IOP PUBLISHING NONLINEARITY

Nonlinearity **21** (2008) 1719–1743 doi:10.1088/0951-7715/21/8/003

Continuity properties of transport coefficients in simple maps

Gerhard Keller[1], **Phil J Howard**[2] **and Rainer Klages**[2]

[1] Mathematisches Institut, Universität Erlangen-Nürnberg, Bismarckstr. $1\frac{1}{2}$, 91054 Erlangen, Germany
[2] School of Mathematical Sciences, Queen Mary, University of London, Mile End Road, London E1 4NS, UK

Received 17 January 2008, in final form 22 April 2008
Published 13 June 2008
Online at stacks.iop.org/Non/21/1719

Recommended by C P Dettmann

Abstract
We consider families of dynamics that can be described in terms of Perron–Frobenius operators with exponential mixing properties. For piecewise C^2 expanding interval maps we rigorously prove continuity properties of the drift $J(\lambda)$ and of the diffusion coefficient $D(\lambda)$ under parameter variation. Our main result is that $D(\lambda)$ has a modulus of continuity of order $\mathcal{O}(|\delta\lambda| \cdot (\log|\delta\lambda|)^2)$, i.e. $D(\lambda)$ is Lipschitz continuous up to quadratic logarithmic corrections. For a special class of piecewise linear maps we provide more precise estimates at specific parameter values. Our analytical findings are quantified numerically for the latter class of maps by using exact series expansions for the transport coefficients that can be evaluated numerically. We numerically observe strong local variations of all continuity properties.

Mathematics Subject Classification: 37C30, 37C40, 37E05, 82C70

PACS numbers: 05.45.Ac, 05.60.Cd

1. Introduction

In simple deterministic dynamical systems physical quantities like transport coefficients can be fractal functions of control parameters. This finding was first reported for a one-dimensional piecewise linear map lifted periodically onto the whole real line, for which the diffusion coefficient was computed by using Markov partitions and topological transition matrices [27,28,30]. A generalization of this result was obtained for a map with both drift and diffusion by deriving exact analytical solutions for the transport coefficients [10,16]. Further maps modelling chemical reaction–diffusion [15] and anomalous diffusion [34] also yielded fractal transport coefficients. Recent work aimed at physically more realistic models like

(Hamiltonian) particle billiards, for which computer simulations yielded transport coefficients that are non-monotonic under parameter variation [32]. Reference [33] contains a summary of this line of research.

These results asked for a more detailed characterization of the 'fractality' of transport coefficients. A first attempt in this direction was reported by Klages and Klauß [31], who used standard techniques from the theory of fractal dimensions for characterizing the drift and diffusion coefficients of the deterministic random walk $S^n(x)$ generated by the map studied in [16]:

$$S : \mathbb{R} \to \mathbb{R},\ S(x) = \begin{cases} ax + b & \text{if } -\frac{1}{2} \leqslant x < \frac{1}{2}, \\ S(x-k) + k & \text{if } -\frac{1}{2} \leqslant x - k < \frac{1}{2},\ k \in \mathbb{Z}. \end{cases} \tag{1}$$

They numerically computed a non-integer box counting dimension for these curves which varied with the parameter interval, leading to the notion of a 'fractal fractal dimension'. These results were questioned by Koza [35], who computed the oscillation of these graphs at specific Markov partition parameter values. His work suggested a dimensionality of one by conjecturing that there exist non-trivial logarithmic corrections to the usual power law behaviour in the oscillation.

This research reveals the need to study the parameter dependence of transport coefficients in a rigorous mathematical setting, which can be formulated as follows: given a parametrized family of chaotic dynamical systems $T_\lambda : I \to I$ on an interval I with unique invariant physical measures μ_λ together with a family of sufficiently regular observables $\psi_\lambda : I \to \mathbb{R}$ one has, under suitable mixing assumptions on the systems (T_λ, μ_λ), a law of large numbers and a central limit theorem for the partial sum processes $S_{\lambda,n}(x) = \sum_{k=0}^{n-1} \psi_\lambda(T_\lambda^k x)$, namely,

$$\lim_{n\to\infty} n^{-1} S_{\lambda,n} = J(\lambda) := \int_I \psi_\lambda(x)\, \mathrm{d}\mu_\lambda(x) \qquad \text{for } \mu_\lambda\text{-a.e. } x$$

(Birkhoff's ergodic theorem) and

$$\mathcal{L}(n^{-\frac{1}{2}} S_{\lambda,n}) \Rightarrow \mathcal{N}(0, 2D(\lambda)),$$

where $D(\lambda) := \lim_{n\to\infty} \frac{1}{2n} \int_I (\sum_{k=0}^{n-1} (\psi_\lambda(T_\lambda^k x) - J(\lambda)))^2\, \mathrm{d}\mu_\lambda(x)$, see, e.g. [19,22,37,39]. For the map $T_\lambda(x) = ax + b \bmod (\mathbb{Z} - \frac{1}{2})$ and the observable $\psi_\lambda(x) = (a-1)x + b$ that are studied in section 4, the process $S_{\lambda,n}$ is just the deterministic random walk S^n from (1), and $J(\lambda)$ and $D(\lambda)$ are the drift and diffusion coefficient of this walk, respectively.

There are a few rigorous results in the literature describing the dependence of μ_λ and of quantities like $J(\lambda)$ for various classes of systems. Without going into the details they can be summarized as follows: if the maps T_λ and the observables ψ_λ depend smoothly on λ and if the topological conjugacy class of T_λ is not changed when λ is varied, then μ_λ (and hence $J(\lambda)$) depends differentiably on λ [4,9,11,20,40,42] (see [21,41] for related results). If the topological class changes, quantities like $J(\lambda)$ may behave less regularly and have a modulus of continuity not better than $|\delta\lambda \cdot \log|\delta\lambda||$, even for very simple maps T_λ like symmetric tent maps [3]. On the other hand, this is always an upper bound for the modulus of continuity when the Perron–Frobenius operator of the system (acting on a suitable space of 'regular' densities) has a spectral gap [23,25].

The goal of this paper is to explicitly relate these mathematical results to transport coefficients. We do so by rigorously proving continuity properties of $J(\lambda)$ and $D(\lambda)$ under parameter variation for certain classes of deterministic maps. In section 2 we give a general estimate for families of dynamics, which can be described in terms of Perron–Frobenius operators with exponential mixing properties. The applicability of these general results to piecewise C^2 expanding interval maps and in particular to the class of piecewise linear maps

discussed in [16, 27, 28, 30, 33] is checked in section 3. The main result is that $D(\lambda)$ has a modulus of continuity of order at most $\mathcal{O}(|\delta\lambda| \cdot (\log|\delta\lambda|)^2)$, i.e. $D(\lambda)$ is Lipschitz continuous up to quadratic logarithmic corrections (proposition 1 in conjunction with lemmas 2 and 3). In section 4, proposition 2, we specialize the general results for transport coefficients to the particular system (1) and provide more precise estimates when the parameter a is an integer (proposition 3). From this we can conclude that, for integer a, the δ-oscillation of the diffusion coefficient as a function of the parameter b is *not* of order $\mathcal{O}(\delta)$ at Lebesgue-almost every b. Although this is weaker than Tricot's [43] definition of fractality[3], we consider this to be a mathematical justification qualifying the parameter dependence of the diffusion coefficient as 'fractal'.

Our analytical findings are quantified by numerical computations in section 5, for which we use exact analytical solutions of the transport coefficients [16]. Particularly, we numerically analyse local variations of these properties, which goes beyond what can be proven rigorously mathematically. Our work corrects and amends previous results reported in [31, 35].

2. The general setting

Let I be a compact interval, m normalized Lebesgue measure on I, L^1_m the space of Lebesgue-integrable functions from I to $\mathbb{R}$, and $\mathrm{BV} \subset L^1_m$ the space of L^1_m-equivalence classes of functions of bounded variation. We use the following simplified notation for the two corresponding norms:

$$|f|_1 := \int |f|\,\mathrm{d}m, \qquad \|f\| := \mathrm{Var}(f), \tag{2}$$

where

$$\mathrm{Var}(f) := \sup\left\{\int f\varphi'\,\mathrm{d}m : \varphi \in C^1(\mathbb{R},\mathbb{R}), |\varphi|_\infty \leqslant 1\right\} \tag{3}$$

is the variation of f as a function from $\mathbb{R} \to \mathbb{R}$ (i.e. extended by $f \equiv 0$ on $\mathbb{R} \setminus I$). If f is differentiable as a function from $\mathbb{R} \to \mathbb{R}$ integration by parts shows easily that $\mathrm{Var}(f) = \int |f'|\,\mathrm{d}m$. Var is obviously a semi-norm, and as $|f|_1 \leqslant |f|_\infty \leqslant \frac{1}{2}\mathrm{Var}(f)$, it is actually a norm. This and more details on functions of bounded variation can be found in [26, section 2.3]. The monograph [2] is a comprehensive reference for most of the background material needed in this section.

We consider a family $\mathcal{T}$ of non-singular maps $T : I \to I$. *Non-singular* means that the *Perron–Frobenius operator* $P_T : L^1_m \to L^1_m$ is well defined, i.e.

$$\int P_T f \cdot g\,\mathrm{d}m = \int f \cdot g \circ T\,\mathrm{d}m \qquad (f \in L^1_m, g \in L^\infty_m). \tag{4}$$

By definition, $|P_T|_1 = 1$ for all $T \in \mathcal{T}$, and we assume

Hypothesis 1. $C_1 := \sup\{\|P^n_T\| : T \in \mathcal{T}, n \in \mathbb{N}\} < \infty$.

Our main assumption is that the maps in $\mathcal{T}$ are *uniformly exponentially mixing* in the following sense:

Hypothesis 2. *Each $T \in \mathcal{T}$ has a unique invariant probability density $h_T \in \mathrm{BV}$ (so $P_T h_T = h_T$), and there are constants $\gamma \in (0,1)$ and $C_2 > 0$ such that, for all $T \in \mathcal{T}$,*

$$|P^n_T f|_1 \leqslant C_2\gamma^n \mathrm{Var}(f) \quad \text{for all } f \in \mathrm{BV} \text{ with } \int f\,\mathrm{d}m = 0 \text{ and for all } n \in \mathbb{N}. \tag{5}$$

[3] More precisely, our proposition 3 implies that $\limsup_{\delta\to 0}\delta^{-1}\mathrm{osc}_\delta(b) = +\infty$ for Lebesgue-a.e. b, whereas Tricot [43, 12.2] requires *uniform convergence* to $+\infty$ for *all* b.

Observe the following consequences of hypotheses 1 and 2:

$$|P_T^n f - h_T|_1 \leqslant C_2\gamma^n(\mathrm{Var}(f) + 2C_1) \qquad \text{for all probability densities } f \in \mathrm{BV} \tag{6}$$

and

$$\mathrm{Var}(h_T) \leqslant 2C_1 \qquad (T \in \mathcal{T}). \tag{7}$$

Indeed, $|P_T^n f - h_T|_1 = |P_T^n(f - h_T)|_1 \leqslant C_2\gamma^n(\mathrm{Var}(f) + \mathrm{Var}(h_T)) \to 0$ as $n \to \infty$ by hypothesis 2 for each probability density $f \in \mathrm{BV}$. In particular $|P_T^n 1 - h_T|_1 \to 0$ so that $\mathrm{Var}(h_T) \leqslant \sup_n \mathrm{Var}(P_T^n 1) \leqslant C_1\mathrm{Var}(1) = 2C_1$ by (3) and hypothesis 1. Now (6) follows at once.

Since it is our goal to investigate the dependence of various dynamical quantities as functions of $T \in \mathcal{T}$, we need to introduce a distance on $\mathcal{T}$. At this stage the following one, which was already considered in [23], is most appropriate. It measures the distance between two maps T_1 and T_2 from $\mathcal{T}$ in terms of a suitable norm of $P_{T_1} - P_{T_2}$:

$$|||P_{T_1} - P_{T_2}||| := \sup\left\{|P_{T_1} f - P_{T_2} f|_1 : f \in \mathrm{BV}, \|f\| \leqslant 1\right\}. \tag{8}$$

This distance can be controlled in terms of a more 'hands-on' distance between the graphs of the maps:

$$\begin{aligned} &d(T_1, T_2) := \inf\{\epsilon > 0 : \exists I_\epsilon \subseteq I \text{ and } \exists \text{ a diffeomorphism } \sigma : I \to I \text{ s.th.} \\ &m(I \setminus I_\epsilon) < \epsilon,\ T_1|_{I_\epsilon} = T_2 \circ \sigma|_{I_\epsilon}, \text{ and} \\ &\forall x \in I_\epsilon :\ |\sigma(x) - x| < \epsilon, |1/\sigma'(x) - 1| < \epsilon\}. \end{aligned} \tag{9}$$

Namely (see [23, lemma 13]),

$$|||P_{T_1} - P_{T_2}||| \leqslant 12 \cdot d(T_1, T_2). \tag{10}$$

Now, as a warm-up exercise, we can prove the following estimate: for $k \geqslant 0$ let

$$\ell_k : (0, \infty) \to (0, \infty), \qquad \ell_k(u) := u \cdot (1 + |\log u|)^k. \tag{11}$$

Lemma 1 (see also [23, proposition 7]). *There exist constants $K_1', K_1 > 0$ such that*

$$|h_{T_1} - h_{T_2}|_1 \leqslant K_1' \cdot \ell_1(|||P_{T_1} - P_{T_2}|||) \leqslant K_1 \cdot \ell_1(d(T_1, T_2)) \qquad (T_1, T_2 \in \mathcal{T}) \tag{12}$$

Proof. Let $\tilde{\eta} := |||P_{T_1} - P_{T_2}|||$, assume without loss of generality that $\tilde{\eta} < 1$, and fix $N \in \mathbb{N}$. For $f \in \mathrm{BV}$,

$$\begin{aligned} |P_{T_1}^N f - P_{T_2}^N f|_1 &\leqslant \sum_{k=0}^{N-1} |P_{T_1}^{N-k-1}(P_{T_1} - P_{T_2})P_{T_2}^k f|_1 \leqslant \sum_{k=0}^{N-1} |(P_{T_1} - P_{T_2})P_{T_2}^k f|_1 \\ &\leqslant \sum_{k=0}^{N-1} |||P_{T_1} - P_{T_2}||| \cdot \|P_{T_2}^k f\| \leqslant \sum_{k=0}^{N-1} \tilde{\eta}\, C_1 \|f\| \leqslant C_1 N \tilde{\eta} \|f\|, \end{aligned} \tag{13}$$

where we used hypothesis 1. Hence,

$$\begin{aligned} |h_{T_1} - h_{T_2}|_1 &\leqslant |P_{T_1}^N 1 - P_{T_2}^N 1|_1 + |P_{T_1}^N(1 - h_{T_1})|_1 + |P_{T_2}^N(1 - h_{T_2})|_1 \\ &\leqslant 2C_1 N\tilde{\eta} + 2 \cdot C_2\gamma^N(2 + 2C_1), \end{aligned}$$

where we used hypothesis 2 and estimate (7). For $N = \lceil \log\tilde{\eta}/\log\gamma \rceil$, this is in the form of the first inequality of (12). □

Remark 1. Even if $\mathcal{T}$ is a family of piecewise linear maps and if T_1 has the Markov property, this estimate can generally not be improved. Examples for this fact within the family of symmetric mixing tent maps are provided in [3, 38].

Suppose now that to each $T \in \mathcal{T}$ there is associated an 'observable' $\psi_T : I \to \mathbb{R}$. We make the following assumptions:

Hypothesis 3. $C_3 := \sup\{\mathrm{Var}(\psi_T) : T \in \mathcal{T}\} < \infty$

Hypothesis 4. *There is* $C_4 > 0$ *such that* $|\psi_{T_1} - \psi_{T_2}|_1 \leqslant C_4 d(T_1, T_2)$ *for all* $T_1, T_2 \in \mathcal{T}$.

Denote

$$J(T) := \int_I \psi_T h_T \, \mathrm{d}m. \tag{14}$$

Then we have immediately from (7) and lemma 1

Corollary 1. *There is some* $K_2 > 0$ *such that, for all* $T_1, T_2 \in \mathcal{T}$,

$$|J(T_1) - J(T_2)| \leqslant K_2 \cdot \ell_1(d(T_1, T_2)). \tag{15}$$

$J(T)$ is the 'drift' of the partial sum process

$$S_{T,n} := \sum_{k=0}^{n-1} \psi_T \circ T^k = n\, J(T) + \sum_{k=0}^{n-1} \hat{\psi}_T \circ T^k$$

under the invariant measure $h_T m$, where $\hat{\psi}_T = \psi_T - J(T)$. Observe that

$$\mathrm{Var}(\hat{\psi}_T) \leqslant 2C_3, \quad |\hat{\psi}_{T_1} - \hat{\psi}_{T_2}|_1 \leqslant 2C_4\, d(T_1, T_2) \qquad \text{for all } T, T_1, T_2 \in \mathcal{T}. \tag{16}$$

In view of hypothesis 2 we can also define the 'diffusion coefficient'[4] of this process:

$$\begin{aligned} D(T) &:= \lim_{n\to\infty} \frac{1}{2n} \int \left(\sum_{k=0}^{n-1} \hat{\psi}_T \circ T^k \right)^2 h_T \, \mathrm{d}m \\ &= \frac{1}{2} \int \hat{\psi}_T^2 \, h_T \, \mathrm{d}m + \sum_{n=1}^{\infty} \int \hat{\psi}_T \cdot \hat{\psi}_T \circ T^n \, h_T \, \mathrm{d}m \\ &= \frac{1}{2} \int \hat{\psi}_T^2 \, h_T \, \mathrm{d}m + \sum_{n=1}^{\infty} \int P_T^n(\hat{\psi}_T h_T)\, \hat{\psi}_T \, \mathrm{d}m. \end{aligned} \tag{17}$$

More precise information on the diffusive behaviour of $S_{T,n}$ is available: for each fixed $T \in \mathcal{T}$ we have the central limit theorem

$$\mathcal{L}(n^{-\frac{1}{2}}(S_{T,n} - nJ(T)) \Rightarrow \mathcal{N}(0, 2D(T)) \qquad \text{as } n \to \infty. \tag{18}$$

(If $D(T) = 0$ this means, as usual, weak convergence to the unit mass at 0.) For Lasota–Yorke type maps looked at in section 3 this follows immediately, e.g. from [19,22,39]. In the slightly more general setting of this section one has to invoke [37, Theorem 1.1][5].

Among physicists (17) is known as the Taylor–Green–Kubo formula for diffusion [33]. For the dependence of $D(T)$ on T we prove the following proposition.

Proposition 1. *There is some* $K_3 > 0$ *such that, for all* $T_1, T_2 \in \mathcal{T}$,

$$|D(T_1) - D(T_2)| \leqslant K_3 \cdot \ell_2(d(T_1, T_2)). \tag{19}$$

[4] This is the convention in the physics literature. In the mathematics literature one would rather call $2D(T)$ the diffusion coefficient.

[5] The assumptions of this theorem are stated in terms of the Koopman operator $\widehat{T}\phi := \phi \circ T$ and its dual $\widehat{T}^*$ on $L^2_{\mu_T}(I)$ where $\mu_T := h_T\, m$. Denote by $\mathcal{F}$ the Borel-σ-algebra on I and by $\mathbb{E}$ the expectation w.r.t. μ_T. As $h_T \widehat{T}^{*n}\phi = P_T^n(h_T\phi)$, the assumptions of [37] are easy to check: $\widehat{T}\widehat{T}^*\phi$ is $T^{-1}\mathcal{F}$-measurable and equals $\mathbb{E}(\phi|T^{-1}\mathcal{F})$, and $\sum_{n=0}^{\infty} \mathbb{E}|\widehat{T}^{*n}\hat{\psi}_T| \leqslant \sum_{n=0}^{\infty} C_2\gamma^n \mathrm{Var}(h_T\hat{\psi}_T) < \infty$ by hypothesis 2, so that $\sum_{n=0}^{\infty} |\mathbb{E}(\hat{\psi}_T \widehat{T}^n \hat{\psi}_T)| < \infty$ and $\sum_{n=0}^{\infty} \mathbb{E}(\widehat{T}^{*n}\hat{\psi}_T)$ converges absolutely and almost surely.

Proof. Observe first that $\mathrm{Var}(\psi h) \leqslant \mathrm{Var}(\psi)|h|_\infty + \mathrm{Var}(h)|\psi|_\infty \leqslant \mathrm{Var}(\psi)\mathrm{Var}(h)$ for all $\psi, h \in \mathrm{BV}$. It follows that, in view of (5),

$$\begin{aligned}\left|\int P_T^n(\hat\psi_T h_T)\,\hat\psi_T\,\mathrm{d}m\right| &\leqslant |P_T^n(\hat\psi_T h_T)|_1 \cdot |\hat\psi_T|_\infty \leqslant C_2C_3\gamma^n\mathrm{Var}(\hat\psi_T h_T)\\ &\leqslant 4C_1C_2C_3^2\gamma^n.\end{aligned} \tag{20}$$

Let $\tilde\eta = \|P_{T_1} - P_{T_2}\|$ as before, denote $\eta := d(T_1, T_2)$ (so that $\tilde\eta \leqslant 12\eta$), and fix $N = \lceil \log\eta/\log\gamma\rceil$. For all $T_1, T_2 \in \mathcal{T}$, equation (20) implies

$$\sum_{n=N}^{\infty}\left|\int P_{T_1}^n(\hat\psi_{T_1}h_{T_1})\,\hat\psi_{T_1}\,\mathrm{d}m - \int P_{T_2}^n(\hat\psi_{T_2}h_{T_2})\,\hat\psi_{T_2}\,\mathrm{d}m\right| \leqslant \frac{8C_1C_2C_3^2}{1-\gamma}\eta. \tag{21}$$

For $0 \leqslant n < N$ we use a different estimate. We decompose

$$\int P_{T_1}^n(\hat\psi_{T_1}h_{T_1})\,\hat\psi_{T_1}\,\mathrm{d}m - \int P_{T_2}^n(\hat\psi_{T_2}h_{T_2})\,\hat\psi_{T_2}\,\mathrm{d}m = \Delta_1^n + \Delta_2^n + \Delta_3^n + \Delta_4^n, \tag{22}$$

where

$$\begin{aligned}|\Delta_1^n| &:= \left|\int P_{T_1}^n(\hat\psi_{T_1}h_{T_1})\,(\hat\psi_{T_1} - \hat\psi_{T_2})\,\mathrm{d}m\right| \leqslant \frac{1}{2}\mathrm{Var}(P_{T_1}^n(\hat\psi_{T_1}h_{T_1}))\,|\hat\psi_{T_1} - \hat\psi_{T_2}|_1\\ &\leqslant 4C_1^2C_3C_4\eta\end{aligned} \tag{23}$$

and

$$\begin{aligned}|\Delta_2^n| &:= \left|\int (P_{T_1}^n - P_{T_2}^n)(\hat\psi_{T_1}h_{T_1})\,\hat\psi_{T_2}\,\mathrm{d}m\right| \leqslant |||P_{T_1}^n - P_{T_2}^n|||\mathrm{Var}(\hat\psi_{T_1}h_{T_1})|\hat\psi_{T_2}|_\infty\\ &\leqslant 4C_1C_3^2|||P_{T_1}^n - P_{T_2}^n||| \leqslant 4(C_1C_3)^2n\tilde\eta,\end{aligned} \tag{24}$$

where the last inequality follows from equation (13). Next,

$$\begin{aligned}|\Delta_3^n| &:= \left|\int P_{T_2}^n\big((\hat\psi_{T_1} - \hat\psi_{T_2})h_{T_1}\big)\hat\psi_{T_2}\,\mathrm{d}m\right| \leqslant |\hat\psi_{T_1} - \hat\psi_{T_2}|_1\,|h_{T_1}|_\infty\,|\hat\psi_{T_2}|_\infty\\ &\leqslant 4C_1C_3C_4\eta\end{aligned} \tag{25}$$

and

$$\begin{aligned}|\Delta_4^n| &:= \left|\int P_{T_2}^n\big(\hat\psi_{T_2}(h_{T_1} - h_{T_2})\big)\hat\psi_{T_2}\,\mathrm{d}m\right| \leqslant |h_{T_1} - h_{T_2}|_1\,|\hat\psi_{T_2}|_\infty\,|\hat\psi_{T_2}|_\infty\\ &\leqslant C_3^2K_1\ell_1(\tilde\eta)\end{aligned} \tag{26}$$

(One may note that the constants $C_1, \ldots, C_4$ originate from hypotheses 1 to 4, respectively.)

From (23)–(26) we see that

$$|\Delta_1^n| + |\Delta_2^n| + |\Delta_3^n| + |\Delta_4^n| \leqslant \tilde K(n\eta + \ell_1(\eta)) \tag{27}$$

for some constant $\tilde K > 0$. Hence, in view of (21) and the choice of N,

$$\begin{aligned}|D(T_1) - D(T_2)| &\leqslant \frac{8C_1C_2C_3^2}{1-\gamma}\eta + \tilde K\left(\frac{1}{2}\ell_1(\eta) + \sum_{n=1}^{N-1}(n\eta + \ell_1(\eta))\right)\\ &\leqslant \frac{8C_1C_2C_3^2}{1-\gamma}\eta + \tilde K\left(\frac{N^2}{2}\eta + N\ell_1(\eta)\right) \leqslant K_3\cdot\ell_2(\eta)\end{aligned} \tag{28}$$

for a suitable constant K_3. (The term $\frac{1}{2}\ell_1(\eta)$ is the $n = 0$ contribution.) □

Remark 2. Quite often slightly stronger forms of hypotheses 2 and 4 are satisfied, where the mixing assumption (5) is replaced by

$$\mathrm{Var}(P_T^n f) \leqslant C_2' \gamma^n \mathrm{Var}(f) \qquad \text{for all } f \in \mathrm{BV} \text{with } \int f \,\mathrm{d}m = 0 \text{ and } n \in \mathbb{N} \tag{29}$$

and the assumption on the T-dependence of ψ_T is strengthened to

$$\mathrm{Var}(\psi_{T_1} - \psi_{T_2}) \leqslant C_4' d(T_1, T_2) \qquad \text{for all } T_1, T_2 \in \mathcal{T}. \tag{30}$$

An inspection of the above estimates shows that $|\Delta_1^n| \leqslant 4C_1C_2'C_3C_4\gamma^n\eta$ and $|\Delta_2^n| \leqslant 4C_1C_2'C_3(1-\gamma)^{-1}\eta$ if (29) is assumed. If additionally (30) is assumed, then $|\Delta_3^n|$ can be estimated as follows: let $\alpha := \int_I (\hat\psi_{T_1} - \hat\psi_{T_2}) h_{T_1} \,\mathrm{d}m$. Then

$$\begin{aligned} |\Delta_3^n| &= \left| \int P_{T_2}^n ((\hat\psi_{T_1} - \hat\psi_{T_2}) h_{T_1} - \alpha h_{T_2}) \hat\psi_{T_2} \,\mathrm{d}m \right| \leqslant C_2 C_3 \, \gamma^n \, \mathrm{Var}((\hat\psi_{T_1} - \hat\psi_{T_2}) h_{T_1} - \alpha h_{T_2}) \\ &\leqslant C_1 C_2 C_3 C_4' (4+2) \gamma^n \eta \,. \end{aligned} \tag{31}$$

Hence, $\sum_{n=0}^{N-1} |\Delta_1^n| + |\Delta_3^n| = \mathcal{O}(\eta)$ uniformly in N and $\sum_{n=0}^{N-1} |\Delta_2^n| = \mathcal{O}(N\eta) = \mathcal{O}(\ell_1(\eta))$. But we see no way, in general, to bound the Δ_4^n-terms in a similar way. However, for particular families of maps T and observables ψ_T sharper estimates of $|\Delta_2^n|$ and $|\Delta_4^n|$ are available. In section 4.2 an example is studied where all $T \in \mathcal{T}$ are conjugate by rotations and leave Lebesgue measure on I invariant.

Remark 3. Instead in terms of $\ell_2(d(T_1, T_2))$ the estimate in proposition 1 can be given more directly in terms of $\ell_2(|||P_{T_1} - P_{T_2}|||) + \ell_2(|\psi_{T_1} - \psi_{T_2}|_1)$. In that way it applies also to stochastic systems, e.g. to stochastic perturbations of many Lasota–Yorke type maps, see e.g. [7, 23].

3. Checking hypotheses 1 and 2

3.1. General piecewise expanding maps

In this subsection we show how the general hypotheses 1 and 2 can be verified in the more particular setting when $\mathcal{T}$ is a parametrized family of piecewise twice continuously differentiable and expanding interval maps. So, from now on, we look at the following setting:

$\Lambda \subset \mathbb{R}^d$ is a compact parameter space, $\mathcal{T} = \{T_\lambda : \lambda \in \Lambda\}$, and (T1)

there is some $L > 0$ such that $d(T_{\lambda_1}, T_{\lambda_2}) \leqslant L|\lambda_1 - \lambda_2|$ for all $\lambda_1, \lambda_2 \in \Lambda$. (T2)

We start with an abstract result which reduces hypothesis 2 essentially to a uniform Lasota–Yorke type inequality plus a mixing assumption.

Lemma 2. *Assume (T1) and (T2). Then hypotheses 1 and 2 (and even the stronger property (29)) are valid if the transformations $T \in \mathcal{T}$ are mixing and satisfy a uniform Lasota–Yorke type inequality: there are constants $C_5, C_6 > 0$ and $\alpha \in (0,1)$ such that*

$$\mathrm{Var}(P_T^n f) \leqslant C_5 \alpha^n \mathrm{Var}(f) + C_6 |f|_1 \qquad \text{for all } T \in \mathcal{T}, \quad n \in \mathbb{N}, \quad f \in \mathrm{BV}. \tag{LY}$$

By 'mixing' we mean that each T has a unique invariant density and that it is weakly mixing w.r.t. this density. Bowen [8] gave sufficient conditions for piecewise monotonic maps to be mixing, see section 3.2 for details.

Proof. As $|P_T|_1 = 1$ and $|f|_1 \leqslant \frac{1}{2}\|f\|$, it is straightforward to check that hypothesis 1 holds with $C_1 = C_5 + \frac{1}{2}C_6$.

We turn to hypothesis 2. Note first that, because of (T1) and (T2), it suffices to show that for each $\lambda \in \Lambda$ there are $\delta(\lambda) > 0$, $C_2(\lambda) > 0$ and $\gamma(\lambda) \in (0, 1)$ such that (29) (and hence also (5)) holds with these constants for all T_{λ_1} with $|\lambda_1 - \lambda| < \delta(\lambda)$. But this is guaranteed by corollary 2 and remark 1c in [25], because property (LY) guarantees that each $P_T : \mathrm{BV} \to \mathrm{BV}$ is quasicompact with spectral radius 1, and the mixing assumption implies that 1 is the only eigenvalue of modulus 1 and that it is a simple eigenvalue. The reader may consult the monograph [2] for a comprehensive discussion of this circle of ideas. □

Our next task is to give sufficient conditions for (LY) and for the mixing property. To this end we specialize further and assume from now on that our maps are piecewise expanding (PE) maps in the following sense:

For each $\lambda \in \Lambda$ there is a finite partition $(I_\lambda^1, \ldots, I_\lambda^{N_\lambda})$ of I into sub-intervals such that all $T_\lambda|_{I_\lambda^j}$ are monotone, C^2, and $\kappa_\lambda := \inf |T_\lambda'| > 2$. (PE)

Already in [36] it was proved that each individual (PE)-map (even if $1 < \kappa_\lambda \leqslant 2$) satisfies (LY) with constants C_5, C_6, α depending on the map. For parametrized families of maps one can generally find uniform constants, but there are counterexamples where this is not possible [5, 6, 23]. Under the above assumption $\kappa_\lambda > 2$ for all $\lambda \in \Lambda$ one can, however, give simple sufficient conditions ensuring the uniform LY-inequality. The proof in [36] (see also [26, proposition 2.1]) shows

$$\mathrm{Var}(P_{T_\lambda} f) \leqslant \frac{2}{\kappa_\lambda}\mathrm{Var}(f) + (E_\lambda + F_\lambda)|f|_1, \tag{32}$$

where

$$E_\lambda = \sup_x \left|\left(\frac{1}{T_\lambda'(x)}\right)'\right|, \qquad F_\lambda = \frac{2}{\kappa_\lambda \min_j |I_\lambda^j|}. \tag{33}$$

From this (LY) follows with $\alpha = \sup_{\lambda\in\Lambda} 2/\kappa_\lambda$, $C_5 = 1$, and $C_6 = \sup_\lambda(\kappa_\lambda/(\kappa_\lambda - 2))(E_\lambda + F_\lambda)$ provided this supremum is finite and $\alpha < 1$.

For maps $T_\lambda : I \to I$ verifying (PE), Bowen [8] gave the following sufficient condition to be mixing:

$$T = T_\lambda \text{ is mixing, if } \sup_{n>0} \mu_T(T^n U) = 1 \qquad \text{for each interval } U = I_\lambda^1, \ldots, I_\lambda^{N_\lambda}. \tag{M}$$

3.2. Piecewise linear modulo 1 maps

We now look at a particular model dealt with in [13,14,17,18] from a mathematical perspective and in [16,27,28,30,33] from a physics point of view. Let $I = [-\frac{1}{2}, \frac{1}{2}]$, $\Lambda = [a_0, a_1] \times [-\frac{1}{2}, \frac{1}{2}]$ for some constants $2 < a_0 < a_1$, and for $\lambda = (a, b) \in \Lambda$ consider

$$T_\lambda(x) = ax + b \bmod (\mathbb{Z} - \tfrac{1}{2}). \tag{34}$$

Hofbauer [17] showed that these maps have always a unique invariant probability density[6]. The parameter space Λ is obviously compact, so (T1) is satisfied. In order to check assumption (T2) on the Lipschitz dependence of the maps on the parameters we estimate $d(T_{a,b}, T_{a',b'})$ for $(a, b), (a', b') \in \Lambda$.

For given $(a, b) \in \Lambda$ let $p = \lceil((a+1)/2) - b\rceil - 1$ and $q = \lceil((a+1)/2) + b\rceil$ where $\lceil x\rceil$ denotes the smallest integer greater or equal to x. Let $A_{-p} := -\frac{1}{2}$, $A_q := \frac{1}{2}$ and denote by

[6] Indeed, Hofbauer shows this for the maximal measure of such maps, but since these maps have constant slope, the maximal measures are just the absolutely continuous ones. For numerical results on the probability densities associated with these measures and how they change under parameter variation see [28].

$A_k := a^{-1}(k - \frac{1}{2} - b)$ $(k = -p+1, \dots, q-1)$ the discontinuities of $T_{a,b}$ in $(-\frac{1}{2}, \frac{1}{2})$. The intervals (A_k, A_{k+1}) $(k = -p, \dots, q-1)$ are the maximal monotonicity intervals of $T_{(a,b)}$. Consider also another $(a', b') \in \Lambda$ and suppose w.l.o.g. that $a' \leqslant a$. With the same p and q as before define $A'_k := a^{-1}(k - \frac{1}{2} - b)$ $(k = -p+1, \dots, q-1)$. Although not all of the A'_k belong necessarily to $(-\frac{1}{2}, \frac{1}{2})$, all these points are discontinuities of $T_{(a',b')}$ when the definition (34) of this map is extended to all $x \in \mathbb{R}$.

Consider the linear map $L : \mathbb{R} \to \mathbb{R}$, $L(x) = (ax+b-b')/a'$ and observe that $L(A_k) = A'_k$ $(-p < k < q)$ and $a'L(x) + b' = ax + b$ for all $x \in \mathbb{R}$. Let $[u, v] := I \cap L^{-1}(I)$ and $I_0 := [u + \delta, v - \delta]$ for some arbitrarily small $\delta > 0$. Define $\sigma : I \to I$ by $\sigma(x) = L(x)$ if $x \in I_0$ and extend σ to a diffeomorphism of I. Then

- $m(I \setminus I_0) \leqslant (1 - \frac{a'}{a}) + |b' - b|/a + 2\delta \leqslant a_0^{-1}(|a' - a| + |b' - b|) + 2\delta$,
- $|\sigma(x) - x| = |L(x) - x| \leqslant \frac{1}{2}|\frac{a}{a'} - 1| + |b - b'|/a' \leqslant a_0^{-1}(\frac{1}{2}|a - a'| + |b - b'|)$ for all $x \in I_0$,
- $|1/\sigma'(x) - 1| = |\frac{a'}{a} - 1| \leqslant a_0^{-1}|a' - a|$ for all $x \in I_0$,
- $T_{(a',b')}(\sigma(x)) = T_{(a,b)}(x)$ for all $x \in I_0$.

So

$$d(T_{(a,b)}, T_{(a',b')}) \leqslant a_0^{-1}(|a - a'| + |b - b'|) \tag{35}$$

as $\delta > 0$ could be chosen arbitrarily small. Hence (T2) is satisfied.

Lemma 3. *The maps T_λ $(\lambda = (a, b) \in \Lambda)$ satisfy hypotheses 1 and 2 and also property (29).*

Proof. As all T_λ are piecewise expanding in the sense of (PE), the discussion in section 3.1 shows that (LY) is satisfied with $C_5 = 1$ and $\alpha = (2/a_0) < 1$. Hence the claim of the lemma follows from lemma 2 once we have shown that the T_λ are mixing. To this end we verify Bowen's mixing condition (M): if U is a non-empty open interval some iterate $T_\lambda^n U$ must contain one of the monotonicity intervals (A_{k-1}, A_k), $k \in \{-p+1, \dots, q-1\}$, introduced above, that are mapped onto $(-\frac{1}{2}, \frac{1}{2})$. (Otherwise the $T_\lambda^n U \subseteq I$ would contain intervals V_n of length $|V_n| \geqslant (\frac{a_0}{2})^n |U|$ growing without limit when $n \to \infty$.) □

Remark 4. For the maps T_λ from lemma 3, hypothesis 2 can be verified by direct elementary calculations along the lines of [24] without invoking the perturbation theory of [25]. One obtains the explicit estimate

$$\mathrm{Var}(P_{T_\lambda} f) \leqslant \frac{2}{a}\mathrm{Var}(f) + 2\left|\int f \,\mathrm{d}m\right| \qquad \text{for all } f \in \mathrm{BV} \text{ and } n \in \mathbb{N}.$$

Alternatively, the mixing property follows also from [13, theorem 5.1].

4. Transport coefficients

We apply the results of the previous sections to determine transport coefficients of the deterministic random walks generated by the maps $T_\lambda = T_{a,b}$ from section 3.2. The random walks in question are $S_{\lambda,n} = \sum_{k=0}^{n-1} \psi_\lambda \circ T_\lambda^k$ with

$$\psi_\lambda(x) = (a - 1)x + b. \tag{36}$$

It is an easy exercise to see that hypotheses 3 and 4 as well as their strengthening (30) are satisfied: $\mathrm{Var}(\psi_{a,b}) \leqslant 2(a - 1 + |b|) < 2(a_1 + 1) =: C_3$ and $|\psi_{a,b} - \psi_{a',b'}|_1 \leqslant \frac{1}{2}\mathrm{Var}(\psi_{a,b} - \psi_{a',b'}) \leqslant |a - a'| + |b - b'|$.

For later use we note that the maps $T_{a,b}$ and $T_{a,-b}$ are conjugate in the sense that $T_{a,b}(-x) = -T_{a,-b}(x)$, in particular $h_{a,-b}(-x)$ is also an invariant density for $T_{a,b}$ and, by uniqueness, $h_{a,b}(x) = h_{a,-b}(-x)$.

We first note the following explicit form of the drift:

$$J(\lambda) = J(a,b) := \int \psi_{a,b} h_{a,b}\,\mathrm{d}m = b + (a-1)\int x h_{a,b}(x)\,\mathrm{d}x. \tag{37}$$

As noted above, $h_{a,0}(x) = h_{a,0}(-x)$. Hence $J(a,0) = 0$.

4.1. Upper bounds for the modulus of continuity of the drift and the diffusion coefficient

Now we apply corollary 1 and proposition 1 to our setting. In view of the estimate in (35) this yields:

Proposition 2. *For the family of maps $(T_\lambda : \lambda \in \Lambda)$ defined in (34), there are constants $K_3, K_4 > 0$ such that the drift $J(\lambda) := J(T_\lambda)$ and the diffusion coefficient $D(\lambda) := D(T_\lambda)$, $\lambda = (a,b)$, satisfy*

$$|J(\lambda) - J(\lambda')| \leqslant K_3 \cdot |\lambda - \lambda'| \cdot (1 + |\log|\lambda - \lambda'||), \qquad (\lambda, \lambda' \in \Lambda), \tag{38}$$

$$|D(\lambda) - D(\lambda')| \leqslant K_4 \cdot |\lambda - \lambda'| \cdot (1 + |\log|\lambda - \lambda'||)^2 \qquad (\lambda, \lambda' \in \Lambda). \tag{39}$$

Corollary 2.

(a) The graph of $D : \Lambda \to \mathbb{R}$ has box- and Hausdorff-dimension 2.

(b) For each $b \in \mathbb{R}$, the graph of $D_b : [a_0, a_1] \to \mathbb{R}$, $D_b(a) = D(a,b)$, has box- and Hausdorff-dimension 1.

(c) For each $a > 2$, the graph of $D_a : [-\frac{1}{2}, \frac{1}{2}] \to \mathbb{R}$, $D_a(b) = D(a,b)$, has box- and Hausdorff-dimension 1.

Proof. This is well known, but we include a proof for completeness. Denote by $\dim_B$ and $\dim_H$ the box and Hausdorff dimension, respectively. Obviously, $2 \leqslant \dim_H(\mathrm{graph}(D)) \leqslant \dim_B(\mathrm{graph}(D))$. So it remains to show that $\dim_B(\mathrm{graph}(D)) \leqslant 2$. To this end subdivide the rectangle Λ into little squares of equal diameter $\approx N^{-1}$. For each such square Q we have

$$\max\{D(\lambda) : \lambda \in Q\} - \min\{D(\lambda) : \lambda \in Q\} \leqslant K_4 N^{-1}(1 + \log N)^2. \tag{40}$$

Hence,

$$\dim_B(\mathrm{graph}(D)) \leqslant \limsup_{N\to\infty} \frac{\log(K_4 N^2 (1+\log N)^2)}{\log N} = 2. \tag{41}$$

The two other claims are proved in the same way. □

Remark 5. Corollary 2 has already been conjectured by Koza [35]. His conjecture was based on calculating the pointwise Minkowski–Bouligand dimension for algebraic Markov partition parameter values of this family of maps by using the exact solutions for drift and diffusion coefficient given in [16]. This led him to conclude that the oscillation [43] of $D(\lambda)$ is linear in the size of the subinterval multiplied with a logarithmic term, cp (20) of [35] with (40) above. The exponent of this logarithmic correction was found to be either one or two depending on the type of Markov partition.

4.2. A closer look at maps with integer slope

We finish this section with a closer look at the functions $D_a(b)$ when a is an integer larger than 2. In this case $T = T_{a,b}$ can be seen as an a-fold covering linear circle map, so it leaves Lebesgue measure invariant. Therefore $h_T = 1$ for all such T so that $J(a, b) = b$ and the estimates on $D_a(b)$ from the proof of proposition 1 simplify drastically. As some of these estimates can be replaced by more precise expansions, this leads to

Proposition 3. *For the family $T_{(a,b)}$ defined in (34) with $a \in \{3, 4, 5, \ldots\}$ we have the following b-dependence of the diffusion coefficient of the generated random walk:*

(1) For even $a \geqslant 4$,

$$D_a(b') - D_a(0) = \mathcal{O}(b')$$

and

$$D_a(b') - D_a\left(\frac{1}{2}\right) = \frac{a-1}{2\log a}\left(b' - \frac{1}{2}\right)\left|\log\left|b' - \frac{1}{2}\right|\right| + \mathcal{O}\left(b' - \frac{1}{2}\right)$$

(2) For odd $a \geqslant 3$,

$$D_a(b') - D_a(0) = -\frac{a-1}{2\log a} b' |\log |b'|| + \mathcal{O}(b')$$

and

$$D_a(b') - D_a\left(\frac{1}{2}\right) = \frac{a-1}{4\log a}\left(b' - \frac{1}{2}\right)\left|\log\left|b' - \frac{1}{2}\right|\right| + \mathcal{O}\left(b' - \frac{1}{2}\right)$$

(3) If b is such that $\hat{b} = -\frac{1}{2} + (b/(a-1))$ is eventually periodic under $T_{a,0}$ and the periodic part of the orbit neither contains $-\frac{1}{2}$ nor $\hat{b}$, then there is a constant $C_{a,b}$ such that

$$D_a(b') - D_a(b) = C_{a,b}\,(b' - b)|\log|b' - b|| + \mathcal{O}(b' - b).$$

This generalizes assertions (1) and (2). See also remark 5.

(4) For fixed δ and random b drawn uniformly from $(0, \frac{1}{2})$ or from $(-\frac{1}{2}, 0)$,

$$D_a(b + \delta(a-1)) - D_a(b) = C_a\,\delta|\log|\delta||^{1/2}\,Z_\delta + \mathcal{O}(\delta) \text{ as } \delta \to 0$$

with a constant $C_a > 0$ and random variables Z_δ which all have the same distribution—a mixture of Gaussians as in (45) depending only on the fixed parameter a.

Remark 6. An immediate consequence of proposition 3(4) is that, for integer a, the graphs of $D_a : [-\frac{1}{2}, \frac{1}{2}] \to \mathbb{R}$ are fractal in a sense only slightly weaker than that of Tricot [43, section 12.2], although they have box- and Hausdorff-dimension 1: following [43], denote $\mathrm{osc}_\delta(b) := \sup_{b',b''} |D_a(b') - D_a(b'')|$ where the supremum extends over all $b', b'' \in [b - \delta, b + \delta]$. Then, for every $M > 0$, $m\{b \in [-\frac{1}{2}, \frac{1}{2}] : \delta^{-1}\mathrm{osc}_\delta(b) \leqslant M\} \to 0$ as $\delta \to 0$ so that $\limsup_{\delta\to 0} \delta^{-1}\mathrm{osc}_\delta(b) = +\infty$ for m-a.e. b.

Proof of proposition 3. Fix $a \in \{3, 4, 5, \ldots\}$. Then $J(a, b) = b$ for all b (see (37) and [16,33]), $\hat{\psi}_{a,b}(x) = (a-1)x =: \hat{\psi}_a(x)$ is the same for all b, and denoting $T_1 = T_{a,b}$ and $T_2 = T_{a,b'}$ we can replace estimates (22)–(27) by

$$\begin{aligned}\int_I P_{T_2}^n(\hat{\psi}_{T_2} h_{T_2})\,\hat{\psi}_{T_2}\,\mathrm{d}m - \int_I P_{T_1}^n(\hat{\psi}_{T_1} h_{T_1})\,\hat{\psi}_{T_1}\,\mathrm{d}m \\ = \int_I \hat{\psi}_a(x)\,\hat{\psi}_a(T_2^n x)\,\mathrm{d}x - \int_I \hat{\psi}_a(x)\,\hat{\psi}_a(T_1^n x)\,\mathrm{d}x \\ = (a-1)^2\left(\int_I x\,T_{a,b'}^n(x)\,\mathrm{d}x - \int_I x\,T_{a,b}^n(x)\,\mathrm{d}x\right).\end{aligned} \tag{42}$$

To evaluate this difference assume henceforth that $0 \leqslant b \leqslant \frac{1}{2}$. There is no loss in doing so, because $T_{a,b}$ and $T_{a,-b}$ are conjugate as obseved above. Define the 'rotation' $R : I \to I$ by $R(x) = x - (b/(a-1)) \bmod (\mathbb{Z} - \frac{1}{2})$. It conjugates $T_{a,b}$ to $T_{a,0}$, namely,

$$R(T_{a,0}^n x) = T_{a,b}^n(Rx) \qquad \text{for all } x \in I \text{ and } n \in \mathbb{N}.$$

Therefore, denoting $\hat{b} = -\frac{1}{2} + (b/(a-1))$ and $\chi_b(x) = 1_{[-\frac{1}{2},\hat{b})}(x) - (b/(a-1))$,

$$\int_I x\, T_{a,b}^n(x)\,\mathrm{d}x = \int_I R(x)\, R(T_{a,0}^n x)\,\mathrm{d}x = \int_I (x + \chi_b(x))(T_{a,0}^n x + \chi_b(T_{a,0}^n x))\,\mathrm{d}x$$

so that

$$\begin{aligned}\int_I x\, T_{a,b'}^n(x)\,\mathrm{d}x - \int_I x\, T_{a,b}^n(x)\,\mathrm{d}x \\ = \int_I P_{a,0}^n x \cdot (\chi_{b'}(x) - \chi_b(x))\,\mathrm{d}x + \int_I P_{a,0}^n(\chi_{b'} - \chi_b)(x) \cdot x\,\mathrm{d}x \\ + \int_I P_{a,0}^n \chi_{b'}(x) \cdot (\chi_{b'}(x) - \chi_b(x))\,\mathrm{d}x + \int_I P_{a,0}^n(\chi_{b'} - \chi_b)(x) \cdot \chi_b(x)\,\mathrm{d}x.\end{aligned}$$

As $\int_I \chi_{b'}(x)\,\mathrm{d}x = 0$ and $\mathrm{Var}(\chi_b) = 2$ for all b, the third term is of order $\mathcal{O}((2/a)^n|b' - b|)$ by lemma 3. As $P_{a,0}^n x = a^{-n}x$, the first term is at most of the same order. Therefore their sums over all n are of the order $\mathcal{O}(|b' - b|)$.

We turn to the two remaining terms. Their sum from $n = 0$ to ∞ is of the form

$$\sum_{n=0}^{\infty} \int_I P_{a,0}^n(\chi_{b'} - \chi_b)(x) \cdot g_b(x)\,\mathrm{d}x = \sum_{n=0}^{\infty} \int_I (\chi_{b'} - \chi_b)(x) \cdot g_b(T_{a,0}^n x)\,\mathrm{d}x \tag{43}$$

with $g_b(x) = x + \chi_b(x)$. Let $\delta = (b' - b)/(a-1)$. As $\int_{-1/2}^{1/2}(\chi_{b'} - \chi_b)(x)\,\mathrm{d}x = 0$, $\mathrm{Var}(\chi_{b'} - \chi_b) \leqslant 4$, and $|g_b| \leqslant 2$, the nth integral is of order $\mathcal{O}((2/a)^n)$. Hence the sum from $n = N_\delta := \lceil |\log|\delta||/\log a \rceil$ to ∞ is of order $|\delta|$, and it remains to estimate the sum from $n = 0$ to $N_\delta - 1$. (For these n we have $a^n|\delta| \leqslant 1$.) In the following we give precise approximations of this sum for various values of b. In all cases one should keep in mind the factor $(a-1)^2$ in (42) and the definition of $\delta = (b' - b)/(a-1)$.

Proof of assertions (1) and (2). We start with the special case $b = 0$ where we have $\chi_b = 0$, so $g_b(x) = x$. Then the nth term in the sum (43) evaluates to

$$\int_{-\frac{1}{2}}^{-\frac{1}{2}+\delta} T_{a,0}^n(x)\,\mathrm{d}x = \begin{cases}\delta \cdot (-\frac{1}{2} + \frac{1}{2}a^n\delta) & \text{if } a \text{ is odd,} \\ \delta \cdot (\frac{1}{2}a^n\delta) & \text{if } a \text{ is even and } n \geqslant 1\end{cases}$$

(Observe that $\int_I T_{(a,0)}^n(x)\,\mathrm{d}x = \int_I x\,\mathrm{d}x = 0$.) It follows that the sum from $n = 0$ to $N_\delta - 1$ in (43) is of the order $\mathcal{O}(|\delta|)$ if a is even and that it is $\delta \cdot (-(|\log|\delta||)/(2\log a) + \mathcal{O}(1))$ if a is odd.

Consider next the case $b = \frac{1}{2}$. As $b = -\frac{1}{2}$ gives rise to the same map, the we may assume w.l.o.g. that $-\frac{1}{2} < b' < b$, i.e. $\delta < 0$. If a is odd, then $T_{a,0}\hat{b} = \hat{b} + \frac{1}{2}$ and $T_{a,0}(\hat{b} + \frac{1}{2}) = \hat{b}$. Therefore

$$\begin{aligned}\int_I (\chi_{b'} - \chi_b)(x) \cdot g_b(T_{a,0}^n x)\,\mathrm{d}x &= \int_{\hat{b}}^{\hat{b}'} g_b(T_{a,0}^n x)\,\mathrm{d}x \\ &= \begin{cases}\int_0^\delta (\hat{b} + a^n t) + \chi_b(\hat{b} + a^n t)\,\mathrm{d}t & \text{if } n \text{ is even} \\ \int_0^\delta (\hat{b} + \dfrac{1}{2} + a^n t) + \chi_b(\hat{b} + \dfrac{1}{2} + a^n t)\,\mathrm{d}t & \text{if } n \text{ is odd}\end{cases} \\ &= \begin{cases}\delta\hat{b} + \dfrac{1}{2}a^n\delta^2 + \delta - \delta\dfrac{b}{a-1} & \text{if } n \text{ is even} \\ \delta(\hat{b} + \dfrac{1}{2}) + \dfrac{1}{2}a^n\delta^2 - \delta\dfrac{b}{a-1} & \text{if } n \text{ is odd}\end{cases}\end{aligned}$$

as long as $a^n|\delta| < b/(a-1)$, i.e. $n < \tilde{N}_\delta := N_\delta - (\log(a-1) - \log b)/\log a$. For the remaining n this identity needs to be modified by at most $|\delta|$. In any case,

$$\sum_{n=0}^{N_\delta-1} \int_I (\chi_{b'} - \chi_b)(x) \cdot g_b(T_{a,0}^n x)\,\mathrm{d}x = \frac{N_\delta}{2} \cdot \frac{\delta}{2} + \frac{N_\delta}{2} \cdot 0 + \mathcal{O}(\delta)$$
$$= \delta|\log|\delta|| \frac{1}{4\log a} + \mathcal{O}(\delta) \quad \text{for odd } a$$

with a constant in '$\mathcal{O}$' that depends on b and a but not on b'. If a is even, then $T_{a,0}\hat{b} = \hat{b}$, and following the argument for odd a and even n we obtain

$$\sum_{n=0}^{N_\delta-1} \int_I (\chi_{b'} - \chi_b)(x) \cdot g_b(T_{a,0}^n x)\,\mathrm{d}x = N_\delta \cdot \frac{\delta}{2} + \mathcal{O}(\delta) = \delta|\log|\delta|| \frac{1}{2\log a} + \mathcal{O}(\delta) \quad \text{for even } a.$$

Proof of assertions (3) and (4). We turn to more general parameters $0 < b < \frac{1}{2}$. We have to estimate

$$\begin{aligned} s(\delta) &:= \sum_{n=0}^{N_\delta-1} \int_I (\chi_{b'} - \chi_b)(x) \cdot g_b(T_{a,0}^n x)\,\mathrm{d}x = \sum_{n=0}^{N_\delta-1} \int_{\hat{b}}^{\hat{b}+\delta} g_b(T_{a,0}^n x)\,\mathrm{d}x \\ &= \sum_{n=0}^{N_\delta-1} \int_0^\delta g_b(T_{a,0}^n(\hat{b}) + a^n t)\,\mathrm{d}t. \end{aligned} \tag{44}$$

The details of this estimate depend strongly on the distributional properties of the orbit of $\hat{b}$ under $T_{a,0}$ and we discuss only two particular but important cases where the situation does not become too complicated. First we look at such b for which the orbit of $\hat{b}$ is eventually periodic but where the periodic part does neither contain $-\frac{1}{2}$ nor $\hat{b}$. (This is a countable dense set of parameters.) In this case one can argue as above for $b = \frac{1}{2}$ and odd a and show that

$$s(\delta) = \delta|\log|\delta|| \frac{C_b}{\log a} + \mathcal{O}(\delta)$$

with $C_b = \int_I g_b(x)\,\mathrm{d}\mu_b(x)$ where μ_b is the equidistribution on the periodic part of the orbit of $\hat{b}$. Exceptionally this may be zero, but typically it will not.

Next we look at Lebesgue typical points $\hat{b}$, i.e. at Lebesgue typical parameters b. For fixed δ we interpret $s(\delta)$ as a random variable where randomness is introduced via the parameter $b \in (0, \frac{1}{2})$. We are going to show that the random variables $\delta^{-1}N_\delta^{-\frac{1}{2}}s(\delta)$ converge in distribution to a mixture of Gaussians, i.e.

$$\mathcal{L}(\delta^{-1}N_\delta^{-\frac{1}{2}}s(\delta)) \Rightarrow \int_0^1 \mathcal{N}(0, \sigma_z^2)\,\mathrm{d}z \qquad \text{as } \delta \to 0 \tag{45}$$

with suitable variances $\sigma_z^2 > 0$ that depend on the fixed parameter a. This shows that approximately $s(\delta) = \delta|\log|\delta||^{\frac{1}{2}} Z$ with a random variable Z that is a mixture of Gaussians which depends only on the fixed integer parameter a.

We start by comparing $s(\delta)$ to $\sum_{n=0}^{N_\delta-1} \delta \cdot g_b(T_{a,0}^n \hat{b})$. As $g_b(x) = x + \chi_b(x)$, we can estimate the difference for the x- and the $\chi_b(x)$-contributions separately. For the x-contribution the difference is easily seen to be of order $\mathcal{O}(\delta)$. For the $\chi_b(x)$-contribution we estimate each of the last $L_\delta := \lceil (3/\log a)\log N_\delta \rceil$ terms of the sum by 2 thus getting a contribution of order

$\mathcal{O}(\delta \log N_\delta)$. For the remaining terms we note the following two estimates which are obvious from a short look at the graph of $T_{a,0}^n$:

$$m\{b \in I : \left|T_{a,0}^n(\hat{b}) - \left(-\frac{1}{2}\right)\right| < a^n\delta\},\ m\{b \in I : |T_{a,0}^n(\hat{b}) - \hat{b}| < a^n\delta\} \leqslant 4a^n\delta\ ,$$

so that

$$m\left\{b \in I : \exists n \in \{0, \ldots, N_\delta - L_\delta - 1\} \text{ s.th. } \left|T_{a,0}^n(\hat{b}) - \left(-\frac{1}{2}\right)\right| < a^n\delta \text{ or } |T_{a,0}^n(\hat{b}) - \hat{b}| < a^n\delta\right\}$$
$$\leqslant \frac{8}{a-1}a^{-L_\delta} = \frac{8}{a-1}N_\delta^{-3}.$$

It follows that $\sum_{n=0}^{N_\delta - L_\delta - 1}\int_0^\delta \chi_b(T_{a,0}^n(\hat{b}) + a^n t)\,\mathrm{d}t = \sum_{n=0}^{N_\delta - L_\delta - 1}\delta\chi_b(T_{a,0}^n\hat{b})$ except on a set of b of Lebesgue measure at most $(8/(a-1))N_\delta^{-3}$. Hence, observing that $L_\delta N_\delta^{-1} \to 0$ as $\delta \to 0$, the convergence in (45) will follow once we have proved

$$\mathcal{L}(Y_{N_\delta}) \Rightarrow \int_0^1 \mathcal{N}(0, \sigma_z^2)\,\mathrm{d}z \qquad \text{as } N_\delta \to \infty \tag{46}$$

where $Y_N(\hat{b}) := N^{-\frac{1}{2}}\sum_{n=0}^{N-1} g_b(T^n\hat{b})$ and b is uniformly distributed in the interval $(0, \frac{1}{2})$. (To ease the notation we abbreviate $T_{a,0}$ by T.) As the single contributions to the sum in Y_N depend on b via $T^n(\hat{b})$ and b itself, this is not the situation of the usual central limit theorem, so we treat the problem in two steps:

Step 1: For fixed $z \in (0, \frac{1}{2})$ consider $Y_N^z(\hat{b}) := N^{-\frac{1}{2}}\sum_{n=0}^{N-1} g_z(T^n\hat{b})$. It is a well known general fact that, for fixed z, the Y_N^z converge in distribution to some $\mathcal{N}(0, \sigma_{2z}^2)$ (see, e.g. [19,22,39])—except for the strict positivity of σ_{2z}^2. To prove this we use [39, lemma 6]: suppose for a contradiction that $\sigma_{2z}^2 = 0$. Then there is a function $\psi : I \to \mathbb{R}$ of bounded variation such that $g_z(x) = \psi(Tx) - \psi(x)$ for Lebesgue-a.e. x. Let $M := 2\sup|\psi|$. Then $\left|\sum_{k=0}^n g_z(T^k x)\right| \leqslant M$ for all n and a.a. x. Looking at suitable periodic orbits it is easy to see that there are $n = n_0$ and $x = x_0$ for which this sum is larger than $M + 2$. But then, as both T and also g_z are at least one-sided continuous, there is a small interval close to x_0 on which the same sum is larger than $M + 1$, which contradicts the above bound that holds for all n and Lebesgue-a.a. x.

Step 2: We need a number of preparations:

(i) Let $J := (-\frac{1}{2}, -\frac{1}{2} + (1/2(a-1)))$ be the interval through which $\hat{b}$ ranges when b is chosen randomly from $(0, \frac{1}{2})$.

(ii) Let $(r_j)_{j>0}$ be any sequence of natural numbers tending to infinity and such that $r_j \leqslant j^{\frac{1}{4}}$ for all j. For each j denote by $C_j \subset I = [-\frac{1}{2}, \frac{1}{2}]$ a set of points that subdivides I into a^{r_j} intervals of the same length which are mapped onto I bijectively by T^{r_j}. (If a is odd take $C_j := T^{-r_j}\{-\frac{1}{2}\}$, if a is even take $C_j := T^{-r_j}\{0\}$.) For $z \in C_j$ denote by $I_z^j = [z, z')$ the subinterval with left endpoint z.

(iii) $|Y_N^z - Y_N^z(T^{r_N}x)| = \mathcal{O}(r_N N^{-\frac{1}{2}}) = o(1)$ as $N \to \infty$, uniformly in x and z, because the two sums involved differ only by $2r_N$ terms.

(iv) Let $f(z) = (z + 1/2)(a - 1)$ so that $f(\hat{b}) = b$ and $\widehat{f(x)} = z$. Then

$$\begin{aligned}
m\{x \in I_z^N : |Y_N^{f(x)}(T^{r_N}x) - Y_N^{f(z)}(T^{r_N}x)| > r_N^{-1} + 2a^{-r_N}/(a-1)\} \\
&\leqslant m\{x \in I_z^N : |Y_N^{f(z')}(T^{r_N}x) - Y_N^{f(z)}(T^{r_N}x)| > r_N^{-1}\} \\
&= a^{-r_N} m\{x \in I : |Y_N^{f(z')}(x) - Y_N^{f(z)}(x)| > r_N^{-1}\} \\
&\leqslant a^{-r_N} r_N^2 \int_I (Y_N^{f(z')} - Y_N^{f(z)})^2 \, \mathrm{d}m \\
&= a^{-r_N} r_N^2 N^{-1} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \int_I (1_{I_z^N} \circ T^k - m(I_z^N))(1_{I_z^N} \circ T^l - m(I_z^N)) \, \mathrm{d}m \\
&\leqslant a^{-r_N} r_N^2 N^{-1} N \sum_{k=0}^{\infty} \left| \int_I (1_{I_z^N} \circ T^k - m(I_z^N))(1_{I_z^N} - m(I_z^N)) \right| \mathrm{d}m
\end{aligned}$$

where, in view of (29), the integral can be estimated by $4C_2'\gamma^k m(I_z^N)$. Hence there is a constant, not depending on z, such that, for large N,

$$m\{x \in I_z^N : |Y_N^{f(x)}(T^{r_N}x) - Y_N^{f(z)}(T^{r_N}x)| > 2r_N^{-1}\} \leqslant \text{const } a^{-r_N} r_N^2 m(I_z^N).$$

In order to prove (45) we now proceed as follows: it suffices to show that for each bounded Lipschitz function $\phi : \mathbb{R} \to \mathbb{R}$ holds

$$\mathbb{E}[\phi(Y_N)] = \int_0^1 \left(\int_I \phi(y) \, \mathrm{d}\mathcal{N}(0, \sigma_z^2)(y) \right) \mathrm{d}z + o(1) \qquad \text{as } N \to \infty \,. \tag{47}$$

To simplify the notation we write $\int_J \phi(Y_N) \, \mathrm{d}m$ instead of $\int_J \phi(Y_N(b)) \, \mathrm{d}b$ etc

$$\begin{aligned}
\mathbb{E}[\phi(Y_N)] &= \frac{1}{m(J)} \int_J \phi(Y_N^{f(x)}(T^{r_N}x)) \, \mathrm{d}x + o(1) \qquad \text{by (iii) for } z = f(x) \\
&= \sum_{z \in C_N} \frac{1}{m(J)} \int_{I_z^N \cap J} \phi(Y_N^{f(z)}(T^{r_N}x)) \, \mathrm{d}x + o(1) \qquad \text{by (iv)} \\
&= \sum_{z \in C_N, I_z^N \subseteq J} \frac{1}{m(J)} a^{-r_N} \int_I \phi(Y_N^{f(z)}) \, \mathrm{d}m + o(1) \\
&= \sum_{z \in C_N, I_z^N \subseteq J} \frac{m(I_z^N)}{m(J)} \int_I \phi \, \mathrm{d}\mathcal{N}(0, \sigma_{2f(z)}^2) + o(1) \qquad \text{by step 1} \\
&= \frac{1}{m(J)} \int_J \left(\int_I \phi \, \mathrm{d}\mathcal{N}(0, \sigma_{2f(z)}^2) \right) \mathrm{d}z + o(1) \\
&= \int_0^1 \left(\int_I \phi \, \mathrm{d}\mathcal{N}(0, \sigma_z^2) \right) \mathrm{d}z + o(1) \qquad \text{as } \mathcal{N} \to \infty
\end{aligned}$$

where one has to choose a sufficiently slowly growing sequence (r_N) in the fourth equality, and where the continuity of σ_{2z}^2 as a function of z is used in the fifth equality which is a simple consequence of our proposition 1. □

Remark 7. In [15,28,33] it has been shown that for $b = 0$ the dynamics of ψ_λ can be expressed in terms of generalized Takagi (or de Rham) functions. Analogous conclusions can be shown to hold for the case of $b \neq 0$. The above results are thus intimately related to continuity properties of this class of functions under parameter variation. These functions are defined by simple functional recursion relations and have been introduced in the literature completely independently from the diffusion problem considered here.

5. Numerical results

In this section we numerically quantify propositions 2 and 3 for the piecewise linear maps (34). By this we mean that we try to recover the functional forms of the upper bounds derived in proposition 2 in numerical box counting for D and J by furthermore determining the values of relevant parameters. Similarly, we try to recover the functional forms for the local modulus of continuity of D derived in proposition 3. We will see that our numerics yields additional information that complements the previous rigorous mathematical analysis.

For our computations we use the series representations of drift J and diffusion coefficient D from equations (73) and (74) in [16] which, for $|a| > 1$, converge quickly under numerical summation[7]. In [31] data sets generated from these expressions were analysed by standard numerical box counting under the assumption that

$$N(\epsilon) \sim \epsilon^{-B} \tag{48}$$

for small enough ϵ [43]. Here N is the number of square boxes of side length ϵ needed to cover the graph of J or D, and $B = \dim_B(\text{graph})$ defines the box (counting) dimension. Analysing $D(a) = D(a, 0)$ on $2 \leqslant a \leqslant 8$ yielded a box dimension of $B \simeq 1.039$ [31]. Computing furthermore the local box dimension $B(a)$ of $D(a)$ on a regular grid of small subintervals Δa centred around a produced locally different values forming an oscillatory structure in $B(a)$, which became more pronounced the smaller Δa, see figure 1 in [31].

5.1. Box counting for the diffusion coefficient

Motivated by proposition 2 and by [35], the numerical results of [31] are now reevaluated. We start with the diffusion coefficient $D(a)$. Corollary 2 states that $B(a) = 1$ for all intervals Δa, which is at variance with the results of [31] mentioned before. However, in contrast to (48), proposition 2 is compatible with the existence of multiplicative logarithmic terms by giving upper bounds for their exponents. Section 4.2 shows that these terms do indeed exist.

In detail, corollary 2 states an upper bound for the box counting function $N(\epsilon)$ of $D(a)$ of

$$N(\epsilon) \leqslant K_4 \epsilon^{-1}(1 - \ln \epsilon)^2 \quad . \tag{49}$$

This motivates us to plot the product $N\epsilon$ as a function of $-\ln \epsilon$: for small enough ϵ and in double-logarithmic representation one should then see a straight line with the slope yielding the exponent of the logarithmic term. Figure 1 numerically demonstrates the existence of this term for $D(a)$ on $2 \leqslant a \leqslant 8$: that is, the exponent is clearly non-zero, however, in the numerics $-\ln \epsilon$ is not large enough to overcome the additive constant in (49) for producing a straight line.

In figure 1 three data sets have been plotted consisting of different numbers of data points for $D(a)$. The bending off of the graphs at larger $-\ln \epsilon$ reflects that box counting starts to resolve the single points of the underlying data sets: from the figure one can estimate that for 10^6 points of $D(a)$ deviations set in around $-\ln \epsilon_{\text{cut}} \simeq 7$ or $\epsilon_{\text{cut}} \simeq 10^{-3}$. Compared with a separation of $\delta a = 6 \times 10^{-6}$ between any two data points along the a-axis, this yields a difference of about three orders of magnitude. The same order of magnitude argument holds if one compares ϵ_{cut} obtained approximately for 10^8 data points from figure 1 with the corresponding separation of $\delta a = 6 \times 10^{-8}$ between any two data points. This leads to the prediction that for the set of 10^9 data points $-\ln \epsilon_{\text{cut}} \simeq 14$ in figure 1.

Inspired by (49), we now fit the box counting results with the function

$$N(\epsilon) = K_5 \epsilon^{-1}(1 + K_6 \ln \epsilon)^{\alpha} \tag{50}$$

[7] Another set of formulae was reported in [10] but only for $D(a)$.

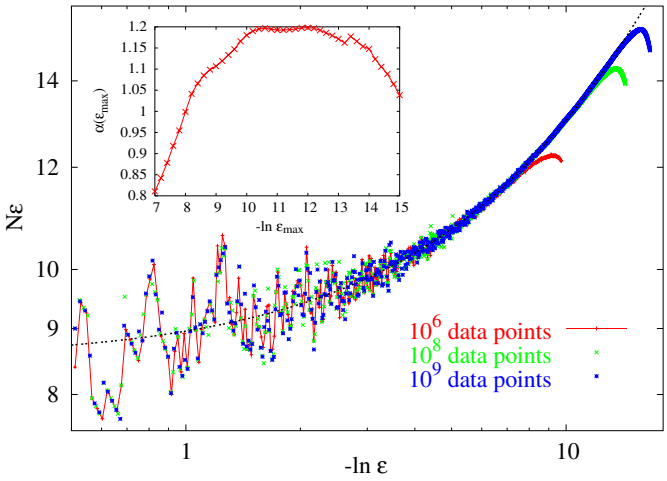

Figure 1. N is the number of boxes of length ϵ needed to cover the graph of $D(a)$ (a) generated from 10^6, 10^8 and 10^9 data points. Motivated by (49) we plot the product $N\epsilon$ as a function of $-\ln\epsilon$ double-logarithmically. The dashed line represents a three-parameter fit for the largest data set over $0.5 \leqslant -\ln\epsilon \leqslant 12$ with the functional form of (50). The inset shows results for the exponent of the logarithmic correction α obtained from fits where we vary the upper bound $\epsilon_{\max}$ of the fit interval.

instead of (48). If this fit function reproduces the numerical $N(\epsilon)$ reasonably well, proposition 2 predicts that $0 \leqslant \alpha \leqslant 2$. However, we emphasize that this proposition only gives us a strict upper bound—it does not actually tell us the 'true' functional form of the whole graph, nor does it exclude $\alpha = 0$. We have indeed checked that fit functions others than (50), which also obey (49), work similarly well. In order to be close to proposition 2 we stick to the fit function (50) in the following.

The dashed line in figure 1 shows a fit of the box counting results for 10^9 data points of $D(a)$ with this functional form[8]. The inset of figure 1 depicts results for the exponent α computed from different fit intervals $[0.5, -\ln\epsilon_{\max}]$ for the same data set of 10^9 points. It indicates convergence towards $\alpha \simeq 1.2$ ($-\ln\epsilon_{\max} \to 12$). The decrease for $-\ln\epsilon_{\max} > 12$ is well in agreement with the cutoff predicted above, which is due to the limited data set. Note that the cutoff sets in much later than the beginning of the plateau. Hence we conclude that for a data set of 10^9 points for $D(a)$, $2 \leqslant a \leqslant 8$, and by assuming the fit function (50), the numerical value for the exponent of the logarithmic term is $\alpha \simeq 1.2$, which is in agreement with proposition 2.[9] However, we emphasize that fits by (50) do not tell the full story: the numerically exact data in figure 1 show the existence of a non-trivial fine structure pointing towards more complicated functional forms for the 'true' $N(\epsilon)$, which should reflect the intricate structure of the underlying $D(a)$, see figure 1 in [31]. These irregularities are not numerical errors.

We now look at local variations of the exponent α, for which we do box counting of $D(a)$ on small intervals around integer values of a. Figure 2(a) reveals that there exist two families of curves, one for even a which is at the bottom of this figure, whereas the one for odd a is on top.

[8] For all fits the nonlinear least-squares Marquardt–Levenberg algorithm as implemented in gnuplot 4.0 has been used.

[9] We have checked that these fit results do not significantly depend on the choice of the initial seeds for our three fit parameters and that the asymptotic standard error for them is less than 10% for $-\ln\epsilon_{\max} > 10$. However, in our view quantitative error estimates are not reliable in this case, because we may not assume that the residua are normally distributed random variables.

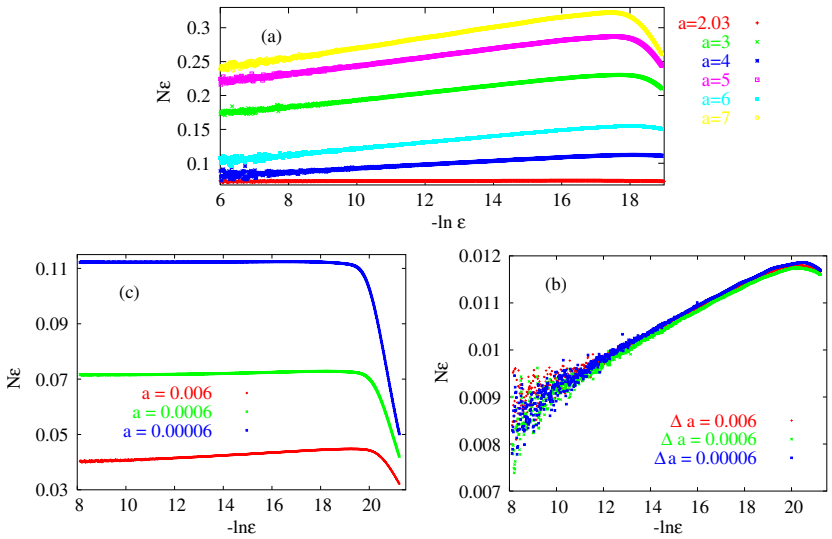

Figure 2. Local variation of the product $N\epsilon$ needed to cover $D(a)$ around integer values of a: (a) shows results for parameter intervals of size $\Delta a = 0.06$ centred around different a, based on 10^8 data points. (b) Displays results for subintervals Δa all centred around $a = 4$, whereas in (c) all subintervals converge towards $a = 5$. In (b) and (c) the graphs have been scaled by multiplying ϵ with the order of magnitude difference between the different values for Δa.

Additionally, graphs for larger slopes are always on top in both groups creating an oscillatory structure. We first consider the special case $a = 2.03$, where according to figure 2(a) $\alpha \simeq 0$. Note that $D(2) = 0$, correspondingly the parameter region just above $a = 2$ marks the onset of diffusion. As described in [28, 29, 33], for $a \to 2$ there is asymptotic convergence of $D(a)$ to the simple random walk solution $D(a) = (a-2)/(2a)$. This physical argument explains why $\alpha \to 0\,(a \to 2)$. There is a trend that larger even integer slopes in (a) give $0 \leqslant \alpha \leqslant 1$ whereas odd a give $1 < \alpha \leqslant 2$. Unfortunately, the fits producing these results are very unstable, hence even these rough estimates should be taken with care. In any case, the indicated order of magnitude of α appears to be in agreement with proposition 2. Our fits furthermore suggest that not only α is a function of a but also that the other two parameters in (50) are locally varying. This agrees with conclusions drawn in [31].

Figures 2(b) and (c) provide a more detailed local analysis by looking at successively smaller subintervals around two specific slopes. While (c) suggests $\alpha \to 0$ $(\Delta a \to 0)$ around $a = 5$, (b) with $a = 4$ yields approximately $\alpha \to 1$ $(\Delta a \to 0)$.[10] Note that the graphs in (b) and (c) have been scaled as described in the figure. Interestingly, this transformation leads to a collapse onto a master curve in (b), whereas it does not work that way in (c). Similar observations have been reported in [35]. Together with proposition 3, figure 2 thus demonstrates remarkable continuity properties of $D(a, b)$ around integer slopes, which strongly depend on the direction in parameter space. These differences between graphs for odd and even a are consistent with the oscillations in the local box counting dimension $B(a)$ reported in figure 1 of [31]: they suggest that the structure of $B(a)$ reflects local variations of the parameters in (50) determining the logarithmic corrections, erroneously being fit in [31] with (48) instead of taking the existence of logarithmic terms into account.

Such local variations of α pose a serious problem to any numerical box counting because of monotonicity of these exponents: if E is a subset of F then $\alpha(E) \leqslant \alpha(F)$. This implies

[10] Again, the fit results are highly unstable, so the latter value should be taken with care.

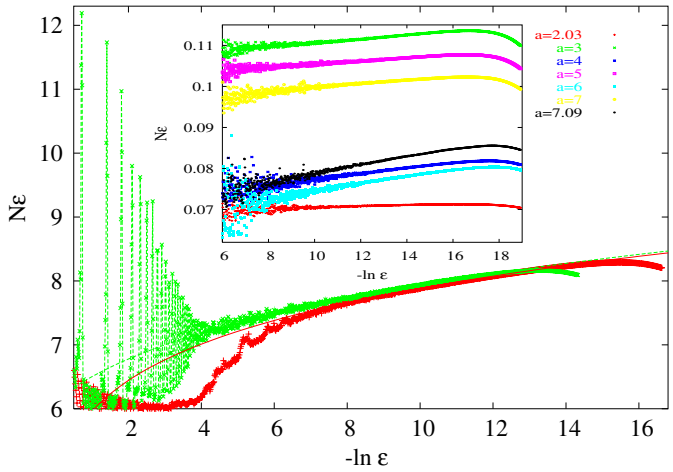

Figure 3. Main graph: product $N\epsilon$ as a function of $-\ln\epsilon$ for the drift $J(a, b)$ over the interval $2 \leqslant a \leqslant 8$ at $b = 0.01$ ('+' symbols, based on 10^9 data points) and at $b = 0.49$ ('x' symbols, based on 10^8 data points). Included are two fits over the intervals $8 \leqslant -\ln\epsilon \leqslant 13$ ($b = 0.01$) and $4 \leqslant -\ln\epsilon \leqslant 11.5$ ($b = 0.49$). Inset: local variation of the product $N\epsilon$ as a function of $-\ln\epsilon$ for parameter intervals of size $\Delta a = 0.06$, mostly centred around integer values of a and based on 10^8 data points. The graph for $a = 7.09$ demonstrates that there exist strong local fluctuations of the box counting functions under variation of the slope a of the map.

that $\alpha(a)$ should eventually converge to the largest local exponent. However, if this exponent is exhibited on a tiny subinterval it could be extremely tedious to detect it numerically. This argument also applies to our previous result of $\alpha \simeq 1.2$ for $D(a)$ on $2 \leqslant a \leqslant 8$, which strictly speaking only holds for the given data set of 10^9 points. We cannot exclude that some tiny interval of $D(a)$ eventually yields a larger α. In other words, the goal of our numerical analysis cannot be to compute unambiguous values for any exponents but rather to demonstrate qualitative and quantitative order of magnitude agreement with proposition 2, and to check for local variations of such exponents.

5.2. Box counting for the drift

We now numerically investigate the parameter dependence of the drift, or current, $J(a, b)$. As for the diffusion coefficient, both the drift and the associated local box dimension $B(a)$ defined by (48) display highly oscillatory structures for fixed b, see figure 4 in [31]. As before, we now reevaluate these findings on the basis of proposition 2 by taking logarithmic terms into account. Figure 3 shows numerically that there are non-zero exponents α for the logarithmic corrections of $J(a, b)$ as allowed by proposition 2. Note particularly the pronounced, different fine structures of both curves displayed in the main part, which are much stronger than for $D(a)$ in figure 1. Due to these oscillations, in case of $J(a, b)$ it is numerically very difficult to extract reliable values for the exponents α by using (50). The two fits included in the main graph yield an order of magnitude of $\alpha \simeq 0.1$, which matches to proposition 2.

The inset of figure 3 is analogous to figure 2(a) in that it shows box counting results for the current $J(a, b)$, $b = 0.01$, mostly at integer values of the slope a. Note that $J(2, 0.01) \simeq 0$ [16], which marks the onset of the drift. As we have argued for the diffusion coefficient, at $a = 2.03$ we are thus in a random walk regime for which one may expect $\alpha \simeq 0$, as shown in the figure. However, $\alpha(a) \neq 0$ in all the other cases of the inset suggesting again

a local variability of α for $J(a, b)$, at least around integer values of a. As in figure 2(a) there exist two family of curves, one for even a at the bottom and one for odd a on top of the figure. There is also again an additional ordering, however, here it is such that curves for larger slopes are always at the bottom in both families of graphs, except at $a = 2.03$. The additional graph for $a = 7.09$ exemplifies the strong local variability of α around $a = 7$ which, as well as the difference between odd and even slopes for box counting results of the drift, agrees with the oscillations in the local box dimension $B(a)$ shown in figure 4 of [31].

Fits for all the inset curves yield a trend towards small exponents around even and somewhat larger values around odd slopes with an order of magnitude of $0 \leqslant \alpha \leqslant 1$, which appears to be consistent with proposition 2. However, we emphasize again that these results give only a rough indication, for the numerical reasons mentioned above. Exact results are only available for special cases: as we have discussed in section 4.2, $J(a, b) = b$ for constant $a \in \mathbb{N}$ under variation of b, where we thus have $\alpha = 0$, linear response and a caricature of Ohm's law. For general a one finds that $J(a, b)/(b|\log|b||)$ is bounded but has no limit for $b \to 0$ [16] pointing towards logarithmic corrections. We have also qualitatively checked graphs of $D(a, b)$ and $J(a, b)$ for other parameter values, that is, by choosing different values for a and b fixed in the parameter plane and studying the resulting functions of the remaining free control parameters. Qualitatively, we obtain results that are analogous to the ones discussed above.

5.3. Continuity properties of the diffusion coefficient at integer slopes

The previous two subsections demonstrated a peculiar behaviour of local box counting results for drift and diffusion coefficient around integer slopes a at fixed values of the bias b. Proposition 3, in turn, gave exact analytical expressions for the difference $D_a(b') - D_a(b)$ of the diffusion coefficient as a function of $\Delta b = b' - b$ at integer a in the limit of small Δb. This suggests to numerically study the continuity properties of $D_a(b)$ at fixed integer values of a in more detail. In order to access suitably small values of the parameter Δb, we have employed the Fortran90 library mpfun90 [1] for arbitrary-precision arithmetic. Using this library we have calculated the difference quotient $(D_a(b') - D_a(b))/(b' - b)$ of D with fixed b at values of Δb down to 10^{-200}. Figure 4(a) shows a subset of our results for $a = 3$ and $a = 4$ at fixed $b \in \{-0.5, 0\}$. There is excellent coarse scale agreement between the numerical results and the analytical observations (1) and (2) of proposition 3 predicting straight lines. This agreement is as good to the limits of attainable precision, and has been checked for other integer values than those shown in figure 4.

Figure 4(b) depicts the diffusion coefficient $D_a(b)$ at $a = 4$, which corresponds to the two curves in (a) at this a value. There is reflection symmetry for $D_a(b)$ with respect to $b = -0.5$ and $b = 0$. One can see that at $b = -0.5$, where the difference quotient in (a) displays a logarithmic correction, $D_a(b)$ in (b) exhibits a global maximum as a sharp cusp. The global minimum at $b = 0$, on the other hand, is rounded-off, see the blowup in (b), reflecting the difference quotient curve in (a) with zero logarithmic term. Observation (3) of proposition 3 generalizes observations (1) and (2) by stating that logarithmic corrections are typical for parameter values of b yielding Markov partitions. In [16, 27, 28, 30] it was shown (for $b = 0$) that Markov partition parameter values identify local maxima and minima of the parameter-dependent diffusion coefficient by relating them to ballistic and localized orbits of the critical points of the lifted map, respectively. In other words, in plots of the parameter-dependent diffusion coefficient $D_a(b)$ the existence or not of local logarithmic corrections seems to be visible in the form of characteristic shapes for the extrema as discussed above; see also [35] for related results. This observation could be interesting for understanding irregularities in

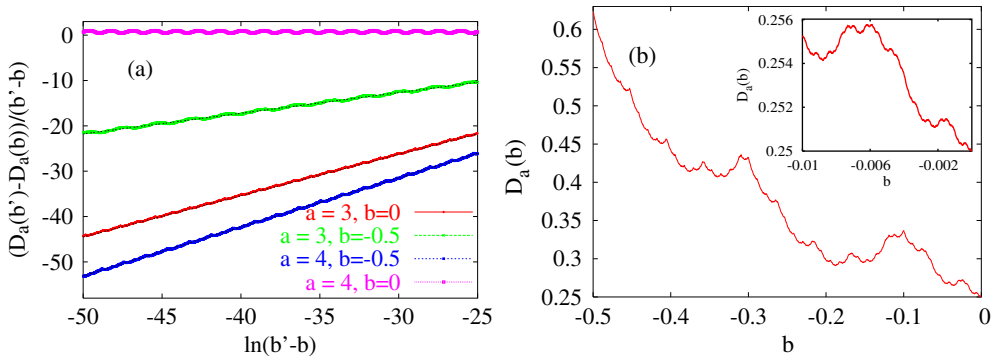

Figure 4. (*a*) Difference quotient $(D_a(b') - D(b))/(b' - b)$ as a function of $\ln(b' - b)$ at integer values of $a \in \{3, 4\}$ and with fixed $b \in \{-0.5, 0\}$. Each curve is based on 10^5 data points. Included are best fit curves (dashed lines, which are almost indistinguishable from the corresponding data points) whose fitted slopes (from bottom to top: $3/(2\ln 4)$, $1/\ln 3$ and $1/(2\ln 3)$) agree with the analytic predictions of section 4.2 to four significant figures. The case $a = 4, b = 0$ clearly has slope zero, as predicted. The barely visible fine scale oscillations of each curve reflect higher order correlations in these quantities. (*b*) Diffusion coefficient $D_a(b)$ at $a = 4$ for $-0.5 \leqslant b \leqslant 0$ and a magnification of the region around $b = 0$. For each curve 2000 data points have been computed from exact analytical solutions for $D_a(b)$ [16]. These curves form the basis for the two graphs at $a = 4$ displayed in (*a*).

the simulation results of transport coefficients in more complicated systems, where rigorous results are not available [33]. In figure 4(*a*) we deliberately restricted the range of $\ln(b' - b)$ so that, upon close scrutiny, a fine structure of all curves can be seen on top of the straight line behaviour. Figure 4(*b*) suggests that this oscillatory fine structure, which yields higher order corrections to the analytical results of section 4.2, is induced by the fine structure of $D_a(b)$.

Finally, we numerically quantify observation (4) in proposition 2. Its main statement is that at fixed Δb and with b values taken uniformly from the interval $[0, 1/2)$, the quantity $(D(b') - D(b))/\Delta b\sqrt{-\ln \Delta b}$ should be distributed like a mixture of centred Gaussians, that distribution being independent of the particular value of Δb. In fact what is typically seen at integer slopes is a distribution rather close to a pure Gaussian. We have tested this using the technique of quantile–quantile plotting (qqplots) as well the standard Shapiro–Wilk normality test. Both tools were implemented in the statistical package R [44]. Figure 5 presents results obtained for three sets of data with the slope fixed at $a = 4$. For larger a, the results become closer to a fixed Gaussian, as the function $g(x)$ in (43) becomes more dominated by the x term which has no b dependence. Here, however, deviations from Gaussianity can be seen, at least for sufficiently small Δb. In the three parts of figure 5, the straight line with slope σ and zero offset μ shows the theoretical result for a Gaussian distribution with standard deviation σ and mean μ, with those parameters here taken as those of our data set. All our distributions show close agreement with this curve. However, the Shapiro–Wilk normality test is more discerning: in (*a*) $\Delta b = 10^{-10}$ and we obtain a p-value of only 0.008, well below the significance level for rejecting the null hypothesis of normality. In (*b*) $\Delta b = 10^{-50}$ and we get a p-value of 0.25, demonstrating that this distribution is indeed very close to a pure Gaussian. It is, however, likely that the deviations from Gaussianity in (*a*) are rather due to deterministic effects arising from the relatively large value of Δb chosen and not from the nature of the true limiting distribution being a mixture of Gaussians predicted by observation (4) in proposition 3. It seems that the dominant behaviour when the distribution seems to have converged is not detectably different from a pure Gaussian. We note that despite this, the two distributions

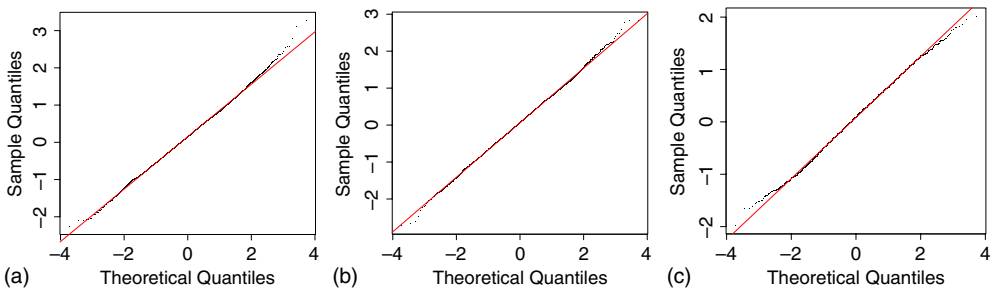

Figure 5. (*a*) Normal quantile–quantile plot at $a = 4$ for the distribution of $D(b') - D(b)$, scaled by $\Delta b\sqrt{-\ln \Delta b}$, with chosen $\Delta b = b' - b = 10^{-10}$ held constant and b picked from a uniform distribution on $[0, 1/2)$. The straight line with slope σ and zero offset μ would be the result for a Gaussian distribution with standard deviation σ and mean μ. Here the numerically obtained values of these parameters were used for the fit. (*b*) As (*a*) but with $\Delta b = 10^{-50}$. (*c*) As (*b*) but with the range of b restricted to $[0, 0.005)$.

in (*a*) and (*b*) are similar and both have mean close to zero, demonstrating that we have no disagreement with observation (4), merely that its details are too sensitive to check numerically.

It is, however, possible to numerically go beyond proposition 3 (4). For example one can also study the nature of the distribution obtained when b is taken from a subinterval of $[0, 1/2)$, which, as can be seen in figure 5(*c*), leads in the case of integer a to distributions with rather more fine structure than the nice curves seen in figures 5(*a*) and (*b*). This is clear evidence of the deterministic nature of the underlying system in the form of strong correlations at fine scales. In this case the Shapiro–Wilk p-value is about 0.001. Furthermore, away from integer values of a quite different behaviour is seen thus clarifying that proposition 3 (4) is rather to be considered atypical. Here the distribution of D—differences is centred around zero still, but with a more sharply peaked and heavily tailed distribution than a Gaussian. These deviations persist even very close to the integer cases (e.g. at $a = 3 + 10^{-50}$), though Gaussian behaviour does appear to be approached slowly in the limit of integer values. These numerical methods can also be used to investigate variation of the continuity of the transport coefficients as b is held fixed and a varies, as considered in [35] and already looked at using box counting in figure 2. Here the maximal exponent of logarithmic correction, i.e. $D(a') - D(a) \sim |a' - a|(\ln |a' - a|)^2$, can be seen for odd a, and though this might appear to be in contradiction to the third part of figure 2 for $a = 5$, in fact arbitrarily close to $a = 5$ the exponent tends locally to zero. Thus the box counting only sees the 'typical' local behaviour and the current method is more suited for picking out atypical behaviour at specific points.

6. Conclusions and outlook

(1) We proved rigorously that the diffusion coefficient of deterministic random walks generated by piecewise expanding interval maps depends continuously on the maps. More precisely, for 'natural' parametrizations of the maps by some parameter λ, the diffusion coefficient as a function of the parameters has a modulus of continuity not worse than $|\delta\lambda|(\log |\delta\lambda|)^2$. The detailed analysis of section 4.2 shows that this modulus of continuity cannot be expected to be better than $|\delta\lambda \cdot \log |\delta\lambda||$ even if all maps have a common smooth invariant density and the drift depends smoothly on a parameter. Comparing this to a detailed study of the parameter dependence of the averages of observables in [4] one might ask whether it is a more common effect in uniformly hyperbolic systems that the

modulus of continuity of the diffusion coefficient is one logarithmic order worse than that of the drift[11].

These rigorous results are intimately related to the question of under which mathematical conditions in simple models physical observables such as transport coefficients may behave regularly or irregularly under parameter variation. Such information is important, e.g. for a detailed understanding of the origin of linear response in chaotic dynamical systems [4, 20, 21, 33].

(2) The logarithmic corrections obtained by our mathematical theory were quantified numerically in the box counting data of both the parameter-dependent drift and diffusion coefficients. This numerical analysis revealed the existence of strong local variations of the parameters governing the logarithmic Lipschitz continuity, partially we computed values for them. It furthermore showed that numerically there exist non-trivial fine structures on top of the coarse functional forms derived in our mathematical propositions.

These new numerical results correct and amend the previous box counting analysis of Klages and Klauß [31] along the lines conjectured by Koza [35]. Our model thus generates very unusual examples of fractals yielding a box counting dimension of one, where the fractality can be understood in terms of multiplicative logarithmic terms that amend the usual box counting power law. We conclude that the (local) non-integer variations of the standard box counting dimension reported in [31] actually reflect non-trivial local variations of the parameters of these logarithmic corrections.

We have furthermore numerically quantified analytical results for the difference quotient of the diffusion coefficient as a function of the bias at integer slopes. Our computations suggest that the existence or not and the strength of local logarithmic corrections is directly visible in plots of the parameter-dependent diffusion coefficient in form of characteristic shapes of the local extrema.

(3) In [34] a nonlinear generalization of our present model has been studied, which exhibits anomalous diffusion generated by marginal fixed points. Computer simulations led to conjecture that the anomalous diffusion coefficient of this map is discontinuous on a dense set of parameter values. It would be interesting to check this conjecture mathematically.

These fractal transport coefficients also seem to provide a nice testing ground for methods of multifractal analysis [12].

Another important problem is to check whether such logarithmic corrections in transport coefficients might also be expected to occur in more 'physical' systems, which are perhaps even accessible experimentally. This seems to be strongly related to the question of whether a family of physical dynamical systems shares the same topological conjugacy class under parameter variation.

Acknowledgments

GK and RK thank C Beck, C Dettmann and M Pollicott, the organizers of the LMS Durham Symposium on *Dynamical Systems and Statistical Mechanics* in July 2006, where this work was started, for their kind invitation. All three authors gratefully acknowledge very detailed and helpful comments by one of the referees. RK and PJH were supported by a grant from the British EPSRC under EP/E00492X/1.

[11] Observe, however, that the situation in [4] differs from ours in that the maps are continuous (though not differentiable) and the observables are differentiable.

References

[1] Bailey D 1995 A Fortran-90 based multiprecision system *ACM Trans. Math. Softw.* **21** 379–87

[2] Baladi V 2000 *Positive Transfer Operators and Decay of Correlations* (*Advanced Series in Nonlinear Dynamics* vol 16) (Singapore: World Scientific)

[3] Baladi V 2007 On the susceptibility function of piecewise expanding interval maps *Commun. Math. Phys.* **275** 839–59

[4] Baladi V and Smania D 2008 Linear response formula for piecewise expanding unimodal maps *Nonlinearity* **21** 677–711
(See also the recent preprint arXiv:0803.3344v1 by the authors.)

[5] Blank M L 1993 Singular effects in chaotic dynamical systems *Russ. Acad. Sci. Dokl. Math.* **47** 1–5

[6] Blank M and Keller G 1997 Stochastic stability versus localization in chaotic dynamical systems *Nonlinearity* **10** 81–107

[7] Blank M and Keller G 1998 Random perturbations of chaotic dynamical systems: Stability of the spectrum *Nonlinearity* **11** 1351–64

[8] Bowen R 1977 Bernoulli maps of the interval *Israel J. Math.* **28** 161–8

[9] Butterley O and Liverani C 2007 Smooth Anosov flows: correlation spectra and stability *J. Mod. Dyn.* **1** 301–22

[10] Cristadoro G 2006 Fractal diffusion coefficient from dynamical zeta functions *J. Phys. A: Math. Gen.* **39** L151–7

[11] Dolgopyat D 2004 On differentiability of SRB states for partially hyperbolic systems *Invent. Math.* **155** 389–449

[12] Faccini A, Wimberger S and Tomadin A 2007 Multifractal fluctuations in the survival probability of an open quantum system *Physica* A **376** 266–74

[13] Flatto L and Lagarias J C 1997 The lap-counting function for linear mod one transformations II. The Markov chain for generalized lap numbers *Ergod. Theory Dyn. Syst.* **17** 123–46

[14] Flatto L and Lagarias J C 1997 The lap-counting function for linear mod one transformations III. The period of a Markov chain *Ergod. Theory Dyn. Syst.* **17** 369–403

[15] Gaspard P and Klages R 1998 Chaotic and fractal properties of deterministic diffusion-reaction processes *Chaos* **8** 409–23

[16] Groeneveld J and Klages R 2002 Negative and nonlinear response in an exacly solved dynamical model of particle transport *J. Stat. Phys.* **109** 821–61

[17] Hofbauer F 1980 Maximal measures for simple piecewise monotonic transformations *Z. Wahrscheinlichkeitstheor. verwandte Geb. (*now *Probab. Theory Relat. Fields)* **52** 289–300

[18] Hofbauer F 1981 The maximal measure for linear mod one transformations *J. Lond. Math. Soc.* **23** 92–112

[19] Hofbauer F and Keller G 1982 Ergodic properties of invariant measures for piecewise monotonic transformations *Math. Z.* **180** 119–40

[20] Jiang M and de la Llave R 2006 Linear response function for coupled hyperbolic attractors *Commun. Math. Phys.* **261** 379–404

[21] Jiang Y and Ruelle D 2005 Analyticity of the susceptibility function for unimodal markovian maps of the interval *Nonlinearity* **18** 2447–53

[22] Keller G 1980 Un théorème de la limite centrale pour une classe de transformations monotones par morceaux *C. R. Acad. Sci. Paris Sér* A **291** 155–8

[23] Keller G 1982 Stochastic stability in some chaotic dynamical systems *Monatshefte Math.* **94** 313–33

[24] Keller G 1999 Interval maps with strictly contracting Perron–Frobenius operators *Int. J. Bifurc. Chaos* **9** 1777–84

[25] Keller G and Liverani C 1999 Stability of the spectrum for transfer operators *Ann. Mat. Sc. Norm. Pisa* **28** 141–52

[26] Keller G and Liverani C 2005 A spectral gap for a one-dimensional lattice of coupled piecewise expanding interval maps *Dynamics of Coupled Map Lattices and of Related Spatially Extended Systems* (*Lecture Notes in Physics* vol 671) ed J-R Chazottes and B Fernandez (Berlin: Springer) pp 115–51

[27] Klages R and Dorfman J R 1995 Simple maps with fractal diffusion coefficients *Phys. Rev. Lett.* **74** 387–90

[28] Klages R 1996 *Deterministic diffusion in one-dimensional chaotic dynamical systems* (Berlin: Wissenschaft & Technik)

[29] Klages R and Dorfman J R 1997 Dynamical crossover in deterministic diffusion *Phys. Rev.* E **55** R1247–50

[30] Klages R and Dorfman J R 1999 Simple deterministic dynamical systems with fractal diffusion coefficients *Phys. Rev.* E **59** 5361–83

[31] Klages R and KlaußT 2003 Fractal fractal dimensions of deterministic transport coefficients *J. Phys. A: Math. Gen.* **36** 5747–64

[32] Klages R, Barna I F and Mátyás L 2004 Spiral modes in the diffusion of a single granular particle on a vibrating surface *Phys. Lett.* A **333** 79–84
[33] Klages R 2007 *Microscopic Chaos, Fractals and Transport in Nonequilibrium Statistical Mechanics* (*Advanced Series in Nonlinear Dynamics* vol 24) (Singapore: World Scientific)
[34] Korabel N, Klages R, Chechkin A V, Sokolov I M and Gonchar V Yu 2007 Fractal properties of anomalous diffusion in intermittent maps *Phys. Rev.* E **75** 036213
[35] Koza Z 2004 Fractal dimension of transport coefficients in a deterministic dynamical system *J. Phys. A: Math. Gen.* **37** 10859–77
[36] Lasota A and Yorke J A 1973 On the existence of invariant measures for piecewise monotonic transformations *Trans. Am. Math. Soc.* **186** 481–8
[37] Liverani C 1996 Central limit theorem for deterministic systems *International Conference on Dynamical Systems (Montevideo, Uruguay, 1995)* (*Pitman Research Notes in Mathematics Series* vol 362) ed F Ledrappier *et al* (a tribute to Ricardo Mañe)
[38] Mazzolena M 2007 Dinamiche espansive unidimensionali: dipendenza della misura invariante da un parametro *Master's Thesis* Roma 2
[39] Rousseau-Egele J 1983 Un théorème de la limite locale pour une classe de transformations dilatantes et monotones par morceaux *Ann. Probab.* **11** 772–88
[40] Ruelle D 1997 Differentiation of SRB states *Commun. Math. Phys.* **187** 227–241
See also Ruelle D 2003 *Commun. Math. Phys.* **234** 185–90
[41] Ruelle D 2005 Differentiating the a.c.i.m. of an interval map with respect to f *Commun. Math. Phys.* **258** 445–53
[42] Ruelle D 2007 Structure and f-dependence of the a.c.i.m. for a unimodal map of Misiurewicz type *Preprint* 0710.2015v1 [math.DS]
[43] Tricot C 1995 *Curves and Fractal Dimension* (Berlin: Springer)
[44] R Development Core Team 2007 *R: A Language and Environment for Statistical Computing* (Vienna: R Foundation for Statistical Computing)