J Stat Phys (2009) 135: 519–534
DOI 10.1007/s10955-009-9747-8

Rare Events, Escape Rates and Quasistationarity: Some Exact Formulae

Gerhard Keller · Carlangelo Liverani

Received: 13 October 2008 / Accepted: 16 April 2009 / Published online: 29 April 2009

Abstract We present a common framework to study decay and exchanges rates in a wide class of dynamical systems. Several applications, ranging from the metric theory of continuos fractions and the Shannon capacity of constrained systems to the decay rate of metastable states, are given.

Keywords Rare event · Decay rate · Metastability · Quasistationarity · Eigenvalue · Perturbation

1 Introduction

In applications of the theory of dynamical systems to concrete situations it is often necessary to study rare events. Examples are open systems with a small chance to escape and metastable states. Although much numerical work exists (e.g. see [6, 10] and references therein) not many rigorous results are available. In principle one can try to apply perturbation theory but the existing theorems [13, 14] do not produce very sharp results. A similar situation occurs in the study of linear response theory. While perturbation theory applies to a wide class of smooth systems [22], this is no longer true when discontinuities are present in

It is a pleasure to thank L. Bunimovich for calling our attention to formula (1.1) in the special case of the doubling map, and V. Baladi and M. Demers for discussions on this formula. We are indebted to the ESI where, during the Workshop on Hyperbolic Dynamical Systems with Singularities (2008), this work was started. L.C. thanks the ENS, Paris, where he was invited during part of this work. Also we like to thank the Institut Henri Poincare–Centre Emile Borel where, during the trimester Mécanique statistique, probabilités et systèmes de particules (2008), this work was finished. Finally, G.K. acknowledges the support by the DFG grant KE 514/7-1, and L.C. by the MIUR grant PRIN 2007B3RBEY.

G. Keller
Department Mathematik, Universität Erlangen-Nürnberg, Bismarckstr. 1 1/2, 91052 Erlangen, Germany
e-mail: keller@mi.uni-erlangen.de

C. Liverani (✉)
Dipartimento di Matematica, II Università di Roma (Tor Vergata), Via della Ricerca Scientifica, 00133 Roma, Italy
e-mail: liverani@mat.uniroma2.it

the system. In that case not only perturbation theory does not imply linear response, but in fact there are cases when linear response itself is violated, see [1, 3] and references therein. Since an open system is typically modeled by a hole in the system (that is by a region in which the dynamics stops once the trajectory enters it), the presence of discontinuities is inevitable.

Accordingly, one could expect that the quasi-invariant measure (which describes the long time distribution of the trajectories conditioned to the event that they have not left the system, i.e. they have not entered the hole) and the escape rate (that measures the rate at which trajectories leave the system) depend in a very erratic (non-differentiable) way on the size and position of the hole. Yet, almost nothing is known about such situations.

In the present paper we prove a general theorem providing a first order expansion of escape rates and exchange rates in terms of the "size" of the rare effect that is investigated (whereby refining the results in [14, 25] even-though limited to the present setting). We derive from this theorem explicit formulae for escape rates (both in one dimensional and two-dimensional cases) and for the exchange rate between two quasi-invariant sets (metastability). To our knowledge such formulae were known only for rather special cases (see [7] where such type of formulae first appeared) and are completely new in the generality presented here.

Just to give an impression of the wide applicability of our main result, we list a few examples that are detailed in Sect. 3.

Let z be a periodic point of period p for the doubling map $x \mapsto 2x \bmod 1$ and consider intervals $I_\varepsilon \ni z$ of length ε. Denote the decay rate for the hole I_ε by λ_ε; i.e. the Lebesgue measure of the set of points that are not trapped by I_ε during the first n iterations of the map decreases asymptotically like λ_ε^n. Then λ_ε has the following first order expansion at $\varepsilon = 0$

$$\lambda_\varepsilon = 1 - \varepsilon \cdot \left(1 - 2^{-p}\right) + o(\varepsilon). \tag{1.1}$$

Similar formulae can be obtained when the hole is a union of several intervals, and analogous results hold of course for a coin tossing process that is stopped once a pattern of heads and tails from a given finite collection is observed.

We turn to continued fraction expansions. Let $m_{k,n}$ be the Lebesgue measure of those points $x \in [0, 1]$ whose continued fraction expansion up to the n-th digit does not contain a block of k consecutive 1's. For each k, these numbers decrease asymptotically like some λ_k^n and, setting $z = \frac{\sqrt{5}-1}{2}$, we show

$$\lim_{k\to\infty} \frac{1-\lambda_k}{z^{2k}} = \frac{z^3(1+z^2)^2}{\ln 2} \approx 0.6504. \tag{1.2}$$

The above are one dimensional examples. Along the same lines one can treat piecewise expanding maps in higher dimensions provided the invariant density is not too irregular in a neighborhood of the holes. This is not guaranteed working only with the usual multivariate BV-spaces, but rather by variants as developed by Blank [4] or Saussol [24]. In principle the present theory also applies to Anosov diffeomorphisms if one can devise the proper functional space setting. Unfortunately, despite recent progress [2, 9], the available settings are still not adequate for the applications of the present results. Nevertheless, it is conceivable that in the near future this result could be applied e.g. to billiards.

A related but different question occurs if a system has two ergodic mixing components that share a common part of their boundary in phase space. In that case small perturbations can cause rare "jumps" over the boundary giving thus rise to quasistationary (also called metastable or nearly invariant) behavior. As a result the double eigenvalue 1 corresponding to the two original mixing components splits into a single eigenvalue 1 and another real eigenvalue λ_ε close to 1 which characterizes the rate of exchange between the two components, see e.g. [8, 21]. In Sect. 3.2 we consider piecewise expanding 1D-maps T on the interval [0, 1] with two mixing components I_1, I_2 having a fixed point z in common. The Markov process obtained by adding to the dynamics, at each time n, independent identically distributed random noise εZ_n shows quasistationary behavior with

$$\lim_{\varepsilon\to 0}\frac{1-\lambda_\varepsilon}{\varepsilon}=\frac{\beta+\alpha}{2}\left(1-\frac{1}{T'(z)}\right)\mathbb{E}[|W|]+\frac{\beta-\alpha}{2}\mathbb{E}[Z_1] \tag{1.3}$$

where $W:=\sum_{n=0}^{\infty}[T'(z)]^{-n}Z_{n+1}$, $\alpha:=\frac{h(z^-)}{2m(I_1)}$, $\beta:=\frac{h(z^+)}{2m(I_2)}$ and $h=\lim_{n\to\infty}P_0^n 1$ is the "natural" invariant density of T.

The paper is organized as follows: in the next section we describe the general setting and state our main Theorem 2.1 whose proof is postponed to Sect. 6. In Sect. 3 we apply the theorem to decay and exchange rates of piecewise expanding interval maps and illustrate its applicability with some specific examples: the doubling map, the Gauss map and the generalized cusp map. The decay rate of a two-dimensional example is studied in Sect. 4. More precisely, we study the rate at which trajectories of two coupled 1D-maps synchronize up to some difference ε in the limit $\varepsilon\to 0$. Finally, in Sect. 5, we indicate relations between some of our formulas and approaches to metastability in molecular dynamics [20], in oceanic structures [12], and with the Shannon capacity of constrained binary codes [16].

2 An Abstract Perturbation Result

Let $(V,\|.\|)$ be a real or complex normed vector space with dual $(V',\|.\|)$. Consider a family $P_\varepsilon: V\to V$ $(\varepsilon\in E)$ of uniformly bounded linear operators where $E\subseteq\mathbb{R}$ is a closed set of parameters with $\varepsilon=0$ as an accumulation point. We make the following assumptions on the operators P_ε: there are $\lambda_\varepsilon\in\mathbb{C}$, $\varphi_\varepsilon\in V$, $\nu_\varepsilon\in V'$ and linear operators $Q_\varepsilon: V\to V$ such that

$$\lambda_\varepsilon^{-1}P_\varepsilon=\varphi_\varepsilon\otimes\nu_\varepsilon+Q_\varepsilon, \tag{A1}$$

$$P_\varepsilon(\varphi_\varepsilon)=\lambda_\varepsilon\varphi_\varepsilon,\ \nu_\varepsilon P_\varepsilon=\lambda_\varepsilon\nu_\varepsilon,\ \ Q_\varepsilon(\varphi_\varepsilon)=0,\ \nu_\varepsilon Q_\varepsilon=0, \tag{A2}$$

$$\sum_{n=0}^{\infty}\sup_{\varepsilon\in E}\|Q_\varepsilon^n\|=:C_1<\infty. \tag{A3}$$

(The summability condition in (A3) can only be satisfied if the operators P_ε have a uniform spectral gap. See Remark 2.2 for a weakening of this requirement.) Observe that assumptions (A1) and (A2) imply $\nu_\varepsilon(\varphi_\varepsilon)=1$ for all ε. As our ultimate goal is to prove a perturbation result for small ε, it is natural to relate the "size" of φ_ε to that of φ_0 by a further assumption:

$$\nu_0(\varphi_\varepsilon)=1\quad\text{and}\quad\sup_{\varepsilon\in E}\|\varphi_\varepsilon\|=:C_2<\infty. \tag{A4}$$

Finally we denote

$$\Delta_\varepsilon:=\nu_0((P_0-P_\varepsilon)(\varphi_0)) \tag{2.1}$$

and we make the following assumptions to control the size of the perturbation: there is $C_3 > 0$ such that

$$\eta_\varepsilon := \|\nu_0(P_0 - P_\varepsilon)\| \to 0 \text{ as } \varepsilon \to 0, \tag{A5}$$

$$\eta_\varepsilon \cdot \|(P_0 - P_\varepsilon)(\varphi_0)\| \le C_3|\Delta_\varepsilon|. \tag{A6}$$

Here η_ε denotes the norm of the linear functional $\nu_0(P_0 - P_\varepsilon) : V \to \mathbb{R}$.

The basic identity is

$$\lambda_0 - \lambda_\varepsilon = \lambda_0 \nu_0(\varphi_\varepsilon) - \nu_0(\lambda_\varepsilon(\varphi_\varepsilon)) = \nu_0((P_0 - P_\varepsilon)(\varphi_\varepsilon)). \tag{2.2}$$

In view of assumptions (A4) and (A5) this implies

$$|\lambda_0 - \lambda_\varepsilon| \le C_2 \eta_\varepsilon, \tag{2.3}$$

in particular, $\lim_{\varepsilon\to 0} \lambda_\varepsilon = \lambda_0$. The main result of this section is the following more accurate approximation for $\lambda_0 - \lambda_\varepsilon$.

Theorem 2.1 *Assume* (A1)–(A6).

(a) *There is* $\varepsilon_0 > 0$ *such that* $\lambda_\varepsilon = \lambda_0$ *if* $\varepsilon \le \varepsilon_0$ *and* $\Delta_\varepsilon = 0$.
(b) *If* $\Delta_\varepsilon \ne 0$ *for all sufficiently small* $\varepsilon \in E$ *and if*

$$q_k := \lim_{\varepsilon \to 0} q_{k,\varepsilon} := \lim_{\varepsilon\to 0} \frac{\nu_0((P_0 - P_\varepsilon)P_\varepsilon^k(P_0 - P_\varepsilon)(\varphi_0))}{\Delta_\varepsilon} \tag{A7}$$

exists for each integer $k \ge 0$, *then*

$$\lim_{\varepsilon\to 0} \frac{\lambda_0 - \lambda_\varepsilon}{\Delta_\varepsilon} = 1 - \sum_{k=0}^{\infty} \lambda_0^{-(k+1)} q_k. \tag{2.4}$$

Remark 2.2 In Sect. 6 we prove this theorem under slightly weaker hypothesis. In particular, we relax the assumption of uniform hyperbolicity (which yields a spectral gap for the transfer operator). Namely, we replace the summability condition $\sum_{n=0}^{\infty} \sup_{\varepsilon\in E} \|Q_\varepsilon^n\| < \infty$ from (A3) by the following condition: there is a second norm $\|\cdot\|_* \ge \|\cdot\|$ on V such that

$$\sum_{n=0}^{\infty} \sup_{\varepsilon\in E} \|Q_\varepsilon^n\|^* =: C_1 < \infty \tag{A3*}$$

where $\|Q_\varepsilon^n\|^* := \sup\{\|Q_\varepsilon^n \psi\| : \|\psi\|_* \le 1\}$. We have to compensate this by slightly stronger assumptions on φ_0, namely $\|\varphi_0\|_* \le C_2 < \infty$ and

$$\eta_\varepsilon \cdot \|(P_0 - P_\varepsilon)(\varphi_0)\|_* \le C_3|\Delta_\varepsilon|. \tag{A6*}$$

3 Applications to Piecewise Expanding Interval Maps

Assume that $T : [0, 1] \to [0, 1]$ is piecewise monotone with (possibly countably many) continuously differentiable branches. (This means that each branch is continuously differentiable in the interior of its domain so that the derivative even of a single branch may be unbounded.) Define $g : \mathbb{R} \to \mathbb{R}$ by $g(x) = 1/|T'(x)|$ if x is in the interior of one of the monotonicity intervals of T and $g(x) = 0$ otherwise, and assume that $\|g\|_\infty < 1$ and that g is of bounded variation. Let BV be the space of real-valued functions of bounded variation on [0, 1]. Rychlik [23] showed that the Perron-Frobenius operator P of T acting on Lebesgue equivalence classes of functions from BV is quasi-compact. (As BV-functions have at most countably many discontinuities, two BV-functions in the same Lebesgue equivalence class have the same discontinuities and differ at most by their values at these countably many points, and we will not distinguish henceforth between BV-functions and their Lebesgue equivalence classes.)

If T is mixing this implies that $P_0 = P$ satisfies (A1)–(A3) for $\varepsilon = 0$ with $\nu_0 = m$ (= Lebesgue measure), $\lambda_0 = 1$ and $0 \le \varphi_0 \in \mathrm{BV}$.

The essential observation behind this is that for $\varepsilon = 0$ a Lasota-Yorke type inequality [17] is satisfied: there are constants $r \in (0, 1)$ and $R > 0$ such that for $\varepsilon = 0$, all $n \in \mathbb{N}$ and all $f \in \mathrm{BV}$,

$$\|P_\varepsilon^n f\| \le R\left(r^n \|f\| + \int |f| dm\right) \tag{3.1}$$

where $\|f\|$ is the variation of the extension of f to the whole real line by setting $f(x) = 0$ if $x \notin [0, 1]$. For the applications in this section we will assume that this inequality holds not only for $\varepsilon = 0$ but, with uniform constants r and R, for all $\varepsilon \in E$. This is mostly the case when P_ε is a small dynamical perturbation of P_0—however there are exceptions, see [14] for a more precise discussion and references.

3.1 Decay Rates

We suppose that T is mixing. Let $(V, \|.\|)$ be the space BV, let $E = [0, \varepsilon_1]$, and consider a family $(I_\varepsilon)_{\varepsilon \in E}$ of compact subintervals of [0, 1] such that $I_\varepsilon \subseteq I_{\varepsilon'}$ if $\varepsilon \le \varepsilon'$. Define the operators P_ε by $P_\varepsilon(f) = P(f 1_{[0,1]\setminus I_\varepsilon})$. If $m(I_{\varepsilon_1})$ is sufficiently small, the perturbation results from [14] apply provided (3.1) holds, see [19, Sect. 7]. In particular, (A1)–(A4) are satisfied for $\varepsilon \in E$. We have $\Delta_\varepsilon = \nu_0(P(1_{I_\varepsilon \setminus I_0}\varphi_0)) = \mu_0(I_\varepsilon \setminus I_0)$ where μ_0 is the probability measure with density φ_0 w.r.t. ν_0. (μ_0 is indeed the equilibrium state for $\log g$ on the "non-trapped" set $X_{nt} := \{x \in [0, 1] : T^n x \notin \mathring{I}_0 \forall n \ge 0\}$. φ_ε is also the conditionally invariant density for the "hole" I_ε, see e.g. [19].)

We need to check assumptions (A5) and (A6). First note that

$$\eta_\varepsilon = \sup_{\|\psi\|\le 1} |\nu_0(P_0(\psi 1_{I_\varepsilon \setminus I_0}))| = |\lambda_0| \sup_{\|\psi\|\le 1} \left| \int_{I_\varepsilon \setminus I_0} \psi \, d\nu_0 \right| \le |\lambda_0| \nu_0(I_\varepsilon \setminus I_0). \tag{3.2}$$

In particular, $|\nu_0(P_0 - P_\varepsilon)(\varphi_0)| = |\lambda_0| \int_{I_\varepsilon \setminus I_0} \varphi_0 \, d\nu_0$. As $\|(P_0 - P_\varepsilon)(\varphi_0)\| \le \mathcal{O}(\|\varphi_0 1_{I_\varepsilon \setminus I_0}\|)$, assumptions (A5) and (A6) will be satisfied if $\nu_0(I_\varepsilon \setminus I_0) \to 0$ when $\varepsilon \to 0$ and if

$$\|\varphi_0 1_{I_\varepsilon \setminus I_0}\| \le \mathrm{const} \frac{1}{\nu_0(I_\varepsilon \setminus I_0)} \int_{I_\varepsilon \setminus I_0} \varphi_0 \, d\nu_0. \tag{3.3}$$

This condition (as well as conditions (A1)–(A4) discussed above) can be checked easily in most cases of interest. It is always satisfied if $\inf \varphi_0|_{I_{\varepsilon_1}} > 0$.

3.1.1 Holes I_ε Shrinking to a Point

We specialize to the case where $I_0 = \{z\}$ for some $z \in [0,1]$ so that P_0 is indeed the Perron-Frobenius operator P for T and $\lambda_0 = 1$, and we assume for simplicity that T and also the invariant density φ_0 are continuous at z. We consider I_ε with length ε, so $m(I_\varepsilon \setminus I_0) = \varepsilon$, and we assume that $\Delta_\varepsilon > 0$. Here are a few examples:

The doubling map: $T(x) = 2x \bmod 1$ with $\varphi_0(x) = 1$.
The Gauss map: $T(x) = \frac{1}{x} \bmod 1$ with $\varphi_0(x) = \frac{1}{\ln 2}\frac{1}{1+x}$.
The generalized cusp map: $T_\gamma(x) = 1 - |2x-1|^\gamma$ for some $\gamma \in (\frac{1}{2}, 1]$. As $|T_\gamma'(x)| = \frac{2\gamma}{|2x-1|^{1-\gamma}} \geq 2\gamma$, this map is a uniformly expanding map with two full branches. The weight function $g(x) = |T_\gamma'(x)|^{-1}$ has two monotone bounded branches and is clearly of bounded variation. (Observe that T_1 is just the tent map. $T_{1/2}$ is known as the cusp map. It has $x = 0$ as a neutral fixed point and is not covered by the present setting.) The invariant density $\varphi_0(x)$ of T_γ behaves like $\frac{\varphi_0(1/2)}{\gamma}(1-x)^{\frac{1}{\gamma}-1}$ near $x = 1$, so it has a zero at $x = 1$ if $\gamma < 1$.[1]

In all three examples, $\Delta_\varepsilon = \mu_0(I_\varepsilon) > 0$. In the first two examples, condition (3.3) is clearly satisfied because $\inf \varphi_0|_{I_\varepsilon} > 0$ if ε is sufficiently small. For the generalized cusp map the same is true if $z \neq 1$. In case $z = 1$, $\|\varphi_0 1_{I_\varepsilon}\| = 2\,\mathrm{const}_\gamma\, \varepsilon^{\frac{1}{\gamma}-1}$ and $\int_{I_\varepsilon} \varphi_0\, dm = \mathrm{const}_\gamma \int_{1-\varepsilon}^1 (1-x)^{\frac{1}{\gamma}-1}\, dx = \mathrm{const}_\gamma \gamma \varepsilon^{\frac{1}{\gamma}}$ so that (3.3) is satisfied as well.

Let

$$U_{k,\varepsilon} := T^{-1}([0,1] \setminus I_\varepsilon) \cap \cdots \cap T^{-k}([0,1] \setminus I_\varepsilon) \cap T^{-(k+1)} I_\varepsilon.$$

As $q_{k,\varepsilon} = \mu_0(I_\varepsilon \cap U_{k,\varepsilon})/\mu_0(I_\varepsilon)$, we find:

If z is not periodic: then $U_{k,\varepsilon} = \emptyset$ for sufficiently small ε so that $q_k = \lim_{\varepsilon \to 0} q_{k,\varepsilon} = 0$ for all k. Therefore,

$$\lim_{\varepsilon \to 0} \frac{1 - \lambda_\varepsilon}{\mu_0(I_\varepsilon)} = 1, \quad \text{in particular } \lim_{\varepsilon \to 0} \frac{1 - \lambda_\varepsilon}{m(I_\varepsilon)} = \varphi_0(z). \tag{3.4}$$

If z is periodic with period p: then $U_{k,\varepsilon} = \emptyset$ for sufficiently small ε except if $k = p - 1$ so that

$$\lim_{\varepsilon \to 0} \frac{1 - \lambda_\varepsilon}{\mu_0(I_\varepsilon)} = 1 - \lim_{\varepsilon \to 0} \frac{\mu_0(I_\varepsilon \cap T^{-p} I_\varepsilon)}{\mu_0(I_\varepsilon)} = 1 - \frac{1}{|(T^p)'(z)|}, \tag{3.5}$$

$$\text{in particular } \lim_{\varepsilon \to 0} \frac{1 - \lambda_\varepsilon}{m(I_\varepsilon)} = \varphi_0(z) \left(1 - \frac{1}{|(T^p)'(z)|}\right). \tag{3.6}$$

Formulas (3.4) and (3.6) imply that the function $\varepsilon \mapsto \lambda_\varepsilon$ is differentiable at $\varepsilon = 0$. Note, however, that in general it is non-differentiable at other values of ε, see Sect. 3.1.2 below.

We look more explicitly at the above three examples. Recall that $I_0 = \{z\}$.

[1] Here is a sketch of the argument: As the map T_γ has full branches, the invariant density $\varphi_0 = \lim_{n \to \infty} P_0^n 1$ is continuous. Also $\varphi_0(\frac{1}{2}) > 0$, because otherwise $\varphi_0(x) = 0$ for all $x \in \bigcup_n T_\gamma^{-n}\{\frac{1}{2}\}$, and this set is dense in $[0,1]$. Therefore, for x close to 1,

$$\varphi_0(x) = P_0 \varphi_0(x) \sim \frac{\varphi_0(1/2)}{\gamma}(1-x)^{\frac{1}{\gamma}-1}.$$

The doubling map: $\lim_{\varepsilon\to 0}\frac{1-\lambda_\varepsilon}{m(I_\varepsilon)} = 1-2^{-p}$ if $T^p(z)=z$. This is (1.1).

The Gauss map: (a) Consider $z=0$ and $\epsilon\in E:=\{\frac{1}{2},\frac{1}{3},\frac{1}{4},\dots\}\cup\{0\}$. Let $I_\varepsilon:=[0,\varepsilon]$. In terms of the continued fraction algorithm this means that the expansion stops as soon as a digit $\geq \varepsilon^{-1}$ is generated. In this case it is easy to see that

$$\mu_0(I_\varepsilon\cap U_{k,\varepsilon}) \leq \frac{1}{\ln 2} m(I_\varepsilon\cap T^{-(k+1)}I_\varepsilon) = \frac{1}{\ln 2}\int_{I_\varepsilon} P_0^k(P_0 1_{I_\varepsilon})\,dm$$
$$= \mathcal{O}(\varepsilon\mu_0(I_\varepsilon))$$

so that $q_{k,\varepsilon}=\mathcal{O}(\varepsilon)$ and hence $q_k=0$ for all k. Hence $\lim_{\varepsilon\to 0}\frac{1-\lambda_\varepsilon}{\varepsilon}=\frac{1}{\ln 2}$.

(b) For the same map we consider $z=\frac{\sqrt{5}-1}{2}$ which is the rightmost fixed point of T. We have $T'(z)=-z^{-2}$. As $-z$ and z^{-1} are the two zeros of x^2-x-1, it is obvious that $1-\frac{1}{|T'(z)|}=z$. Hence, for intervals I_ε of length ε around z we have $\lim_{\varepsilon\to 0}\frac{1-\lambda_\varepsilon}{\varepsilon} = \frac{1}{\ln 2}\frac{z}{1+z}=\frac{z^2}{\ln 2}$.

(c) Denote $f(x)=\frac{1}{x}-1$. Then f is the rightmost branch of T, and the interval around z which is mapped by T^k onto $(0,1)$ has endpoints $f^{-k}(1)$ and $f^{-k}(0)$. Denote the length of this interval by ε_k and the interval itself by I_{ε_k}. As $-z$ and z^{-1} are the eigenvalues of the coefficient matrix of f^{-1}, a calculation shows that $\varepsilon_k = z^{2k+1}(1+z^2)^2(1+\mathcal{O}(z^{2k+2}))$.[2] In terms of the continued fraction algorithm the hole I_{ε_k} means that an expansion stops at time $n+k$ as soon as at least k consecutive digits 1 are generated. So it is natural to rewrite the limit from b) as in formula (1.2), namely

$$\lim_{k\to\infty}\frac{1-\lambda_{\varepsilon_k}}{z^{2k}} = \frac{z^3(1+z^2)^2}{\ln 2}\approx 0.6504. \tag{3.7}$$

The generalized cusp map: We focus on $z=1$, where the invariant density φ_0 vanishes, and consider holes $I_\varepsilon=[1-\varepsilon,1]$. Then, for ε close to 0, we have $\mu_0(I_\varepsilon)\sim\frac{\varphi_0(1/2)}{\gamma}\int_{1-\varepsilon}^1(1-x)^{\frac{1}{\gamma}-1}\,dx = \varphi_0(\frac{1}{2})\,\varepsilon^{\frac{1}{\gamma}}$ so that $\lim_{\varepsilon\to 0}\frac{1-\lambda_\varepsilon}{\varepsilon^{1/\gamma}}=\varphi_0(\frac{1}{2})$.

3.1.2 Holes Shrinking to a Nontrivial Hole

We assume now that I_0 is not a single point but an interval of some fixed length $\ell>0$ and the intervals $I_\varepsilon\supseteq I_0$ have length $\ell+\varepsilon$. To simplify the discussion we assume more specifically that $I_\varepsilon=[a-\varepsilon,a+\ell]$ where $a\in[0,1]$ is a continuity point of T and also of φ_0. Assume furthermore that a is not periodic for T (the periodic case can be dealt with analogously). Now, as μ_0 is supported by the non-trapped set X_{nt}, we have in particular $\mu_0(I_0)=0$ and hence $\Delta_\varepsilon=\mu_0(I_\varepsilon)$. It follows from Theorem 2.1 that either $\mu_0(I_\varepsilon)=0$ and hence $\lambda_\varepsilon=\lambda_0$ for all sufficiently small ε, or $\lim_{\varepsilon\to 0}\frac{\lambda_0-\lambda_\varepsilon}{\mu_0(I_\varepsilon)}=1$. But observe that μ_0 is of fractal nature, so

[2] Denote the coefficient matrix $\begin{pmatrix}0&1\\1&1\end{pmatrix}$ of f^{-1} by M and let $M^k=:\begin{pmatrix}a_k&b_k\\c_k&d_k\end{pmatrix}$. Then $f^{-k}(x_1)-f^{-k}(x_0)=\det(M^k)\frac{x_1-x_0}{(c_kx_1+d_k)(c_kx_0+d_k)}$ so that $\varepsilon_k=|f^{-k}(1)-f^{-k}(0)|=\frac{1}{|(c_k+d_k)d_k|}$. As $c_{k+1}=d_k$ and $d_{k+1}=(c_k+d_k)$, we have $\varepsilon_k=\frac{1}{|d_{k+1}d_k|}$ and $d_{k+1}=d_k+d_{k-1}$. With $d_0=d_1=1$ this yields $d_k=\frac{1}{1+z^{-2}}(z^{-(k+2)}+(-z)^k)$. Hence

$$\varepsilon_k=\frac{z^{2k}(1+z^{-2})^2}{(z^{-3}+(-1)^{k+1}z^{2k+1})(z^{-2}+(-1)^kz^{2k})}=z^{2k+1}(1+z^2)^2(1+\mathcal{O}(z^{2k+2})).$$

typically $\mu_0(I_\varepsilon)$ depends on ε in a devil's staircase manner. Hence either $\lambda_\varepsilon = \lambda_0$ for small ε (which happens if a itself is trapped), or $\lim_{\varepsilon\to 0} \frac{\varepsilon}{\mu_0(I_\varepsilon)} = 0$ and $\varepsilon \mapsto \lambda_\varepsilon$ is not differentiable at $\varepsilon = 0$.

3.2 Exchange Rates

We suppose that T has two ergodic components and that its restriction to each of these components is mixing. So the eigenvalue 1 of P has (geometric) multiplicity 2 and the rest of its spectrum is contained in a disk of radius smaller than some $\gamma \in (0, 1)$. Let $(\Pi_\varepsilon)_{\varepsilon\in E}$ be a family of Markov operators close to the identity with $\Pi_0 = \mathbb{I}$, and denote $P_\varepsilon := P \circ \Pi_\varepsilon$. (One could as well consider $\Pi_\varepsilon \circ P$ since that operator has the same eigenvalues as P_ε.) Under rather weak regularity assumptions on the Π_ε, the spectral perturbation results from [14] apply again. This is true, for example, if the Π_ε are convolutions with smooth densities $k_\varepsilon(x) = \varepsilon^{-1}k(\varepsilon^{-1}x)$ (modeling random perturbations) or if they are conditional expectations w.r.t. m and a finite partition into intervals of length ε (modeling Ulam's discretization scheme.) As the P_ε are also Markov operators, this means that 1 is an isolated eigenvalue of each P_ε. If it has multiplicity 2 there is nothing more to say about it. If it is a simple eigenvalue, however, then there is a second simple eigenvalue λ_ε close to 1 to which we will apply Theorem 2.1.

Let $(V, \|.\|)$ be the space

$$\mathrm{BV}_0 := \{f \in \mathrm{BV} : m(f) = 0\}.$$

BV_0 is invariant under all P_ε, and the previous discussion implies that assumptions (A1)–(A4) are satisfied. More precisely, $\lambda_0 = 1$, and there is an increasing function $\chi : [0, 1] \to \{-1, 1\}$ such that $\chi \circ T = \chi$, $|\varphi_0| = \chi\varphi_0$ is an invariant density for $P = P_0$, and $\nu_0 = \chi m$, so that $\mu_0 := \varphi_0\nu_0 = |\varphi_0|m$ is an invariant probability measure that gives equal mass to both ergodic components of T. Let $\tilde{I}_1 = \{\chi = -1\}$ and $\tilde{I}_2 = \{\chi = 1\}$ be the two invariant components of T.

In order to apply Theorem 2.1 let

$$p_\varepsilon(x) := \frac{1}{2}\left(1 - \left(\chi(x) \cdot (\Pi_\varepsilon^*\chi)(x)\right)\right)$$

where Π_ε^* is the dual of Π_ε with respect to Lebesgue measure on [0, 1]: It is easy to see that $p_\varepsilon(x)$ is the probability that the Markovian dynamics Π_ε move the system from the state x (that belongs to one of the two invariant components of T) to some state in the other component.

We have to check assumptions (A5) and (A6). Routine calculations show that $\nu_0(P_0 - P_\varepsilon)(\psi) = 2m(\chi \cdot p_\varepsilon \cdot \psi)$ for each $\psi \in V$. This implies

$$\Delta_\varepsilon = \nu_0(P_0 - P_\varepsilon)(\varphi_0) = 2m(|\varphi_0| \cdot p_\varepsilon) = 2\mu_0(p_\varepsilon) \tag{3.8}$$

and $\nu_0(P_0 - P_\varepsilon)(\psi) = 2m(\chi p_\varepsilon \psi) \leq 2\|\psi\| \int_0^1 |p_\varepsilon| dm$ so that $\eta_\varepsilon \leq 2m(p_\varepsilon)$. Therefore we require that the average probability $m(p_\varepsilon)$ to change the invariant component under the action of Π_ε tends to 0 as $\varepsilon \to 0$ and that

$$\|(\mathbb{I} - \Pi_\varepsilon)(\varphi_0)\| \leq \mathrm{const} \cdot \frac{1}{m(p_\varepsilon)} \int |\varphi_0| p_\varepsilon \, dm. \tag{3.9}$$

In the following we will assume

$$\inf |\varphi_0| > 0.$$

This trivially implies (3.9) although the latter can be verified in many other cases. It remains to check assumption (A7). Observing (3.8) and

$$\begin{aligned}\nu_0\left((P_0-P_\varepsilon)P_\varepsilon^k(P_0-P_\varepsilon)(\varphi_0)\right) &= m((\chi-\Pi_\varepsilon^*\chi)\cdot P_\varepsilon^k(P_0-P_\varepsilon)(\varphi_0))\\ &= 2m\left(p_\varepsilon\chi\cdot P_\varepsilon^k P_0(\varphi_0-\Pi_\varepsilon\varphi_0)\right)\end{aligned}\tag{3.10}$$

we get the following expression for the $q_{k,\varepsilon}$:

$$q_{k,\varepsilon}=\frac{1}{m(|\varphi_0|p_\varepsilon)}m\left(p_\varepsilon\chi\cdot P_\varepsilon^k P_0(\varphi_0-\Pi_\varepsilon\varphi_0)\right).\tag{3.11}$$

The evaluation of the limit as $\varepsilon\to 0$ depends strongly on the details of the map T and of the perturbation. We therefore make some further simplifying assumptions:

- The Π_ε are local perturbations, i.e., for each x, $\Pi_\varepsilon\delta_x$ is supported in a $C\varepsilon$-neighborhood of x.
- T is continuous.

As the restrictions of T to its two ergodic components are mixing, the continuity of T implies that the non-wandering part of these components are just two single intervals $I_i\subset\tilde I_i$. If these intervals do not have a common end point, then $\Delta_\varepsilon=2\mu_0(p_\varepsilon)=0$ for small ε so that $\lambda_\varepsilon=1$ for such ε by our main theorem. Otherwise I_1 and I_2 have a common endpoint z. Since two interval can have at most one common endpoint and since the map is continuous, it follows by the invariance of I_1, I_2 that z is a fixed point. In this case, $p_\varepsilon(x)=0$ unless x belongs to the $C\varepsilon$-neighborhood of z. As an example let us consider the special (but still rather general) class of examples characterized by the following properties

- $\overline{I_1\cup I_2}=[0,1]$.[3]
- Assume $\Pi_\varepsilon f(x)=\int_0^1 K_\varepsilon(y,x)f(y)dy$ where K_ε is a positive kernel such that, for all $y\in[0,1]$, $\int_0^1 K_\varepsilon(y,x)dx=1$. In order to satisfy assumptions (A1)–(A4) one should suppose that the kernels are bistochastic or close to convolution kernels in the sense of [5, Corollary 3.20].
- There exists $a>0$ such that $K_\varepsilon(y,x)=\varepsilon^{-1}K(\varepsilon^{-1}(x-y))$ provided $|z-y|\le a$. Here K is a smooth probability density supported in $[-1,1]$.[4]
- φ_0 is continuous in each ergodic component and T is differentiable at z.

Note that, since $|\varphi_0|$ must give the same weight to the two ergodic components it will, in general, be discontinuous at z. Let α,β be the left and right limit respectively. Then, introducing coordinates $x=z+\varepsilon\zeta$ and setting $\theta(y)=\operatorname{sign} y$, we have $\varphi_0(z+\varepsilon\zeta)=\frac{\beta+\alpha}{2}\theta(\zeta)+\frac{\beta-\alpha}{2}+o(1)$, uniformly for ζ in a compact set. Next, for ε small enough,

$$p_\varepsilon(z+\varepsilon\zeta)=\frac{1}{2}\int_{\mathbb{R}}K(y-\zeta)[1-\theta(y)\theta(\zeta)]dy=:p(\zeta).$$

Note that $p\ge 0$ and $p(\zeta)=0$ if $|\zeta|>1$. Accordingly,

$$\mu_0(p_\varepsilon)=m(|\varphi_0|p_\varepsilon)=\varepsilon\frac{\beta-\alpha}{2}m(\theta p)+\varepsilon\frac{\beta+\alpha}{2}m(p)+o(\varepsilon)=:\varepsilon\Gamma+o(\varepsilon).$$

[3] If the wandering part is present, the final result still holds with I_i substituted by $\bigcup_{n\in\mathbb{N}}T^n I_i$ in (3.13).

[4] One can consider the more general case $K_\varepsilon(y,x)=\varepsilon^{-1}\tilde K(y,\varepsilon^{-1}(x-y))$, for some smooth function $\tilde K$. The final formula then holds with $K(\cdot)$ replaced by $\tilde K(z,\cdot)$.

In addition, for each function Ψ_ε such that $\Psi_\varepsilon(z+\varepsilon\zeta)=\psi(\zeta)+o(1)$, for some fixed compact support function ψ, holds, in the limit $\varepsilon\to 0$, $(\Pi_\varepsilon\Psi_\varepsilon)(z+\varepsilon\zeta)=\int_{\mathbb{R}}K(\zeta'-\zeta)\psi(\zeta')d\zeta'+o(1)$ and $(P_0\Psi_\varepsilon)(z+\varepsilon\zeta)=\Lambda\cdot\psi(\Lambda\zeta)+o(1)$ where $\Lambda:=\frac{1}{T'(z)}>0$. Hence,

$$\begin{aligned}(P_\varepsilon\Psi_\varepsilon)(z+\varepsilon\zeta)&=\Lambda\int_{\mathbb{R}}K(\zeta\Lambda-\zeta')\psi(\zeta')d\zeta'+o(1)\\&=:(\mathcal{K}\psi)(\zeta)+o(1).\end{aligned}$$

The above setting applies to $\Psi_\varepsilon=P_0(\mathbb{I}-\Pi_\varepsilon)\varphi_0$, namely

$$[P_0(\mathbb{I}-\Pi_\varepsilon)\varphi_0](z+\varepsilon\zeta)=\frac{\beta+\alpha}{2}[\Lambda\theta(\zeta)-\mathcal{K}\theta(\zeta)]+o(1).$$

Indeed, $\mathcal{K}\theta(\zeta)=\Lambda\theta(\zeta)$ for $|\zeta|\geq\Lambda^{-1}$, hence $\Lambda\theta-\mathcal{K}\theta$ is compactly supported. Thus

$$q_k=\frac{(\beta+\alpha)\langle\theta p,\mathcal{K}^k(\Lambda\mathbb{I}-\mathcal{K})\theta\rangle}{2\Gamma}.$$

Since the operator $\mathcal{K}$ has L^∞ norm smaller than $|\Lambda|$, the latter equality implies

$$\lim_{\varepsilon\to 0}\frac{\lambda_0-\lambda_\varepsilon}{\varepsilon}=2\Gamma-\frac{(\beta+\alpha)}{2}\sum_{n=0}^{\infty}\langle 2\theta p,\mathcal{K}^n(\Lambda\mathbb{I}-\mathcal{K})\theta\rangle.$$

On the other hand, a direct computation (observing the fact that $\theta(y)=\theta(\Lambda y)$) shows that $2\theta p=(\mathbb{I}-\mathcal{K}^*)\theta$, so

$$\lim_{\varepsilon\to 0}\frac{\lambda_0-\lambda_\varepsilon}{\varepsilon}=2\Gamma-\frac{\beta+\alpha}{2}\Big[\langle\theta,(\Lambda\mathbb{I}-\mathcal{K})\theta\rangle-\lim_{n\to\infty}\langle(\mathcal{K}^*)^n\theta,(\Lambda\mathbb{I}-\mathcal{K})\theta\rangle\Big].$$

Note that $(\mathcal{K}^*)^n\theta$ converges pointwise to a function θ_∞ such that $\theta-\theta_\infty$ is supported in the interval $[-(1-\Lambda)^{-1},(1-\Lambda)^{-1}]$, see (3.15). In particular, $\mathcal{K}^*\theta_\infty=\theta_\infty$. Then,

$$\begin{aligned}\lim_{\varepsilon\to 0}\frac{1-\lambda_\varepsilon}{\varepsilon}&=2\Gamma-\frac{\beta+\alpha}{2}\langle\theta-\theta_\infty,(\Lambda\mathbb{I}-\mathcal{K})\theta\rangle\\&=2\Gamma-\frac{\beta+\alpha}{2}\left[\langle(\mathbb{I}-\mathcal{K}^*)(\theta-\theta_\infty),\theta\rangle+(\Lambda-1)\langle\theta-\theta_\infty,\theta\rangle\right]\\&=2\Gamma-\frac{\beta+\alpha}{2}\left[\langle 2\theta p,\theta\rangle+(\Lambda-1)m(1-\theta\theta_\infty)\right]\\&=\frac{\beta+\alpha}{2}\left(1-\frac{1}{T'(z)}\right)m(1-\theta\theta_\infty)+\frac{\beta-\alpha}{2}m(2\theta p).\end{aligned}\tag{3.12}$$

To make the formula more explicit and transparent let us make some further remarks. The original dynamical system has a natural invariant measure with density by $h=\lim_{n\to\infty}P_0^n 1$. By our assumptions, $\{h,\chi h\}$ is a basis for the eigenspace of the eigenvalue one of the operator P_0. Thus $\varphi_0=a\chi h+bh$ for suitable $a,b\in\mathbb{R}$. Recall that $\int_{I_1}|\varphi_0|\,dm=\int_{I_2}|\varphi_0|\,dm=\frac{1}{2}$. Hence $\frac{1}{2}=|b-a|\int_{I_1}h\,dm=|b-a|\lim_{n\to\infty}\int_0^1\frac{1-\chi}{2}\circ T^n\,dm=|b-a|m(I_1)$ and, analogously, $\frac{1}{2}=|b+a|m(I_2)$. Therefore,

$$\alpha=|\varphi_0(z^-)|=\frac{h(z^-)}{2m(I_1)}\quad\text{and}\quad\beta=|\varphi_0(z^+)|=\frac{h(z^+)}{2m(I_2)}.\tag{3.13}$$

To describe the meaning of the two factors involving θ and θ_∞, let Z be a random variable whose distribution has probability density K. Then

$$m(2\theta p) = -2\mathbb{E}[Z]. \tag{3.14}$$

Next let $Z_1, Z_2, \dots$ be independent copies of Z. The kernel $\mathcal{K}^*$ describes a Markov process

$$X_n = \Lambda^{-1}(X_{n-1} + Z_n) = \dots = \Lambda^{-n}\left(X_0 + \sum_{k=1}^{n} \Lambda^{k-1} Z_k\right).$$

The asymptotic behavior of the process (X_n) is determined by the random variable $W := \sum_{k=1}^{\infty} \Lambda^{k-1} Z_k$. Indeed, let $X_0 = \zeta$. Then $X_n \to +\infty$ if $W > -\zeta$ and $X_n \to -\infty$ if $W < -\zeta$. (As the Z_k have density, $W = 0$ has probability 0.) As $((\mathcal{K}^*)^n\theta)(\zeta)$ is the conditional expectation of $\theta(X_n)$ given $X_0 = \zeta$, it follows readily that

$$\begin{aligned}\theta_\infty(\zeta) &= \mathbb{P}(X_n \to +\infty | X_0 = \zeta) - \mathbb{P}(X_n \to -\infty | X_0 = \zeta) \\ &= 1 - 2\mathbb{P}(W < -\zeta). \end{aligned} \tag{3.15}$$

Hence

$$m(1 - \theta\theta_\infty) = 2\int_0^\infty [\mathbb{P}(W > \zeta) + \mathbb{P}(-W > \zeta)]\, d\zeta = 2\mathbb{E}[|W|]. \tag{3.16}$$

Note that $(1 - 1/T'(z))m(1 - \theta\theta_\infty) = 2(1 - \Lambda)\mathbb{E}[|W|] \geq 2(1 - \Lambda)|\mathbb{E}[W]| = 2|\mathbb{E}[Z]| = |m(2\theta p)|$, so the r.h.s. of (3.12) is clearly positive.

We finish this section with a comment on the term *exchange rate*. Let $A_\varepsilon^+ := \{\varphi_\varepsilon > 0\}$, $A_\varepsilon^- := \{\varphi_\varepsilon < 0\}$, and $\tilde{p}_\varepsilon := 1_{A_\varepsilon^+} \cdot P_\varepsilon^* 1_{A_\varepsilon^-} + 1_{A_\varepsilon^-} \cdot P_\varepsilon^* 1_{A_\varepsilon^+}$. $\tilde{p}_\varepsilon(x)$ is the probability to exchange the sets $A_\varepsilon^\pm$ under the action of P_ε^*. Now Proposition 5.7 from [8] can be rephrased in our setting as $1 - \lambda_\varepsilon = 2\int \tilde{p}_\varepsilon |\varphi_\varepsilon|\, dm$,[5] so it is nearly twice the "stationary exchange rate" $\int \tilde{p}_\varepsilon h_\varepsilon\, dm$ where $h_\varepsilon = P_\varepsilon h_\varepsilon$ is the unique invariant probability density of the perturbed system. If all A_ε are identical (e.g. under suitable symmetry assumptions on the system as in [8, Corollary 5.9]), then $\tilde{p}_\varepsilon$ coincides with p_ε from above.

4 An Application to Two Coupled Interval Maps

Let $T : [0, 1] \to [0, 1]$ be a mixing piecewise expanding map as in Sect. 3. To simplify the discussion we assume that $\gamma := \inf |T'| > 4$. Let $M := [0, 1]^2$ and define, for $\delta \in [0, \frac{1}{4} - \frac{1}{\gamma})$, the two-dimensional coupled map

$$\hat{T} : M \to M, \quad \hat{T}(x, y) = ((1 - \delta)T(x) + \delta T(y), (1 - \delta)T(y) + \delta T(x)).$$

It is uniformly piecewise expanding with minimal expansion strictly larger than 2 in the sense that

$$\|(D\hat{T})^{-1}\| \leq \frac{1}{\gamma(1 - 2\delta)} < \frac{1}{4}.$$

[5]The operator P in [8] corresponds to our P_ε and the signed measure ν to our $\varphi_\varepsilon m$.

As discussed in great detail in [15] there is $\delta_1 \in (0, \frac{1}{2} - \frac{1}{\gamma}]$ such that, for $\delta \in [0, \delta_1]$, $\hat{T}$ is mixing in the sense that its Perron-Frobenius operator $\hat{P} : \mathrm{BV}(M) \to \mathrm{BV}(M)$ has a unique invariant probability density $\hat{h}$ and a spectral gap. Here $\mathrm{BV}(M)$ is the space of functions of bounded variation on $\mathbb{R}^2$ that vanish outside M.

For $\varepsilon \in E := [0, \varepsilon_1]$ let $S_\varepsilon := \{(x, y) \in M : |x - y| \le \varepsilon\}$. If we interpret S_ε as a hole in the phase space M, this means that we stop a trajectory as soon as the two components have synchronized up to a difference of at most ε. The corresponding Perron-Frobenius operator $\hat{P}_\varepsilon : \mathrm{BV}(M) \to \mathrm{BV}(M)$ is defined by $\hat{P}_\varepsilon(\psi) = \hat{P}(\psi \cdot 1_{M \setminus S_\varepsilon})$. Denote the (two-dimensional) variation of a function $\psi \in \mathrm{BV}(M)$ by $\mathrm{Var}(\psi)$. It is easy to check that $\mathrm{Var}(\hat{P}_\varepsilon \psi) \le 2\,\mathrm{Var}(\hat{P}\psi)$ so that the family of operators $\hat{P}_\varepsilon$ satisfies a uniform Lasota-Yorke inequality. (Observe that we made the generous assumption $\gamma > 4$ and consult [15].) In view of the spectral stability results of [14], assumptions (A1)–(A4) are satisfied with $\nu_0 = m$ (the Lebesgue measure on M), $\varphi_0 = \hat{h}$, $\mu_0 = \hat{h}m$ and $\lambda_0 = 1$.

We turn to assumptions (A5) and (A6). Observe first that

$$\nu_0(P_0 - P_\varepsilon)(\psi) = m(\psi 1_{S_\varepsilon}) \le C\varepsilon\,\mathrm{Var}(\psi). \tag{4.1}$$

(The constant C depends on the details of the definition of the variation.) So in particular $\eta_\varepsilon \le C\varepsilon$ and (A5) is satisfied. As $\hat{h}$ is of bounded variation, we may assume that it is regularized along the diagonal of M in the sense that for 1D-Lebesgue-almost every x the value $\hat{h}(x, x)$ is the average of the limits of $\hat{h}(x - u, x + u)$ and $\hat{h}(x + u, x - u)$ as $u \searrow 0$. In view of (4.1) we therefore conclude

$$\lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \Delta_\varepsilon = \lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \int_{S_\varepsilon} \hat{h}\, dm = \int_0^1 \hat{h}(x, x)\, dx. \tag{4.2}$$

As $\mathrm{Var}(\hat{h} 1_{S_\varepsilon}) \le 2\,\mathrm{Var}(\hat{h})$, we conclude that (A6) is satisfied if $\int_0^1 \hat{h}(x, x)\, dx > 0$. It remains to evaluate the q_k. As in Sect. 3 let

$$\hat{U}_{k,\varepsilon} := \hat{T}^{-1}(M \setminus S_\varepsilon) \cap \cdots \cap \hat{T}^{-k}(M \setminus S_\varepsilon) \cap \hat{T}^{-(k+1)} S_\varepsilon.$$

Then $q_{k,\varepsilon} = \mu_0(S_\varepsilon \cap \hat{U}_{k,\varepsilon}) / \mu_0(S_\varepsilon)$ and, since the diagonal of M is invariant under $\hat{T}$, we find

$$q_0 = \lim_{\varepsilon \to 0} q_{0,\varepsilon} = \frac{1}{\int_0^1 \hat{h}(x, x)\, dx} \int_0^1 \hat{h}(x, x) \frac{1}{(1 - 2\delta)|T'(x)|}\, dx$$

and $q_k = \lim_{\varepsilon \to 0} q_{k,\varepsilon} = 0$ for all $k \ge 1$. So finally,

$$\lim_{\varepsilon \to 0} \frac{1 - \lambda_\varepsilon}{2\varepsilon} = \int_0^1 \hat{h}(x, x) \left(1 - \frac{1}{(1 - 2\delta)|T'(x)|}\right) dx. \tag{4.3}$$

5 Related Results

5.1 Metastable States in Molecular Dynamics and Oceanic Structures

Phase space methods to characterize biomolecular conformations as metastable states are used in molecular dynamics (see e.g. [20] and references cited there). Very roughly, if the Markov operator P_ε describes the discrete time evolution of such a system in a fixed time scale and if $X = D_1 \cup D_2$ is a decomposition (up to null sets) of the underlying phase space, then the *metastability measure* of this decomposition is defined as $\text{meta}(D_1, D_2) = \frac{1}{2}[\mu_\varepsilon(D_1)^{-1} \int_{D_1} P_\varepsilon 1_{D_1} d\mu_\varepsilon + \mu_\varepsilon(D_2)^{-1} \int_{D_2} P_\varepsilon 1_{D_2} d\mu_\varepsilon]$. Theorem 1 of [20] relates $\text{meta}(D_1, D_2)$ to the second eigenvalue of P_ε in a way very similar to formula (3.12). Reference [11] is an up-to-date review of the phase space decomposition approach to metastability in general flow dynamical systems, and [12] is an application of these ideas to the detection of coherent oceanic structures. In the framework of weakly coupled rapidly mixing Markov chains, reference [21] also relates the second largest eigenvalue of a system to the exchange probabilities between its components.

5.2 Shannon Capacity of Constrained Systems of Binary Sequences

In information theory, the topological entropy of subshifts of $\{0, 1\}^{\mathbb{N}}$ that are determined by a (short) list $\mathbb{L}_m = (B_1^{(m)}, \dots, B_p^{(m)})$ of distinct blocks of length m which are not allowed to occur [18] is called the Shannon capacity of the system. It is closely related to the rate of periodic prefix-synchronized (PPS) codes with markers $B_i^{(m)}$, $i = 1, \dots, p$ (see e.g. [16]). For each sequence $\mathbb{L}_1, \mathbb{L}_2, \dots$ of such lists with fixed length p there are a subsequence (m_j) and $z_1, \dots, z_p \in \{0, 1\}^{\mathbb{N}}$ such that the $B_i^{(m_j)}$ converge to z_i as $j \to \infty$. We will assume without loss that the full sequences $(B_i^{(m)})_m$ converge. As the full two-shift is isomorphic (for each invariant measure of positive entropy) to the doubling map,[6] the shift constrained by the forbidden blocks in $\mathbb{L}_m$ is isomorphic to the doubling map T with "hole" I_m being the union of those monotonicity intervals of T^m labeled by the words in $\mathbb{L}_m$. Hence the topological entropy $h(\mathbb{L}_m)$ of this shift equals $\log(2\lambda_m)$ where λ_m is the leading eigenvalue of the Perron-Frobenius operator of T with hole I_m, compare Sect. 3.1.

If the limit points z_i belong to $B_i^{(m)}$ for all i and m, one can analyze the situation just as in Sect. 3.1. Some elementary reasoning yields the following: for $i = 1, \dots, p$ let $\ell(i) = \min\{j \geq 1 : T^j z_i \in \{z_1, \dots, z_p\}\}$ with the convention that $\ell(i) = +\infty$ if no such j exists. Then

$$\lim_{m\to\infty} \frac{\log 2 - h(\mathbb{L}_m)}{2^{-m}} = \sum_{i=1}^{p} (1 - 2^{-\ell(i)}).$$

This is minimal when all $\ell(i) = 1$, e.g. if the z_i form just one periodic orbit. In that case $h(\mathbb{L}_m) = \log 2 - p2^{-(m-1)} + o(2^{-m})$, which supports the conjecture on the precise values of $h(\mathbb{L}_m)$ for $p = 2^k$ and $m \geq k + 1$ stated in [16].

6 Proof of the Main Theorem

As announced in Remark 2.2 we prove Theorem 2.1 under the weaker summability assumption (A3*). The reader who does not want to follow this slight generalization of the argument may just neglect all "$*$" attached to the norms. We use the following notation:

$$\kappa_N := \sum_{n=N}^{\infty} \sup_{\varepsilon\in E} \|Q_\varepsilon^n\|^*.$$

[6] Just associate to each trajectory $x \in [0, 1]$ the sequence $(\sigma_n) \in \{0, 1\}^{\mathbb{N}}$ where $\sigma_n = 0$ iff $T^n x \in [0, 1/2]$. This induces a measurable isomorphism (modulo null sets) between the doubling map and the full one-sided shift.

Lemma 6.1 *There is a constant $C>0$ such that, for all $\varepsilon\in E$ and all $N\ge 0$,*

(a) $|1-\nu_\varepsilon(\varphi_0)|\le C\eta_\varepsilon$,
(b) $\|Q_\varepsilon^N\varphi_0\|\le C\kappa_N(\|(P_0-P_\varepsilon)(\varphi_0)\|_*+|\lambda_0-\lambda_\varepsilon|)$.

Proof (a) As $(\mathbb{I}-\lambda_\varepsilon^{-1}P_\varepsilon)(\lambda_\varepsilon^{-1}P_\varepsilon)^k(\varphi_0)=(\mathbb{I}-\lambda_\varepsilon^{-1}P_\varepsilon)Q_\varepsilon^k(\varphi_0)$ for all $k\ge 0$,

$$\begin{aligned}|1-\nu_\varepsilon(\varphi_0)|&=\lim_{n\to\infty}\left|\nu_0\left(\varphi_0-(\lambda_\varepsilon^{-1}P_\varepsilon)^n(\varphi_0)\right)\right|\\&\le\sum_{k=0}^{\infty}\left|\nu_0\left((\mathbb{I}-\lambda_\varepsilon^{-1}P_\varepsilon)Q_\varepsilon^k(\varphi_0)\right)\right|\\&=\sum_{k=0}^{\infty}\left|\nu_0\left((\lambda_0^{-1}P_0-\lambda_\varepsilon^{-1}P_\varepsilon)Q_\varepsilon^k(\varphi_0)\right)\right|\\&\le|\lambda_0|^{-1}\eta_\varepsilon\sum_{k=0}^{\infty}\|Q_\varepsilon^k\|^*\|\varphi_0\|_*+|\lambda_0^{-1}||\lambda_\varepsilon-\lambda_0|\|\nu_0\|\sum_{k=1}^{\infty}\|Q_\varepsilon^k\|^*\|\varphi_0\|_*\\&=\mathcal{O}(\eta_\varepsilon)+\mathcal{O}(\lambda_0-\lambda_\varepsilon)=\mathcal{O}(\eta_\varepsilon)\end{aligned}$$

where we used (A3*) and (2.3) for the last estimate.

(b) For each $N\ge 0$ we have

$$\begin{aligned}\|Q_\varepsilon^N\varphi_0\|&\le\limsup_{n\to\infty}\|Q_\varepsilon^N(\varphi_0-(\lambda_\varepsilon^{-1}P_\varepsilon)^n(\varphi_0))\|+\limsup_{n\to\infty}\|Q_\varepsilon^{N+n}\varphi_0\|\\&\le\sum_{k=0}^{\infty}\|Q_\varepsilon^N(\lambda_\varepsilon^{-1}P_\varepsilon)^k(\mathbb{I}-\lambda_\varepsilon^{-1}P_\varepsilon)(\varphi_0)\|+\limsup_{n\to\infty}\kappa_{N+n}\|\varphi_0\|_*\\&\le|\lambda_0^{-1}|\sum_{k=0}^{\infty}\left(\|Q_\varepsilon^{N+k}(P_0-P_\varepsilon)(\varphi_0)\|+|\lambda_0-\lambda_\varepsilon|\|Q_\varepsilon^{N+k+1}\varphi_0\|\right)\\&\le|\lambda_0^{-1}|\sum_{k=0}^{\infty}\|Q_\varepsilon^{N+k}\|^*\left(\|(P_0-P_\varepsilon)(\varphi_0)\|_*+|\lambda_0-\lambda_\varepsilon|\|\varphi_0\|_*\right)\\&=\mathcal{O}(\kappa_N)\big(\|(P_0-P_\varepsilon)(\varphi_0)\|_*+|\lambda_0-\lambda_\varepsilon|\big).\end{aligned}$$

□

Proof of Theorem 2.1 Observe first that by (2.2), for each $n>0$,

$$\begin{aligned}&\nu_\varepsilon(\varphi_0)(\lambda_0-\lambda_\varepsilon)\\&\quad=\nu_\varepsilon(\varphi_0)\nu_0((P_0-P_\varepsilon)(\varphi_\varepsilon))\\&\quad=\Delta_\varepsilon-\nu_0\left((P_0-P_\varepsilon)(\mathbb{I}-(\lambda_\varepsilon^{-1}P_\varepsilon)^n)(\varphi_0)\right)-\nu_0\left((P_0-P_\varepsilon)Q_\varepsilon^n(\varphi_0)\right)\\&\quad=\Delta_\varepsilon-\sum_{k=0}^{n-1}\nu_0\left((P_0-P_\varepsilon)(\lambda_\varepsilon^{-1}P_\varepsilon)^k(\mathbb{I}-\lambda_\varepsilon^{-1}P_\varepsilon)(\varphi_0)\right)+\mathcal{O}(\eta_\varepsilon\|Q_\varepsilon^n\varphi_0\|)\\&\quad=\Delta_\varepsilon-\lambda_0^{-1}\sum_{k=0}^{n-1}\nu_0\left((P_0-P_\varepsilon)(\lambda_\varepsilon^{-1}P_\varepsilon)^k(P_0-P_\varepsilon)(\varphi_0)\right)\end{aligned}$$

$$+\lambda_0^{-1}(\lambda_0-\lambda_\varepsilon)\sum_{k=1}^{n}\nu_0\big((P_0-P_\varepsilon)(\lambda_\varepsilon^{-1}P_\varepsilon)^k(\varphi_0)\big)$$

$$+\mathcal{O}(\kappa_n)\big(|\Delta_\varepsilon|+\eta_\varepsilon|\lambda_0-\lambda_\varepsilon|\big)\quad \text{(by Lemma 6.1b and (A6*))}$$

$$=\Delta_\varepsilon\left(1-\lambda_0^{-1}\sum_{k=0}^{n-1}\lambda_\varepsilon^{-k}q_{k,\varepsilon}\right)$$

$$+\mathcal{O}(\eta_\varepsilon)|\lambda_0-\lambda_\varepsilon|\sum_{k=1}^{n}\big(|\nu_\varepsilon(\varphi_0)|\|\varphi_\varepsilon\|+\|Q_\varepsilon^k\varphi_0\|\big)+\mathcal{O}(\kappa_n)\big(|\Delta_\varepsilon|+\eta_\varepsilon|\lambda_0-\lambda_\varepsilon|\big)$$

where

$$q_{k,\varepsilon}:=\frac{\nu_0((P_0-P_\varepsilon)P_\varepsilon^k(P_0-P_\varepsilon)(\varphi_0))}{\nu_0((P_0-P_\varepsilon)(\varphi_0))}. \tag{6.1}$$

Observing Lemma 6.1a, (A4) and Lemma 6.1b, the error terms can be estimated by $\mathcal{O}(\eta_\varepsilon)n|\lambda_0-\lambda_\varepsilon|+\mathcal{O}(\kappa_n)|\Delta_\varepsilon|$ so that, in view of Lemma 6.1a, this yields, for each $n>0$,

$$(1+\mathcal{O}(\eta_\varepsilon))(\lambda_0-\lambda_\varepsilon)(1+n\mathcal{O}(\eta_\varepsilon))=\Delta_\varepsilon\left(1-\lambda_0^{-1}\sum_{k=0}^{n-1}\lambda_\varepsilon^{-k}q_{k,\varepsilon}\right)+\mathcal{O}(\kappa_n)|\Delta_\varepsilon|. \tag{6.2}$$

If $\Delta_\varepsilon=0$ and η_ε is small, it follows that $\lambda_\varepsilon=\lambda_0$. Otherwise we assumed in (A7) that $q_k=\lim_{\varepsilon\to 0}q_{k,\varepsilon}$ exists for each k, and we conclude

$$\lim_{\varepsilon\to 0}\frac{\lambda_0-\lambda_\varepsilon}{\Delta_\varepsilon}=1-\sum_{k=0}^{n-1}\lambda_0^{-(k+1)}q_k+\mathcal{O}(\kappa_n)$$

for each $n>0$. From this the claim (2.4) follows in the limit $n\to\infty$. □

References

1. Baladi, V.: On the susceptibility function of piecewise expanding interval maps. Commun. Math. Phys. **275**, 839–859 (2007)
2. Baladi, V., Gouezel, S.: Good Banach spaces for piecewise hyperbolic maps via interpolation, arXiv:0711.1960v1
3. Baladi, V., Smania, D.: Linear response formula for piecewise expanding unimodal maps. Nonlinearity **21**, 677–711 (2008)
4. Blank, M.: Stochastic properties of deterministic dynamical systems. Sov. Sci. Rev. C Math./Phys. **6**, 243–271 (1987)
5. Blank, M., Keller, G.: Stochastic stability versus localization in one-dimensional chaotic dynamical systems. Nonlinearity **10**, 81–107 (1997)
6. Bunimovich, L.A., Dettmann, C.P.: Peeping at chaos: nondestructive monitoring of chaotic systems by measuring long-time escape rates. Europhys. Lett. EPL **80**(4), 6 (2007), Art. 40001
7. Bunimovich, L.A., Yurchenko, A.: Where to place an hole to achieve maximal escape rate. Preprint
8. Dellnitz, M., Junge, O.: On the approximation of complicated dynamical behavior. SIAM J. Numer. Anal. **36**, 491–515 (1999)
9. Demers, M., Liverani, C.: Stability of statistical properties in two-dimensional piecewise hyperbolic maps. Trans. Am. Math. Soc. **360**, 4777–4814 (2008)
10. Dettmann, C.P., Howard, T.B.: Asymptotic expansions for the escape rate of stochastically perturbed dynamical systems, arXiv:0805.4570v1
11. Froyland, G., Padberg, K.: Almost-invariant sets and invariant manifolds—connecting probabilistic and geometric descriptions of coherent structures in flows. Preprint (2008)

12. Froyland, G., Padberg, K., England, M.H., Treguier, A.M.: Detection of coherent oceanic structures via transfer operators. Phys. Rev. Lett. **98**, 224503 (2007)
13. Gouezel, S., Liverani, C.: Banach spaces adapted to Anosov systems. Ergod. Theory Dyn. Syst. **26**(1), 189–217 (2006)
14. Keller, G., Liverani, C.: Stability of the spectrum for transfer operators. Ann. Sc. Norm. Super. Pisa, Sci. Fis. Mat. (4) **XXVIII**, 141–152 (1999)
15. Keller, G., Liverani, C.: A spectral gap for a one-dimensional lattice of coupled piecewise expanding interval maps. In: Chazottes, J.-R., Fernandez, B. (eds.) Dynamics of Coupled Map Lattices and of Related Spatially Extended Systems. Lecture Notes in Physics, vol. 671, pp. 115–151. Springer, Berlin (2005)
16. Kashyap, N.: Maximizing the Shannon capacity of constrained systems with two constrainets. SIAM J. Discrete Math. **17**, 298–319 (2003)
17. Lasota, A., Yorke, J.: Existence of invariant measures for piecewise monotonic transformations. Trans. Am. Math. Soc. **186**, 481–488 (1973)
18. Lind, D.: Perturbations of shifts of finite type. SIAM J. Discrete Math. **2**, 350–365 (1989)
19. Liverani, C., Maume-Deschamps, V.: Lasota-Yorke maps with holes: conditionally invariant probability measures and invariant probability measures on the survivor set. Ann. Inst. Henri Poincaré (B) Probab. Stat. **39**, 385–412 (2003)
20. Meerbach, E., Dittmer, E., Horenko, I., Schtte, Ch.: Multiscale modelling in molecular dynamics: biomolecular conformations as metastable states. In: Ferrario, M., Ciccotti, G., Binder, K. (eds.) Computer Simulations in Condensed Matter: From Materials to Chemical Biology, Vol. I. Lecture Notes in Physics, vol. 703, pp. 475–497. Springer, Berlin (2006)
21. Meerbach, D., Schütte, Ch., Fischer, A.: Eigenvalue bounds on restrictions of reversible nearly uncoupled Markov chains. Linear Algebra Appl. **398**, 141–160 (2005)
22. Ruelle, D.: Differentiation of SRB states. Commun. Math. Phys. **187**, 227–241 (1997). See also Commun. Math. Phys. **234**, 185–190 (2003)
23. Rychlik, M.: Bounded variation and invariant measures. Stud. Math. **76**, 69–80 (1983)
24. Saussol, B.: Absolutely continuous invariant measures for multidimensional expanding maps. Isr. J. Math. **116**, 223–248 (2000)
25. Bahsoun, W.: Rigorous numerical approximation of escape rates. Nonlinearity **19**(11), 2529–2542 (2006)