0|_{I_r}$ are weakly covering†; see [**Liv95**].

(I4) *Restriction on periodic critical points:*
one of the following holds:
(I4a) $\inf_{x \in I \setminus \mathcal{C}_0} |T_0'(x)| > 2$;
(I4b) T_0 has no periodic critical points, except possibly a fixed point at 0 or 1.

Because T_0 may be bi-valued at points in $\mathcal{C}_0$, a critical point $c_{i,0}$ is considered periodic if there exists $n > 0$ such that $c_{i,0} \in T_0^n(c_{i,0})$. Condition (I4) is necessary in order to ensure that the perturbed systems defined below satisfy uniform Lasota–Yorke estimates. Since we cannot exclude the possibility of the forward orbit of a critical point containing other critical points, these uniform estimates do not follow directly from the original paper [**LY73**], but rather from later works; see §4.2.

In what follows, we consider C^2-small perturbations $T_\epsilon : I \circlearrowleft$ of T_0 for $\epsilon > 0$. This means that a critical set for T_ϵ may be chosen to be

$$\mathcal{C}_\epsilon = \{0 = c_{0,\epsilon} < c_{1,\epsilon} < \cdots < c_{d,\epsilon} = 1\}$$

where, for each i, $\epsilon \mapsto c_{i,\epsilon}$ is a C^2 function for $\epsilon \geq 0$. Furthermore, there exists $\delta > 0$ such that for all sufficiently small ϵ there is a C^2 extension $\hat{T}_{i,\epsilon} : [c_{i,0} - \delta, c_{i+1,0} + \delta] \to \mathbb{R}$ of $T_\epsilon|_{[c_{i,\epsilon}, c_{i+1,\epsilon}]}$, with $\hat{T}_{i,\epsilon} \to \hat{T}_{i,0}$ in the C^2 topology. We also assume the following properties.

(P1) *Unique ACIM:*
for each $\epsilon > 0$, T_ϵ has only one ACIM μ_ϵ, with density $\phi_\epsilon := d\mu_\epsilon/dx$.

(P2) *Boundary condition:*
the boundary point does not move, and no holes are created near the boundary. To be precise:
(P2a) if $b \notin \mathcal{C}_0$, then necessarily $T_0(b) = b$, and we assume further that for all $\epsilon > 0$, $T_\epsilon(b) = b$;
(P2b) if $b \in \mathcal{C}_0$, we assume that $T_0(b_-) < b < T_0(b_+)$ and also that $b \in \mathcal{C}_\epsilon$ for all ϵ.

If the boundary point does move under the perturbation, condition (P2) can often still be satisfied upon performing a smooth change of coordinates close to the identity; see §2.4.

† A piecewise expanding map $T : I \circlearrowleft$ with $\mathcal{C} = \{0 = c_0 < c_1 < \cdots < c_d = 1\}$ as a critical set is weakly covering if there is some N such that for every i, $\bigcup_{k=0}^N T^k([c_i, c_{i+1}]) = I$.

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

2.2. *Main results.* The central question of this study is the following: for small ϵ, how can we asymptotically approximate μ_ϵ by a convex combination of μ_l and μ_r? To that end, let $H_{l,\epsilon} := I_l \cap T_\epsilon^{-1}(I_r)$ and $H_{r,\epsilon} := I_r \cap T_\epsilon^{-1}(I_l)$. We refer to these sets as *holes*. Once a T_ϵ-orbit enters a hole, it leaves one of the invariant sets for T_0 and continues in the other. As $\epsilon \to 0$, the holes converge (in the Hausdorff metric) to the infinitesimal holes from which they arise.

Condition (P1) ensures that for $\epsilon > 0$, at least one of the holes has positive Lebesgue measure. In view of (I3), without loss of generality we can suppose that $\mu_l(H_{l,\epsilon}) > 0$ and define

$$l.h.r. = \lim_{\epsilon \to 0} \frac{\mu_r(H_{r,\epsilon})}{\mu_l(H_{l,\epsilon})}$$

whenever the limit exists. (Here *l.h.r.* stands for *limiting hole ratio*.)

THEOREM 1. (Approximation of the invariant density) *Consider the family of perturbations T_ϵ of T_0 under the assumptions stated in §2.1. Suppose that l.h.r. (as defined above) exists. Then, as $\epsilon \to 0$,*

$$\phi_\epsilon \xrightarrow{L^1} \alpha\phi_l + (1-\alpha)\phi_r \quad \textit{where } \frac{\alpha}{1-\alpha} = l.h.r.$$

We allow for the possibility that $l.h.r. = +\infty$, in which case $\alpha = 1$. Several straightforward generalizations of the above result are discussed in §2.4.

Remark 1. The limit *l.h.r.* will always exist as long as the perturbations open up holes $H_{l,\epsilon}$ whose size is truly of first order in ϵ. For simplicity, suppose that there are only two infinitesimal holes, $h_l \in I_l$ and $h_r \in I_r$. Then we can always write

$$H_{*,\epsilon} = (h_* - a_*\epsilon + o(\epsilon),\ h_* + b_*\epsilon + o(\epsilon)) \quad \text{for } * \in \{l, r\},$$

and if $a_l + b_l > 0$, then

$$l.h.r. = \frac{\phi_r(h_r)(a_r + b_r)}{\phi_l(h_l)(a_l + b_l)}.$$

For example, this will be the case if $T_\epsilon = T_0 + \epsilon g + o(\epsilon)$ for some smooth function g with $g(h_l) > 0$.

The case where the limit *l.h.r.* does not exist is addressed in §2.4.

Remark 2. An alternative definition of *l.h.r.* is as a limit of a ratio of escape rates. For $* \in \{l, r\}$, we can consider a dynamical system with a hole, where orbits stop upon entering the hole, by taking the unperturbed map $T_0|_{I_*}$ with $H_{*,\epsilon}$ as the hole. See [**DY06**] for an exposition of such systems. Let $R_{*,\epsilon}$ be the exponential escape rate of Lebesgue measure and suppose that there is only one infinitesimal hole in each initially invariant interval. Then, as $\epsilon \to 0$, $\mu_*(H_{*,\epsilon})/R_{*,\epsilon} \to 1$ (see [**BY08, KL09**]).

Next, let $\mathcal{L}_\epsilon$ be the Perron–Frobenius operator associated with T_ϵ acting on the Banach space $BV = \{f : I \to \mathbb{C} : \mathrm{var}(f) < \infty\}$ with the variation norm†, and let $\sigma(\mathcal{L}_\epsilon)$ denote the spectrum of $\mathcal{L}_\epsilon$. It follows from [**Kel89**, Theorem 8.3(b)], for instance, that $\mathcal{L}_0$ has 1

† In fact, we work in the quotient space obtained by identifying two functions of bounded variation if they differ on at most a countable set. As confusion is unlikely to arise, we use the same notation for a function and its equivalence class.

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

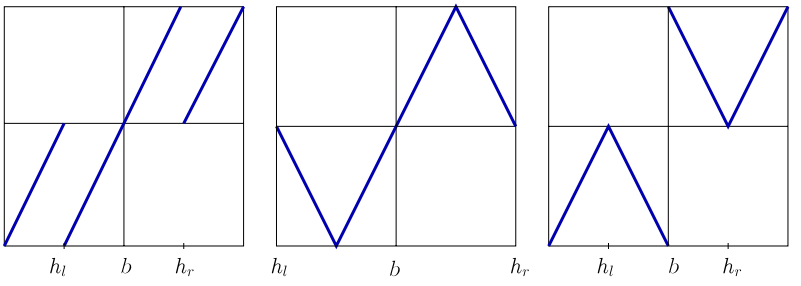


FIGURE 3. Piecewise linear maps giving rise to metastable systems.

as an isolated eigenvalue of multiplicity two. Furthermore, in [**KL99**] it was shown that for fixed small $\delta > 0$ and every $\epsilon > 0$ small enough, $\sigma(\mathcal{L}_\epsilon) \cap B_\delta(1)$ consists of exactly two eigenvalues, 1 and $\rho_\epsilon < 1$, each of multiplicity one. As $\epsilon \to 0$, $\rho_\epsilon \to 1$ and the total spectral projection of $\mathcal{L}_\epsilon$ associated with $\sigma(\mathcal{L}_\epsilon) \cap B_\delta(1)$ converges (at a given rate in an appropriate norm) to the total spectral projection of $\mathcal{L}_0$ associated with $\sigma(\mathcal{L}_0) \cap B_\delta(1)$. In §4.2 we will verify that a uniform Lasota–Yorke inequality (7) holds in our setting; this is the only assumption of [**KL99**] that is neither trivial nor well known in our context.

THEOREM 2. (Characterization of the eigenspace corresponding to the smaller eigenvalue) *For each $\epsilon > 0$ small enough, there is a unique real-valued function $\psi_\epsilon \in BV$ satisfying $\mathcal{L}_\epsilon \psi_\epsilon = \rho_\epsilon \psi_\epsilon$, $|\psi_\epsilon|_{L^1} = 1$ and $\int_{I_l} \psi_\epsilon \, dx > 0$. As $\epsilon \to 0$,*

$$\psi_\epsilon \xrightarrow{L^1} \tfrac{1}{2}\phi_l - \tfrac{1}{2}\phi_r.$$

Remark 3. Suppose that μ_l and μ_r are both mixing for T_0. Given a typical initial density $f \in BV$ (i.e. one with non-zero coefficient of ψ_ϵ when expressed as a linear combination of eigenvectors), as $\mathcal{L}_\epsilon^n f \to \phi_\epsilon$ the deviation $\mathcal{L}_\epsilon^n f - \phi_\epsilon$ becomes approximately proportional to ψ_ϵ for n large. In this case, Theorem 2 implies, roughly, that for ϵ small, $\mathcal{L}_\epsilon^n f$ becomes close to a linear combination of ϕ_l and ϕ_r more quickly than it comes close to the specific linear combination $\alpha\phi_l + (1-\alpha)\phi_r$.

2.3. *Examples.* The three piecewise linear maps shown in Figure 3 satisfy assumptions (I1)–(I4) from §2.1. In all three cases, normalized Lebesgue measure restricted to the left or right intervals is the unique ACIM of the corresponding restricted system.

By adding a small C^2 perturbation $g : I \times [0, \epsilon_0) \to I$ such that $g(\cdot, 0) \equiv 0$ and, for any $\epsilon \neq 0$, $g(b, \epsilon) = 0$, $g(h_l, \epsilon) > 0$ and $g(h_r, \epsilon) < 0$, we obtain a one-parameter family of perturbations $T_\epsilon := T_0 + g(\cdot, \epsilon)$ satisfying assumptions (P1) and (P2).

If

$$\lim_{\epsilon \to 0} \frac{\text{Leb}(H_{r,\epsilon})}{\text{Leb}(H_{l,\epsilon})} = l.h.r.,$$

then by Theorem 1 the invariant densities ϕ_ϵ associated with T_ϵ satisfy

$$\phi_\epsilon \xrightarrow{L^1} \alpha \, \text{Leb}|_{I_l} + (1-\alpha)\text{Leb}|_{I_r} \quad \text{where } \frac{\alpha}{1-\alpha} = l.h.r.$$

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

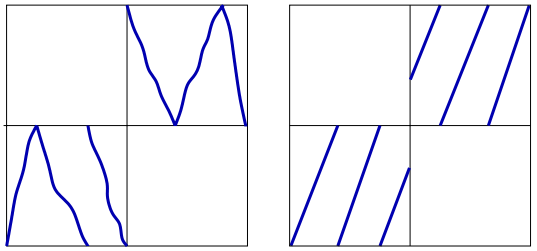


FIGURE 4. Examples of initial maps T_0 that give rise to metastable systems for which our results hold.

The possibility $l.h.r. = \infty$ is allowed, and in this case we have

$$\phi_\epsilon \xrightarrow{L^1} \text{Leb}|_{I_l}.$$

Two other initial maps T_0 to which Theorems 1 and 2 are applicable are shown in Figure 4.

2.4. *Generalizations.* Our results can be extended, with essentially the same proofs, to yield the following straightforward generalizations.

Multiple invariant sets. We can allow T_0 to have $m \geq 2$ invariant sets $I_1, \ldots, I_m$ of positive Lebesgue measure, provided it has a unique ACIM $\phi_i = d\mu_i/dx$ on each I_i. Each invariant set can be an interval or a union of intervals; see Figure 5. For simplicity, we limit ourselves to the case where, for $\epsilon > 0$, all of the transitions between the initially invariant sets are of first order in ϵ; that is, for $i \neq j$, $\mu_i(I_i \cap T_\epsilon^{-1} I_j)$ is either identically 0 or equals $\epsilon \cdot \beta_{i,j} + o(\epsilon)$ for some $\beta_{i,j} > 0$. In this case, under assumptions that are straightforward generalizations of (I1)–(I4), (P1) and (P2), the unique invariant density ϕ_ϵ of T_ϵ converges as $\epsilon \to 0$ to a convex combination of the ϕ_i. The coefficients can be determined from the corresponding coefficients of the stationary measure for the finite-state Markov chain whose transition matrix has the off-diagonal entries $\epsilon \cdot \beta_{i,j}$. (One can easily check that the stationary measure for this Markov chain is independent of ϵ for all small $\epsilon > 0$, and also that our assumptions imply that the transition matrix is irreducible and hence has a unique stationary measure.)

If $m > 2$, the analogue of Theorem 2 says only that the eigenfunctions for T_ϵ whose eigenvalues approach 1 but are distinct from 1 limit on the space of linear combinations of invariant densities for T_0 with integral 0.

Relaxed boundary condition. The condition that the boundary point not move when T_0 is perturbed is inessential; relaxing this restriction simply means that the metastable states for T_ϵ will be slight perturbations of the initial invariant sets. In this case, a smooth change of coordinates restores the hypothesis (P2). For example, when $b \notin \mathcal{C}_0$, assumption (P2a) is actually superfluous, although the definitions in the statement of Theorem 1 must be modified slightly. As remarked earlier, we necessarily have $T_0(b) = b$. Furthermore, the graph of T_0 intersects the diagonal transversely at this point. Thus,

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

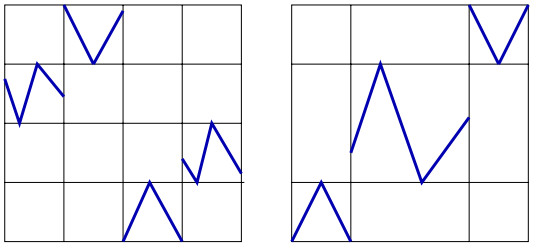


FIGURE 5. Initial maps T_0 that give rise to metastable systems to which our results can be generalized. The initially invariant sets are $I_1 = [0, 1/4] \cup [1/2, 3/4]$ and $I_2 = [1/4, 1/2] \cup [3/4, 1]$ in the left diagram and $I_1 = [0, 1/4]$, $I_2 = [1/4, 3/4]$ and $I_3 = [3/4, 1]$ in the right diagram.

for all small $\epsilon > 0$, there is a unique point b_ϵ near b satisfying $T_\epsilon(b_\epsilon) = b_\epsilon$. Then the quasi-invariant sets for T_ϵ are $I_{l,\epsilon} := [0, b_\epsilon]$ and $I_{r,\epsilon} := [b_\epsilon, 1]$, and the corresponding holes are defined by $H_{l,\epsilon} := I_{l,\epsilon} \cap T_\epsilon^{-1}(I_{r,\epsilon})$ and $H_{r,\epsilon} := I_{r,\epsilon} \cap T_\epsilon^{-1}(I_{l,\epsilon})$. Aside from these minor modifications, the statements and proofs of our main results remain the same.

When $b \in \mathcal{C}_0$, (P2b) can be relaxed by no longer requiring that $b \in \mathcal{C}_\epsilon$ for all ϵ. In this case, when $\epsilon > 0$, $\mathcal{C}_\epsilon$ contains a point b_ϵ that converges to b as $\epsilon \to 0$, and the quasi-invariant sets and holes must be redefined as above. However, it is still essential to assume that no holes are created near the boundary, and we enforce this with the assumption that $T_0(b_-) < b < T_0(b_+)$. For example, if

$$T_\epsilon(x) = [(3x \bmod 1/2) + 3\epsilon] \cdot 1_{x<1/2} + [(-3x \bmod 1/2) + 1/2 - \epsilon] \cdot 1_{x>1/2},$$

then all of our assumptions aside from (P2) hold, with $b = 1/2$, $\mu_* = \text{Leb}|_{I_*}$ and $l.h.r. = 1/3$. However, as $\epsilon \to 0$, $\phi_\epsilon \xrightarrow{L^1} \phi_r$. The difficulty is that orbits ejected from I_r by T_ϵ immediately return to I_r.

Multiple limiting densities. When the limit $l.h.r.$ in §2.2 does not exist, we let

$$\underline{l.h.r.} = \liminf_{\epsilon \to 0} \frac{\mu_r(H_{r,\epsilon})}{\mu_l(H_{l,\epsilon})}, \qquad \overline{l.h.r.} = \limsup_{\epsilon \to 0} \frac{\mu_r(H_{r,\epsilon})}{\mu_l(H_{l,\epsilon})}.$$

Since the function $\mu_r(H_{r,\epsilon})/\mu_l(H_{l,\epsilon})$ is continuous in $\epsilon > 0$, our arguments show that the set of limit points for ϕ_ϵ as $\epsilon \to 0$ is precisely

$$\left\{ \tilde{\alpha}\phi_l + (1 - \tilde{\alpha})\phi_r : \frac{\tilde{\alpha}}{1 - \tilde{\alpha}} \in [\, \underline{l.h.r.}\,,\, \overline{l.h.r.}\,] \right\}.$$

3. *Proofs of the main theorems*

In this section, we state the main properties of the invariant densities of the initial system and its perturbations. Then, we present the proofs of Theorems 1 and 2. For notational convenience, we will assume that there are only two infinitesimal holes, $h_l \in I_l$ and $h_r \in I_r$; the proof without this restriction is essentially the same.

3.1. *Properties of the invariant densities.*

Here we record some of the relevant characteristics of the density functions ϕ_l, ϕ_r and ϕ_ϵ. First, if $f \in BV$, we can (and will)

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

choose a representative of f with only regular discontinuities; that is, for each x, $f(x) = (\lim_{y\to x^-} f(y) + \lim_{y\to x^+} f(y))/2$. Then, following [**Bal07**], we can uniquely decompose f into the sum of a regular and a singular (or saltus) part: $f = f^{\text{reg}} + f^{\text{sal}}$. Here f^{reg} is continuous with $\text{var}(f^{\text{reg}}) \le \text{var}(f)$, and f^{sal} is the sum of at most countably many step functions. We write $f^{\text{sal}} = \sum_{u\in\mathcal{S}} s_u H_u$, where $\mathcal{S}$ is the discontinuity set of f, s_u is the *jump* of f at u, and $H_u(x)$ is -1 if $x < u$, $-1/2$ if $x = u$ and 0 if $x > u$. This representation imposes the boundary condition $f^{\text{sal}}(1) = 0$. Furthermore,

$$\text{var}(f^{\text{sal}}) = \sum_{u\in\mathcal{S}} |s_u| \le \text{var}(f).$$

PROPOSITION 1. (Key facts about the invariant densities) *There exists $\epsilon_0 > 0$ such that the following hold.*

(i) *Uniform bound on the variations of the invariant densities:*

$$\sup_{0<\epsilon<\epsilon_0} \text{var}(\phi_\epsilon) < +\infty.$$

Also, $\text{var}(\phi_l), \text{var}(\phi_r) < +\infty$.

(ii) *Uniform bound on the Lipschitz constant of the regular parts:*
for $0 < \epsilon < \epsilon_0$, each of the $\phi_\epsilon^{\text{reg}}$ is Lipschitz continuous with constant $\text{Lip}(\phi_\epsilon^{\text{reg}})$*, and*

$$\sup_{0<\epsilon<\epsilon_0} \text{Lip}(\phi_\epsilon^{\text{reg}}) < +\infty.$$

Also, ϕ_l^{reg} and ϕ_r^{reg} are Lipschitz.

(iii) *Approximate continuity near the infinitesimal holes:*
for $* \in \{l, r\}$ and each $\eta > 0$, there exists $\delta > 0$ such that for all $0 < \epsilon < \epsilon_0$,

$$\text{var}_{[h_*-\delta,h_*+\delta]}(\phi_\epsilon^{\text{sal}}) := \textit{the variation of } \phi_\epsilon^{\text{sal}} \textit{ over } [h_* - \delta, h_* + \delta] < \eta.$$

Also, ϕ_* is continuous at h_*.

The proof of Proposition 1 is technical, so we defer it until §4.2.

3.2. *Proofs.* We recall that for any $C_1, C_2 > 0$,

$$\{f \in BV : |f|_{L^1} \le C_1,\ \text{var}(f) \le C_2\}$$

is pre-compact in L^1. This fact will be used repeatedly in what follows.

Proof of Theorem 1. By (i) of Proposition 1, we can choose a sequence of values ϵ' converging to 0 such that $\phi_{\epsilon'}$ converges in L^1 to some function, which we denote by ϕ_0. Using the fact that ϕ_ϵ is a fixed point of the Perron–Frobenius operator $\mathcal{L}_\epsilon$ associated with T_ϵ (see §4.1 for the definition), one can verify that ϕ_0 is an invariant density for T_0, and so there exists α such that $\phi_0 = \alpha\phi_l + (1-\alpha)\phi_r$. We will show that necessarily $\alpha/(1-\alpha) = l.h.r.$. This implies that there is exactly one limit point of ϕ_ϵ as $\epsilon \to 0$, and Theorem 1 follows.

Now, for $\epsilon' > 0$, $\mu_{\epsilon'}(H_{l,\epsilon'}) = \mu_{\epsilon'}(H_{r,\epsilon'})$ because $I_l = H_{l,\epsilon'} \cup (I_l \backslash H_{l,\epsilon'})$ and $T_{\epsilon'}^{-1}(I_l) = H_{r,\epsilon'} \cup (I_l \backslash H_{l,\epsilon'})$ are disjoint unions (modulo sets of zero Lebesgue measure),

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

and $\phi_{\epsilon'} = d\mu_{\epsilon'}/dx$ is an invariant density for $T_{\epsilon'}$. We will show that as $\epsilon' \to 0$,

$$\mu_{\epsilon'}(H_{l,\epsilon'}) = \alpha\mu_l(H_{l,\epsilon'}) + o(1)\cdot\mu_l(H_{l,\epsilon'}), \tag{1}$$

$$\mu_{\epsilon'}(H_{r,\epsilon'}) = (1-\alpha)\mu_r(H_{r,\epsilon'}) + o(1)\cdot\mu_r(H_{r,\epsilon'}), \tag{2}$$

from which the equation $\alpha/(1-\alpha) = l.h.r.$, and hence Theorem 1, follows immediately.

We prove only equation (1), since the proof of (2) is analogous. Write

$$\begin{aligned}\mu_{\epsilon'}(H_{l,\epsilon'}) &= \int_{H_{l,\epsilon'}} \phi_{\epsilon'}\,dx = \alpha\int_{H_{l,\epsilon'}} \phi_l\,dx + \int_{H_{l,\epsilon'}} (\phi_{\epsilon'} - \alpha\phi_l)\,dx \\ &= \alpha\mu_l(H_{l,\epsilon'}) + O\left(\sup_{x\in H_{l,\epsilon'}} |\phi_{\epsilon'}(x) - \alpha\phi_l(x)|\right)\cdot \mathrm{Leb}(H_{l,\epsilon'}).\end{aligned}$$

But as $\epsilon' \to 0$, $H_{l,\epsilon'} \to h_l$ in the Hausdorff metric, and then

$$\mu_l(H_{l,\epsilon'})/\mathrm{Leb}(H_{l,\epsilon'}) \to \phi_l(h_l) > 0$$

because ϕ_l is continuous at h_l. Thus our proof is completed by the following lemma.

LEMMA 1. *As $\epsilon' \to 0$,*

$$\sup_{x\in H_{l,\epsilon'}} |\phi_{\epsilon'}(x) - \alpha\phi_l(x)| \to 0.$$

Although this uniform convergence might at first seem surprising, properties (ii) and (iii) of Proposition 1 essentially say that near h_l, $\{\phi_{\epsilon'}\}$ behaves like a family of equicontinuous functions.

Proof. We proceed by contradiction. Suppose that there exists $C > 0$ and a subsequence $\epsilon'' \to 0$ of the ϵ' values such that for each ϵ'', there is a point $x_{\epsilon''} \in H_{l,\epsilon''}$ with $|\phi_{\epsilon''}(x_{\epsilon''}) - \alpha\phi_l(x_{\epsilon''})| > C$. Necessarily, $x_{\epsilon''} \to h_l$ as $\epsilon'' \to 0$.

We restrict all functions of interest to the left subinterval I_l. Set $\gamma_{\epsilon''} := \phi^{\mathrm{reg}}_{\epsilon''} - \alpha\phi^{\mathrm{reg}}_l$ and $\omega_{\epsilon''} := \phi^{\mathrm{sal}}_{\epsilon''} - \alpha\phi^{\mathrm{sal}}_l$, so that $\phi_{\epsilon''} - \alpha\phi_l = \gamma_{\epsilon''} + \omega_{\epsilon''}$. Using (ii) of Proposition 1, let L be such that for all sufficiently small ϵ'', $\mathrm{Lip}(\gamma_{\epsilon''}) < L$. Next, we use Proposition 1(iii) with $\eta = C/5$ and make sure to choose the corresponding $\delta < C/(5L)$ small enough so that $\mathrm{var}_{[h_l-\delta,h_l+\delta]}(\alpha\phi^{\mathrm{sal}}_l) < C/5$ as well. Thus $\mathrm{var}_{[h_l-\delta,h_l+\delta]}(\omega_{\epsilon''}) < 2C/5$. Then, if $x \in [h_l-\delta, h_l+\delta]$ and ϵ'' is sufficiently small, $x_{\epsilon''} \in [h_l-\delta, h_l+\delta]$ and

$$\begin{aligned}|\gamma_{\epsilon''}(x) + \omega_{\epsilon''}(x)| &\geq |\gamma_{\epsilon''}(x_{\epsilon''}) + \omega_{\epsilon''}(x_{\epsilon''})| \\ &\quad - |\gamma_{\epsilon''}(x) + \omega_{\epsilon''}(x) - \gamma_{\epsilon''}(x_{\epsilon''}) - \omega_{\epsilon''}(x_{\epsilon''})| \\ &\geq C - [L\cdot 2\delta + 2C/5] \geq C/5.\end{aligned}$$

But this contradicts the fact that

$$\gamma_{\epsilon''} + \omega_{\epsilon''} = \phi_{\epsilon''} - \alpha\phi_l \xrightarrow{L^1} 0. \qquad \square$$

Proof of Theorem 2. First, we observe that the results of [**KL99**] guarantee that for small $\epsilon > 0$, $\rho_\epsilon < 1$ is a simple eigenvalue of multiplicity one. Hence there are exactly two real-valued eigenfunctions, $\pm\psi_\epsilon$, satisfying $\mathcal{L}_\epsilon\psi_\epsilon = \rho_\epsilon\psi_\epsilon$ and $|\psi_\epsilon|_{L^1} = 1$. But for such functions,

$$\int \psi_\epsilon\,dx = \int \mathcal{L}_\epsilon\psi_\epsilon\,dx = \rho_\epsilon\int \psi_\epsilon\,dx,$$

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

and so $\int \psi_\epsilon \, dx = 0$. We have the following uniform bound on their variations, whose proof we defer until §4.2.

LEMMA 2. (Uniform bound on the variations of the ψ_ϵ) *There exists* $\epsilon_1 > 0$ *such that* $\sup_{0<\epsilon<\epsilon_1} \operatorname{var}(\psi_\epsilon) < +\infty$.

Let ψ_0 be any limit point in L^1 of ψ_ϵ as $\epsilon \to 0$. Then, since $\rho_\epsilon \to 1$, it follows that ψ_0 is invariant under $\mathcal{L}_0$ and is therefore a linear combination of ϕ_l and ϕ_r. Since $|\psi_0|_{L^1} = 1$ and $\int \psi_0 \, dx = 0$, we necessarily have $\psi_0 = \pm\phi_l/2 \mp \phi_r/2$. Hence we can uniquely specify ψ_ϵ by the condition $\int_{I_l} \psi_\epsilon \, dx > 0$, and Theorem 2 follows.

4. *Proofs of the properties of the densities*

In order to prepare for the proofs of Proposition 1 and Lemma 2, it will be convenient to first show how to derive such properties for an invariant density of a single, fixed piecewise expanding map. We do this in §4.1. Then, in §4.2, we prove Proposition 1 and Lemma 2 by showing how such estimates can be made uniformly for the family of maps T_ϵ, $\epsilon \geq 0$.

Before beginning, we remark that if $f \in BV$, then for each x,

$$|f|_\infty \leq |f(x)| + \operatorname{var}(|f|) \leq |f(x)| + \operatorname{var}(f).$$

Upon integrating, we find that $|f|_\infty \leq |f|_{L^1} + \operatorname{var}(f)$. We will use this fact repeatedly in what follows.

4.1. *Properties of an invariant density for a single piecewise expanding map.*

Let $T : I \circlearrowleft$ be a piecewise C^2 uniformly expanding map, with $\mathcal{C} = \{0 = c_0 < c_1 < \cdots < c_d = 1\}$ as a critical set. Let $\mathcal{L}$ be the associated Perron–Frobenius operator, i.e. the transfer operator acting on densities. We begin by briefly reviewing a method for finding an invariant density of T. Such a method was introduced in [**LY73**]; see [**Bal00**, Ch. 3] for a more modern exposition. Let $\lambda_T = \inf_{x \in I \setminus \mathcal{C}} |T'(x)| > 1$ be the minimum expansion and $D_T = \sup_{x \notin \mathcal{C}} |T''(x)|/|T'(x)|$ the distortion of T. Then, if $f \in BV$ and $x \notin T\mathcal{C}$,

$$\mathcal{L}f(x) = \sum_{i=1}^{d} f(\xi_i(x))|\xi_i'(x)|1_{J_i}(x) \tag{3}$$

where

$$J_i = T|_{[c_{i-1}, c_i]}([c_{i-1}, c_i]) \quad \text{and} \quad \xi_i = (T|_{[c_{i-1}, c_i]})^{-1} : J_i \to [c_{i-1}, c_i].$$

One can show that there exist constants $\beta \in (0, 1)$ and C_{LY} such that for each $n \geq 1$ and any $f \in BV$, the following Lasota–Yorke inequality holds:

$$\operatorname{var}(\mathcal{L}^n f) \leq C_{\mathrm{LY}}\beta^n \operatorname{var}(f) + C_{\mathrm{LY}}|f|_{L^1}. \tag{4}$$

In fact, β can be chosen to be any number greater than λ_T^{-1}, although we will not use this fact. Set $F_n = (1/n)\sum_{k=0}^{n-1} \mathcal{L}^k 1$. Then $F_n \xrightarrow{L^1} \phi$, where $\phi \in BV$ is the density of an ACIM for T. Using Helly's theorem, one obtains that $\operatorname{var}(\phi) \leq C_{\mathrm{LY}}$.

We wish to characterize the properties of the regular and singular terms in the decomposition $\phi = \phi^{\mathrm{reg}} + \phi^{\mathrm{sal}}$. First, let us define a hierarchy on the set of points in the postcritical orbits $\mathcal{S} = \bigcup_{k \geq 1} T^k\mathcal{C}$ by $\#(u) := \inf\{k \geq 1 : u \in T^k\mathcal{C}\}$. The following

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

characterization is motivated by the discussion of the invariant densities for unimodal expanding maps found in [**Bal07, BS08**]. In particular, in [**BS08**, §3.3] a norm is introduced on the sequence of jumps of ϕ along the postcritical orbit with weights that grow exponentially in $\#(u)$.

LEMMA 3. *Given the hypotheses above, the following properties hold.*

(a) *The term ϕ^{reg} is Lipschitz continuous. Furthermore, there exists a constant $C_{\mathrm{dis}} = C_{\mathrm{dis}}(\lambda_T, D_T)$ such that*

$$\mathrm{Lip}(\phi^{\mathrm{reg}}) \le C_{\mathrm{dis}}(1 + C_{\mathrm{LY}}).$$

This constant $C_{\mathrm{dis}}(\lambda_T, D_T)$ can be defined so that it depends continuously on $\lambda_T > 1$ and $D_T \ge 0$.

(b) *The discontinuity set of ϕ is a subset of $\mathcal{S} = \bigcup_{k\ge 1} T^k\mathcal{C}$. If we write $\phi^{\mathrm{sal}} = \sum_{u\in\mathcal{S}} s_u H_u$, then for each $m \ge 0$,*

$$\sum_{\{u\in\mathcal{S}:\#(u)>m\}} |s_u| \le \lambda_T^{-m}\, C_{\mathrm{LY}}.$$

Proof. We begin by noting from equation (3) that F_n is smooth except possibly at points in $\bigcup_1^{n-1} T^k\mathcal{C}$. Write $F_n^{\mathrm{sal}} = \sum_{u\in\mathcal{S}} s_{u,n} H_u$. Then we can show that $|s_{u,n}|$ decays uniformly exponentially fast in $\#(u)$; that is, we make the following claim.

SUBLEMMA. *For any $m, n \ge 0$,*

$$\sum_{\{u\in\mathcal{S}:\,\#(u)>m\}} |s_{u,n}| \le \lambda_T^{-m}\, C_{\mathrm{LY}}.$$

Proof of the sublemma. If $m \ge n$, then

$$\sum_{\#(u)>m} |s_{u,n}| = 0 \quad \text{and} \quad \sum_{\#(u)>0} |s_{u,n}| = \mathrm{var}(F_n^{\mathrm{sal}}) \le \mathrm{var}(F_n) \le C_{\mathrm{LY}}.$$

Since

$$F_n = \frac{n-1}{n}\mathcal{L}F_{n-1} + \frac{1}{n},$$

if $\#(u) > 1$ we see from equation (3) with $f = F_{n-1}$ that

$$|s_{u,n}| \le \frac{n-1}{n}\lambda_T^{-1} \sum_{\{v\in\mathcal{S}:Tv=u\}} |s_{v,n-1}|.$$

Thus, if $0 < m < n$, then

$$\begin{aligned}
\sum_{\#(u)>m} |s_{u,n}| &\le \sum_{\#(u)>m} \frac{n-1}{n}\lambda_T^{-1} \sum_{\{v\in\mathcal{S}:Tv=u\}} |s_{v,n-1}| \\
&\le \frac{n-1}{n}\lambda_T^{-1} \sum_{\#(u)>m-1} |s_{u,n-1}| \\
&\le \cdots \\
&\le \frac{n-m}{n}\lambda_T^{-m} \sum_{\#(u)>0} |s_{u,n-m}| \le \lambda_T^{-m} C_{\mathrm{LY}}.
\end{aligned}$$

For the inequalities above, we have used the fact that if $\#(u) > 1$, then $T^{-1}(u)$ does not contain any critical points. □

Using a diagonalization argument, we can find a subsequence n_j such that for each u, s_{u,n_j} converges as $n_j \to \infty$ to some number, which we write as $\hat{s}_u$. In

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

particular, for each m we have $\sum_{\#(u)>m} |\hat{s}_u| \le \lambda_T^{-m} C_{\mathrm{LY}}$, and $F_{n_j}^{\mathrm{sal}} \xrightarrow{L^1} F^{\mathrm{sal}}$ where $F^{\mathrm{sal}} := \sum_{u\in\mathcal{S}} \hat{s}_u H_u$. Furthermore, a standard distortion estimate (see, for example, [**Bal07**, proof of Proposition 3.3]) gives that there exists a constant C_{dis} such that for each $n \ge 0$,

$$\mathrm{Lip}((\mathcal{L}^n 1)^{\mathrm{reg}}) \le C_{\mathrm{dis}} |\mathcal{L}^n 1|_\infty \le C_{\mathrm{dis}}(1 + \mathrm{var}(\mathcal{L}^n 1)).$$

Here, C_{dis} depends only on the minimum expansion and the distortion of T. In particular, $\sup_{n\ge1} \mathrm{Lip}(F_n^{\mathrm{reg}}) \le C_{\mathrm{dis}}(1 + C_{\mathrm{LY}})$. By the Arzelà–Ascoli theorem, we can find a continuous function F^{reg} such that some subsequence of $\{F_{n_j}^{\mathrm{reg}}\}$ converges in L^∞ to F^{reg}.

By the uniqueness of the decomposition $\phi = \phi^{\mathrm{reg}} + \phi^{\mathrm{sal}}$, we conclude that $\phi^{\mathrm{reg}} = F^{\mathrm{reg}}$ and $\phi^{\mathrm{sal}} = F^{\mathrm{sal}}$. Lemma 3 follows. □

4.2. *Proofs of Proposition 1 and Lemma 2.* We prove only the claims about ϕ_ϵ for $\epsilon > 0$, and leave the claims about ϕ_l and ϕ_r to the reader.

Let $\mathcal{L}_\epsilon$ be the Perron–Frobenius operator (3) associated with T_ϵ. The first key step is to prove that the $\mathcal{L}_\epsilon$ with ϵ sufficiently small satisfy Lasota–Yorke inequalities with uniform constants. Let λ_ϵ and D_ϵ be the minimum expansion and distortion of T_ϵ, respectively. Then, as $\epsilon \to 0$, $\lambda_\epsilon \to \lambda_0$ and $D_\epsilon \to D_0$. Furthermore, T_ϵ is a piecewise C^2 uniformly expanding map that is a small C^2 perturbation of T_0, and the two critical sets $\mathcal{C}_\epsilon$ and $\mathcal{C}_0$ are ϵ-close to each other. This is not sufficient to guarantee uniform Lasota–Yorke inequalities; see, for example, [**Kel82**, §6] or [**Bla92**]. However, such uniform inequalities do follow if we make the additional assumption (I4), which guarantees that either (a) we have $\lambda_0 > 2$ or (b) T_0 has no periodic critical points, except possibly the points in ∂I as fixed points. We assume case (a) in our presentation here, and comment on case (b) at the end of this section.

Fix $\lambda \in (2, \lambda_0)$. The original proof from [**LY73**] shows that if $f \in BV$ is real-valued, then

$$\mathrm{var}(\mathcal{L}_\epsilon f) \le (2\lambda_\epsilon^{-1})\mathrm{var}(f) + C_\epsilon |f|_{L^1},$$

where

$$C_\epsilon = D_\epsilon/\lambda_\epsilon + 2 \max_i |c_{i+1,\epsilon} - c_{i,\epsilon}|^{-1}. \tag{5}$$

(See also [**Liv95**, §2].) Upon iterating, we find that for sufficiently small ϵ, all such f and $n \ge 1$,

$$\mathrm{var}(\mathcal{L}_\epsilon^n f) \le \beta^n \mathrm{var}(f) + C_{\mathrm{LY}} |f|_{L^1} \tag{6}$$

with $\beta = 2\lambda^{-1}$ and $C_{\mathrm{LY}} = 2C_0/(1 - 2\lambda^{-1})$. Similar estimates can be made for complex-valued f by applying (6) to the real and imaginary parts separately. Since each T_ϵ has a unique ACIM, we know from our discussion in §4.1 that for sufficiently small $\epsilon > 0$, $(1/n)\sum_{k=0}^{n-1} \mathcal{L}_\epsilon^k 1 \xrightarrow{L^1} \phi_\epsilon$ as $n \to \infty$. It follows from Lemma 3 that $\mathrm{var}(\phi_\epsilon)$ and $\mathrm{Lip}(\phi_\epsilon^{\mathrm{reg}})$ are uniformly bounded.

Next, we prove (iii). Given $\eta > 0$, choose n large enough that $\lambda^{-n} C_{\mathrm{LY}} < \eta$. Using (I2), we can choose $\delta > 0$ so small that for $0 < k \le n$, $(T_0^k \mathcal{C}_0) \cap [h_* - 2\delta, h_* + 2\delta] = \emptyset$. It follows that for ϵ sufficiently small, $(T_\epsilon^k \mathcal{C}_\epsilon) \cap [h_* - \delta, h_* + \delta] = \emptyset$ as well. Using part (b) of Lemma 3 with $m = n$, we then see that $\mathrm{var}_{[h_*-\delta, h_*+\delta]}(\phi_\epsilon^{\mathrm{sal}}) < \eta$.

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

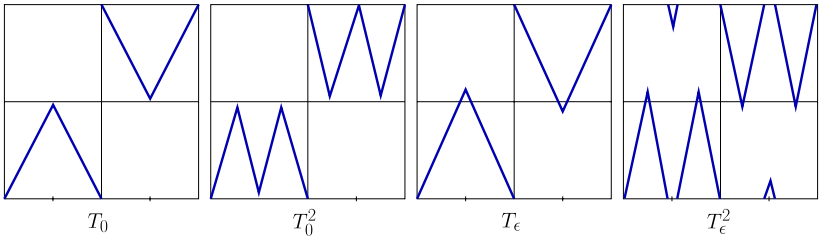


FIGURE 6. Creation of small intervals of differentiability.

Finally, to prove Lemma 2, we use equation (6) with $f = \psi_\epsilon$, n chosen so large that $\beta^n < 1/2$, and ϵ chosen so small that $\rho_\epsilon^n > 3/4$. It follows that $\mathrm{var}(\psi_\epsilon) \le C_{\mathrm{LY}}/(\rho_\epsilon^n - \beta^n) \le 4C_{\mathrm{LY}}$.

Modifications when the minimum expansion is not bigger than two. If, in assumption (I4), the minimum expansion of T_0 is $\lambda_0 \le 2$, one derives Lasota–Yorke estimates for $\mathcal{L}_0$ by first fixing N large enough so that $\lambda_0^N > 2$. Then the arguments from [**LY73**] used above will yield a Lasota–Yorke estimate for $\mathcal{L}_0^N$, and this can be interpolated to give similar estimates for $\mathcal{L}_0$. One can try to obtain uniform estimates for $\mathcal{L}_\epsilon$, but the arguments used above will work only if the critical points for T_ϵ^N are in a one-to-one correspondence with and very close to those of T_0^N (cf. equation (5)), as would be the case if $\mathcal{C}_0 \cap (\bigcup_{k=1}^{N-1} T^k \mathcal{C}_0) = \emptyset$.

Unfortunately, this will never be the case in our setting, at least when $b \in \mathcal{C}_0$. This is because the infinitesimal holes in H_0 are necessarily critical points, and they are mapped to b by T_0. Because at least some of the infinitesimal holes must be mapped across the boundary point when $\epsilon > 0$, this means that necessarily T_ϵ^2 will have more critical points than T_0^2, and these additional critical points will create very short intervals on which T_ϵ^2 is smooth; see Figure 6. However, this problem can be dealt with by using assumption (I4b). Specifically, in [**BY93**] it is shown that because of the restriction on the periodic critical points, the growth in the number of very short intervals on which T_ϵ^n is smooth as n increases can be controlled, and uniform Lasota–Yorke estimates can still be made. More precisely, there exists $\epsilon_0 > 0$ and constants $\beta \in (0, 1)$ and C_{LY} such that for any $\epsilon \in [0, \epsilon_0]$, $n \ge 1$ and $f \in BV$,

$$\mathrm{var}(\mathcal{L}_\epsilon^n f) \le C_{\mathrm{LY}} \beta^n \,\mathrm{var}(f) + C_{\mathrm{LY}} |f|_{L^1}. \tag{7}$$

The proof of this is essentially identical to [**Bal00**, proof of Lemma 3.2] (see also [**Bal00**, Remark 3.4] and [**BY93**, proof of Lemma 8]), so we omit it. The rest of the proofs of Proposition 1 and Lemma 2 proceed as above.

Acknowledgements. The authors would especially like to acknowledge D. Dolgopyat for suggesting this project and for helpful conversations. They also thank M. Demers for helpful conversations and the referees for thoughtful comments. The first author's research was supported by CONACyT, México. The third author was partially supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship.

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press

References

[AAV07] J. F. Alves, V. Araújo and C. H. Vásquez. Stochastic stability of non-uniformly hyperbolic diffeomorphisms. *Stoch. Dyn.* **7**(3) (2007), 299–333.

[AT05] V. Araújo and A. Tahzibi. Stochastic stability at the boundary of expanding maps. *Nonlinearity* **18**(3) (2005), 939–958.

[Bal00] V. Baladi. *Positive Transfer Operators and Decay of Correlations (Advanced Series in Nonlinear Dynamics, 16)*. World Scientific, River Edge, NJ, 2000.

[Bal07] V. Baladi. On the susceptibility function of piecewise expanding interval maps. *Comm. Math. Phys.* **275**(3) (2007), 839–859.

[Bla92] M. L. Blank. Chaotic mappings and stochastic Markov chains. *Mathematical Physics, X (Leipzig, 1991)*. Springer, Berlin, 1992, pp. 341–345.

[BS08] V. Baladi and D. Smania. Linear response formula for piecewise expanding unimodal maps. *Nonlinearity* **21**(4) (2008), 677–711.

[BY93] V. Baladi and L.-S. Young. On the spectra of randomly perturbed expanding maps. *Comm. Math. Phys.* **156**(2) (1993), 355–385.

[BY08] L. Bunimovich and A. Yurchenko. Where to place a hole to achieve a fastest escape rate. *Preprint*, 2008, arXiv:0811.4438 [math.DS], *Israel J. Math.* at press.

[CM06] N. Chernov and R. Markarian. *Chaotic Billiards (Mathematical Surveys and Monographs, 127)*. American Mathematical Society, Providence, RI, 2006.

[DY06] M. Demers and L. S. Young. Escape rates and conditionally invariant measures. *Nonlinearity* **19**(2) (2006), 377–397.

[FP08] G. Froyland and K. Padberg. Almost-invariant sets and invariant manifolds—connecting probabilistic and geometric descriptions of coherent structures in flows. *Phys. D* **238**(16) (2009), 1507–1523.

[FPET07] G. Froyland, K. Padberg, M. H. England and A. M. Treguier. Detection of coherent oceanic structures via transfer operators. *Phys. Rev. Lett.* **98** (2007), 224503.

[Kel82] G. Keller. Stochastic stability in some chaotic dynamical systems. *Monatsh. Math.* **94**(4) (1982), 313–333.

[Kel89] G. Keller. Markov extensions, zeta functions, and Fredholm theory for piecewise invertible dynamical systems. *Trans. Amer. Math. Soc.* **314**(2) (1989), 433–497.

[KL99] G. Keller and C. Liverani. Stability of the spectrum for transfer operators. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (4) **28**(1) (1999), 141–152.

[KL09] G. Keller and C. Liverani. Rare events, escape rates and quasistationarity: some exact formulae. *J. Stat. Phys.* **135**(3) (2009), 519–534.

[Liv95] C. Liverani. Decay of correlations for piecewise expanding maps. *J. Stat. Phys.* **78**(3–4) (1995), 1111–1129.

[LY73] A. Lasota and J. A. Yorke. On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.* **186** (1973), 481–488.

[LY78] T. Y. Li and J. A. Yorke. Ergodic transformations from an interval into itself. *Trans. Amer. Math. Soc.* **235** (1978), 183–192.

[MDHS06] E. Meerbach, E. Dittmer, I. Horenko and C. Schütte. Multiscale modelling in molecular dynamics: biomolecular conformations as metastable states. *Computer Simulations in Condensed Matter: From Materials to Chemical Biology.* Vol. I *(Lecture Notes in Physics, 703)*. Springer, Berlin, 2006, pp. 475–497.

[MSF05] E. Meerbach, C. Schütte and A. Fischer. Eigenvalue bounds on restrictions of reversible nearly uncoupled Markov chains. *Linear Algebra Appl.* **398** (2005), 141–160.

[Rue98] D. Ruelle. General linear response formula in statistical mechanics, and the fluctuation–dissipation theorem far from equilibrium. *Phys. Lett. A* **245**(3–4) (1998), 220–224.

[Vás07] C. H. Vásquez. Statistical stability for diffeomorphisms with dominated splitting. *Ergod. Th. & Dynam. Sys.* **27**(1) (2007), 253–283.

[Via97] M. Viana. Stochastic dynamics of deterministic systems. *Lecture Notes*, Brazilian Mathematics Colloquium, IMPA, Rio de Janeiro, 1997.

https://doi.org/10.1017/S0143385710000337 Published online by Cambridge University Press