MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

"Hitting- and Return-time Statistics of Piecewise Expanding Dynamical Systems"

verfasst von / submitted by

Luisa Weginger

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2022 / Vienna 2022

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066821
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Mathematik
Betreut von / Supervisor:	Assoz. Prof. Mag. Dr. Roland Zweimüller, Privatdoz.
Mitbetreut von / Co.-Supervisor:	Dipl.-Math. Dr. Tanja Schindler, B.Sc.

Abstract

This Master's Thesis provides insight into the hitting- and return-time statistics of some dynamical systems and investigates which conditions are sufficient to infer convergence to exponential limit laws.

We show that Rychlik systems admit exponential limit laws for intervals if their invariant density is bounded away from 0. We give an example where the invariant density is not bounded away from 0 but exponential hitting- and return-time statistics still arise. Furthermore this example itself is Rychlik but the system with respect to its invariant measure is not. This gives rise to the question, whether or not this system still admits the same properties as a Rychlik system. For higher dimensions the family of piecewise expanding systems admits exponential hitting- and return-time statistics for balls around a non-periodic point x^* if again the invariant density is bounded away from 0.

Zusammenfassung

Diese Master-Arbeit gibt Einblicke in das Langzeitverhalten von den Treffer- und Rückkehrzeit Verteilungen von stückweise expandierenden Systemen und wir untersuchen welche Vorausetzungen notwendig sind um Konvergenz zu einer Exponentialverteilung herzuleiten.
Wir zeigen, dass die Treffer-und Rückkehrzeit Verteilungen von Rychlik Systemen gegen die Exponentialverteilung konvergieren, falls die invariante Dichte von 0 weg beschränkt ist. Außerdem geben wir ein Beispiel an, in welchem die Dichte nicht von 0 weg beschränkt ist, aber das gleiche Langzeitverhalten aufkommt. Dieses Beispiel ist selbst Rychlik, das System bezüglich des invarianten Maßes ist jedoch nicht Rychlik. Die Frage, ob dieses System die gleichen Eigenschaften wie ein Rychlik System hat, kommt auf. Für höhere Dimensionen betrachten wir die Familie von stückweise expandierenden Systemen, die wiederum für Bälle um einen nicht periodischen Punkt x^* exponentielles Langzeitverhalten von Treffer- und Rückkehrzeit Verteilungen aufweist, falls die invariante Dichte von 0 weg beschränkt ist.

Contents

1 Collection of Results from Measure and Ergodic Theory

In the following chapter we include some definitions and results from measure and ergodic theory to then talk about the hitting- and return time statistics of different families of dynamical systems. We start by including different notions of convergence, the connection between the space of probability measures and the space of their densities and the fact that weak convergence is metrizable.

1.1 Convergence of measures

Definition 1.1. Let $(\mathbb{M}, d_{\mathbb{M}})$ be a metric space endowed with the Borel σ-algebra $\mathcal{B}_{\mathbb{M}}$ and let ν_n, $n \geq 1$ and ν be measures on $(\mathbb{M}, \mathcal{B}_{\mathbb{M}})$. Furthermore let $(\mathbb{X}, \mathcal{A})$ and $(\mathbb{Y}, \mathcal{C})$ be measure spaces, $\mu_{\mathbb{X}}$, $\mu_{\mathbb{Y}}$ probability measures on $(\mathbb{X}, \mathcal{A})$, $(\mathbb{Y}, \mathcal{C})$ respectively and $(\mu_{\mathbb{X}}^n)_{n\geq 1}$ a sequence of probability measures on $(\mathbb{X}, \mathcal{A})$. Let $R_n : \mathbb{X} \longrightarrow \mathbb{M}$ and $R : \mathbb{Y} \longrightarrow \mathbb{M}$ be random elements of $\mathbb{M}$.

(i) We say $\underline{\nu_n \text{ converges weakly to } \nu}$ (write $\nu_n \Longrightarrow \nu$), if for all bounded continuous functions $\psi : \mathbb{M} \longrightarrow \mathbb{R}$ it holds that $\int_{\mathbb{M}} \psi d\nu_n \longrightarrow \int_{\mathbb{M}} \psi d\nu$ as $n \to \infty$.

(ii) We say $\underline{R_n \text{ converges in distribution to } R}$ (write $R_n \overset{\mu_{\mathbb{X}}}{\Longrightarrow} R$), if $\mu_{\mathbb{X}} \circ R_n^{-1} \Longrightarrow \mu_{\mathbb{Y}} \circ R^{-1}$.

(iii) We write $R_n \overset{\mu_{\mathbb{X}}^n}{\Longrightarrow} R$ if $\mu_{\mathbb{X}}^n \circ R_n^{-1} \Longrightarrow \mu_{\mathbb{Y}} \circ R^{-1}$.

Remark 1.2. The following two properties are equivalent:

(i) $\nu_n \Longrightarrow \nu$

(ii) $\int_{\mathbb{M}} \psi d\nu_n \longrightarrow \int_{\mathbb{M}} \psi d\nu$ as $n \to \infty$, for all bounded Lipschitz-continuous functions $\psi : \mathbb{M} \to \mathbb{R}$.

If we look at the distributional convergence of random elements R_n of $\mathbb{R}$, we get the following equivalence:

(i) $R_n \overset{\mu_{\mathbb{X}}^n}{\Longrightarrow} R$

(ii) $F_n(t) := \mu_{\mathbb{X}}^n(R_n \leq t) \to F(t) := \mu_{\mathbb{Y}}(R \leq t)$ as $n \to \infty$, for all continuity points t of F.

In the following we will often switch between talking about a compact set of measures and the compact set of the corresponding densities. We first introduce some notation and then argue why this is alright.

Definition 1.3. Let $(\mathbb{X}, \mathcal{A}, \lambda)$ be a measure space. We define the set
$\mathcal{P}(\lambda) := \{\nu : \nu \text{ is a probability measure on } (\mathbb{X}, \mathcal{A}, \lambda) \text{ and } \nu << \lambda\}$ to be the set of probability measures on $(\mathbb{X}, \mathcal{A})$ absolutely continuous with respect to λ.
$\mathcal{P}(\lambda)$ equipped with the total variation distance $d_{\mathcal{P}}(\nu, \tilde{\nu}) := 2 \sup\limits_{A\in\mathcal{A}} |\nu(A) - \tilde{\nu}(A)|$ is a metric space.
Furthermore we define $\mathcal{D}(\lambda) \subseteq L^1(\lambda)$ to be the set of all probability densities with respect to λ.

Remark 1.4. Take ν, $\tilde{\nu} \in \mathcal{P}(\lambda)$ and define $u := \frac{d\nu}{d\lambda}$ and $\tilde{u} := \frac{d\tilde{\nu}}{d\lambda}$. Then

$$d_{\mathcal{P}}(\nu, \tilde{\nu}) = 2 \sup_{A\in\mathcal{A}} |\int_A u \, d\lambda - \int_A \tilde{u} \, d\lambda| \leq 2 \sup_{A\in\mathcal{A}} \int_A |u - \tilde{u}| d\lambda \leq 2\|u - \tilde{u}\|_1 \tag{1.1}$$

and

$$\|u - \tilde{u}\|_1 = \int\limits_B u - \tilde{u}\, d\lambda + \int\limits_{B^C} \tilde{u} - u\, d\lambda \leq 2 \sup_{A \in \mathcal{A}} |\nu(A) - \tilde{\nu}(A)| \tag{1.2}$$

with $B = \{x \in \mathbb{X} : u - \tilde{u} \geq 0\}$.
With these inequalities we get that for $\mathcal{U} \subseteq \mathcal{P}(\lambda)$, the set $\{\frac{d\nu}{d\lambda} : \nu \in \mathcal{U}\} \subseteq \mathcal{D}(\lambda)$ is compact in $L^1(\lambda)$ iff $\mathcal{U}$ is compact in $\mathcal{P}(\lambda)$.

Later we are going to use that weak-convergence is metrizable, which we discuss in the next remark.

Remark 1.5. For a compact separable metric space $(\mathbb{M}, d_{\mathbb{M}})$ we denote by $\mathcal{M}(\mathcal{B}_{\mathbb{M}})$ the set of all finite signed measures on the Borel σ-algebra $\mathcal{B}_{\mathbb{M}}$ on $\mathbb{M}$. Let $\mathcal{C}(\mathbb{M})$ be the set of all continuous functions from $\mathbb{M}$ to $\mathbb{R}$ and denote by $\mathcal{C}(\mathbb{M})^*$ the set of all linear functionals on $\mathcal{C}(\mathbb{M})$. Whenever we talk about a metric on $\mathcal{C}(\mathbb{M})$ we mean the supremum norm $\|\cdot\|_\infty$. By the Riesz representation Theorem we know that $\mathcal{C}(\mathbb{M})^* = \mathcal{M}(\mathcal{B}_{\mathbb{M}})$.
From general theory on the weak*-topology, it follows that the weak*-topology on the closed unit ball $\overline{B_{\mathcal{M}(\mathcal{B}_{\mathbb{M}})}}$ is metrizable with metric $D_{\mathbb{M}}$ defined as

$$D_{\mathbb{M}}(\nu, \tilde{\nu}) := \sum_{j=1}^{\infty} 2^{-(j+1)} | \int\limits_{\mathbb{M}} \chi_j d\nu - \int_{\mathbb{M}} \chi_j d\tilde{\nu}|,$$

for some dense subset $\{\chi_j,\ j \geq 1\}$ of the closed unit ball $\overline{B_{\mathcal{C}(\mathbb{M})}}$.
For any such metric and for any sequence $(\nu_l)_{l\geq 1}$ and any ν in $\overline{B_{\mathcal{M}(\mathcal{B}_{\mathbb{M}})}}$, $D_{\mathbb{M}}(\nu_l, \nu) \longrightarrow 0$ as $l \to \infty$ iff $\nu_l \Longrightarrow \nu$. (Since convergence in the weak*-topology corresponds to the weak convergence as defined in Definition 1.1.) For more information see for example [12].

From now on, whenever we use the metric $D_{\mathbb{M}}$, we take $(\chi_j)_{j\geq 1}$ to be a fixed, dense sequence of Lipschitz functions in $\overline{B_{\mathcal{C}(\mathbb{M})}}$.

1.2 Ergodic Dynamical Systems and their transfer operator

In the following chapter we are going to introduce and recall some basic definitions and theorems from ergodic theory. The content can be found in the lecture notes of the course on ergodic theory given by Roland Zweimüller in the fall term 2021/2022 at the University of Vienna. Alternatively one can take a look at the introductory book "A First Course in Ergodic Theory", see [5].

We start by defining a general measurable dynamical system and some often used properties of such systems.

Definition 1.6. We call $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T)$ a (measurable) dynamical system, if

- $(\mathbb{X}, \mathcal{A}, \lambda)$ is a measure space
- $T : \mathbb{X} \longrightarrow \mathbb{X}$ is measurable.

If T is null preserving, i.e. $\lambda(A)=0$ implies $\lambda(T^{-1}A)=0$ for all $A\in\mathcal{A}$, we call $\mathcal{S}$ a null-preserving dynamical system.
If T is measure preserving, i.e. $\lambda(T^{-1}A)=\lambda(A)$ for all $A\in\mathcal{A}$, we call $\mathcal{S}$ a measure-preserving dynamical system.
If additionally $\lambda(\mathbb{X})=1$, we call $\mathcal{S}$ a probability-preserving dynamical system.
We call a measure $\mu << \lambda$ invariant with respect to T, if $\mu\circ T^{-1}=\mu$.

If not stated otherwise, we assume from now on that measures are finite or σ-finite and the system $\mathcal{S}$ is null-preserving.

For a system $\mathcal{S}$ we may look at the effect the map T has on densities with respect to the measure λ via the transfer operator (also called the Perron-Frobenius operator).

Definition 1.7. For a dynamical system $\mathcal{S}=(\mathbb{X},\mathcal{A},\lambda,T)$ we define the transfer operator $\hat{T}:L^1(\lambda)\to L^1(\lambda)$ as the operator, which pushes forward densities, i.e. for a signed measure $\nu << \lambda$ with density $u:=\frac{d\nu}{d\lambda}$ we define

$$\hat{T}u=\frac{d(\nu\circ T^{-1})}{d\lambda}.$$

In the above definition we use the Radon-Nikodym Theorem to define $\hat{T}u$. Every $u\in L^1(\lambda)$ corresponds to a signed measure $\nu(A):=\int\limits_A u\,d\lambda$ and since the system $\mathcal{S}$ is null preserving we get that the density $\frac{d(\nu\circ T^{-1})}{d\lambda}$ exists and is unique in $L^1(\lambda)$.
Below we are going to list some important properties of the transfer operator. For further information see Chapter 6 in [5].

Proposition 1.8. *(i) $\hat{T}$ is positive and linear*

(ii) $\hat{T}$ preserves the integral, i.e. $\int\limits_{\mathbb{X}} u\,d\lambda=\int\limits_{\mathbb{X}}\hat{T}u\,d\lambda$

(iii) $\|\hat{T}\|=\sup\{\frac{\|\hat{T}u\|_{L^1(\lambda)}}{\|u\|_{L^1(\lambda)}}:\,u\neq 0\}=1$

(iv) $\mu << \lambda$ is an invariant measure iff $\frac{d\mu}{d\lambda}=\hat{T}\frac{d\mu}{d\lambda}$

(v) Take $f\in L_\infty(\lambda)$ and $u\in L_1(\lambda)$, then $\int\limits_{\mathbb{X}}(f\circ T)u\,d\lambda=\int\limits_{\mathbb{X}} f\,\hat{T}u\,d\lambda$

Remark 1.9. If one knows, that the transfer operator for a system $\mathcal{S}=(\mathbb{X},\mathcal{A},\lambda,T)$ is given by $\hat{T}:L^1(\lambda)\to L^1(\lambda)$, one can infer that the transfer operator
$P:L^1(\nu)\to L^1(\nu)$ of the system $(\mathbb{X},\mathcal{A},\nu)$ for some $\nu\sim\lambda$ is given by

$$Pu(x)=\frac{\hat{T}(uh)(x)}{h(x)}\text{ for }\nu\text{ a.e. }x\in\mathbb{X},$$

where $h:=\frac{d\nu}{d\lambda}$.

We need to restrict general measurable dynamical systems further to be able to derive the later needed results. For this we define the following properties.

Definition 1.10. Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T)$ be a dynamical system. We call a set $W \in \mathcal{A}$ wandering for $\mathcal{S}$, if almost no point $x \in W$ returns to W, i.e. $\lambda(W \cap \bigcup_{n\geq 1} T^{-n}W) = 0$.
We call a set $A \in \mathcal{A}$ recurrent, if almost every point $x \in A$ returns to A, i.e. $A \subseteq \bigcup_{n\geq 1} T^{-n}A \mod \lambda$.

Definition 1.11. Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T)$ be a dynamical system. We call $\mathcal{S}$

(i) ergodic, if $T^{-1}A = A$ for $A \in \mathcal{A}$ implies $0 \in \{\lambda(A), \lambda(A^c)\}$.

(ii) conservative, if all wandering sets $W \in \mathcal{A}$ have measure 0.

One can show that in (i) we could exchange the condition $T^{-1}A = A$ implies $0 \in \{\lambda(A), \lambda(A^c)\}$ with the (as it seems) more restricting condition $T^{-1}A = A \mod \lambda$ implies $0 \in \{\lambda(A), \lambda(A^c)\}$ and would get an equivalent definition.
Next we include an important characterization of ergodicity. The theorem and its proof can be found in the lecture notes of the course on ergodic theory given by Roland Zweimüller in the fall term 2021/2022 at the University of Vienna or alternatively in [16].

Theorem 1.12. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T)$ be a null-preserving dynamical system. $\mathcal{S}$ is ergodic iff for all probability densities u, $v \in \mathcal{D}(\lambda)$ it holds that $\|\frac{1}{n}\sum_{j=0}^{n-1} \hat{T}^j u - \frac{1}{n}\sum_{j=0}^{n-1} \hat{T}^j v\|_{L^1(\lambda)} \longrightarrow 0$ as $n \to \infty$.*

1.3 Asymptoticallly T-invariant sequences and their long-time behaviour

We are going to look at sequences of functions R_n for which the distance $d(R_n, R_n \circ T)$ converges in measure to 0.

Definition 1.13. Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T)$ be a dynamical system, and let
$R_n : \mathbb{X} \longrightarrow \mathbb{M}$, for $n \in \mathbb{N}$ be measurable functions into some metrizable space $(\mathbb{M}, d_{\mathbb{M}})$ endowed with the Borel σ-algebra. We call $(R_n)_{n\geq 1}$ asymptotically T-invariant in measure or just asymptotically T-invariant, if for all $P << \lambda$ and for all $\varepsilon > 0$,
$P(d_{\mathbb{M}}(R_n, R_n \circ T) > \varepsilon) \to 0$ as $n \to \infty$.

For the following theorem recall Remark 1.5. Its proof can be found in [15, pp. 17–19].

Theorem 1.14. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic probability-preserving system, $(\mathbb{M}, d_{\mathbb{M}})$ a compact separable metric space endowed with the Borel σ-algebra $\mathcal{B}_{\mathbb{M}}$ and $R_l : \mathbb{X} \longrightarrow \mathbb{M}$ an asymptotically T-invariant sequence of measurable maps. Furthermore let $\mathcal{U}$ be a compact set in the space $(\mathcal{P}(\mu), d_{\mathcal{P}})$. Then*

$$D_{\mathbb{M}}(\nu \circ R_l^{-1}, \tilde{\nu} \circ R_l^{-1}) \longrightarrow 0 \tag{1.3}$$

as $l \to \infty$, uniformly for $\nu, \tilde{\nu} \in \mathcal{U}$.

This implies that if $R_l \overset{\nu}{\Longrightarrow} R$ for some $\nu \in \mathcal{U}$ then $R_l \overset{\nu_l}{\Longrightarrow} R$ for all sequences $(\nu_l)_{l\geq 1} \subseteq \mathcal{U}$. Before we start with the proof we include a remark which states that if a sequence of maps converges pointwise to a constant map and the sequence is equicontinuous it converges uniformly.

Remark 1.15. Let $(\mathfrak{K}, d_{\mathfrak{K}})$ be a compact metric space and $f_n : \mathfrak{K} \to \mathfrak{K}$, $n \geq 1$ a sequence of maps which converges pointwise to a constant map $c : \mathfrak{K} \to \mathfrak{K}$, where $c(x) = \tilde{c}$ for all $x \in \mathfrak{K}$. If $(f_n)_{n\geq 1}$ is equicontinuous, then $f_n \overset{n\to\infty}{\longrightarrow} k$ uniformly in $\mathfrak{K}$.
Since $\mathfrak{K}$ is compact, $(f_n)_{n\geq 1}$ is uniformly equicontinuous. Let $\varepsilon > 0$ be arbitrary, then there exists $\delta > 0$ such that for all x, y for which $d_{\mathfrak{K}}(x, y) < \delta$ holds, $d_{\mathfrak{K}}(f_n(x), f_n(y)) < \frac{\varepsilon}{2}$ follows for all $n \geq 1$. Since $\mathfrak{K}$ is compact, there exists a finite δ-dense subset $\{x_1, ..., x_k\}$ of $\mathfrak{K}$. Because of pointwise convergence we find $n_0 \in \mathbb{N}$ such that $d_{\mathfrak{K}}(f_n(x_i), \tilde{c}) < \frac{\varepsilon}{2}$, for all $n \geq n_0$ and for all $i \in \{1, ..., k\}$.
If we now take $y \in \mathfrak{K}$ arbitrary, then there exists $i \in \{1, ..., k\}$ such that $d_{\mathfrak{K}}(y, x_i) < \delta$ and therefore $d_{\mathfrak{K}}(f_n(y), \tilde{c}) \leq d_{\mathfrak{K}}(f_n(y), f_n(x_i)) + d_{\mathfrak{K}}(f_n(x_i), \tilde{c}) < \varepsilon$ for all $n \geq n_0$.

Proof of Theorem 1.14. We are going to show that for any Lipschitz function χ in $\overline{B_{\mathcal{C}(\mathbb{M})}}$ and every $\varepsilon > 0$, there exists $l_0 \in \mathbb{N}$, such that for all $l \geq l_0$ and for all $\nu, \tilde{\nu} \in \mathcal{U}$

$$\left| \int_{\mathbb{X}} \chi \circ R_l \, d\nu - \int_{\mathbb{X}} \chi \circ R_l \, d\tilde{\nu} \right| < \varepsilon \tag{1.4}$$

holds.
If we assume equation (1.4) then (1.3) follows from the definition of $D_{\mathbb{M}}$.
We now want to show (1.4). Take $\varepsilon > 0$. We want to use an approximation of $\int_{\mathbb{X}} \chi \circ R_l \, d\nu$ and $\int_{\mathbb{X}} \chi \circ R_l \, d\tilde{\nu}$ by $\int_{\mathbb{X}} \chi \circ R_l \, d(\frac{1}{N} \sum_{j=0}^{N-1} \nu \circ T^{-j})$ and $\int_{\mathbb{X}} \chi \circ R_l \, d(\frac{1}{N} \sum_{j=0}^{N-1} \tilde{\nu} \circ T^{-j})$ for some N large enough, which follows from the T-invariance of R_l.

From the ergodicity of $\mathcal{S}$ we infer that $\|\frac{1}{n} \sum_{j=0}^{n-1} \hat{T}^j (u - \tilde{u})\|_1$ converges to 0 as $n \to \infty$, for all u, $\tilde{u} \in \mathcal{D}(\mu)$.
With Remark 1.4 it follows that $f_n(\nu)$, where $f_n : \mathcal{P}(\mu) \to \mathcal{P}(\mu)$ is defined as $f_n(\nu) := \frac{1}{n} \sum_{j=0}^{n-1} \nu \circ T^{-j}$, converges to $f(\nu) = \mu$ for every $\nu \in \mathcal{P}(\mu)$. (The density of $\mu \in \mathcal{P}(\mu)$ with respect to μ is $\mathbb{1}_{\mathbb{X}}$, which is invariant with respect to $\hat{T}$ since μ is the invariant measure and so $\|\frac{1}{n} \sum_{j=0}^{n-1} \hat{T}^j u - \mathbb{1}_{\mathbb{X}}\|_1 \to 0$ for all $u \in \mathcal{D}(\mu)$.)

The dual operator on $\mathcal{D}(\mu)$, $\tilde{f}_n(u) := \frac{1}{n} \sum_{j=0}^{n-1} \hat{T}^j u$, has norm equal to 1 for all $n \geq 1$, so we can use Remark 1.4 to get that every f_n has norm less than 2 and therefore the family $(f_n)_{n\geq 1}$ is equicontinuous. Now we can use Remark 1.15 to infer that f_n converges to f uniformly in $\mathcal{P}(\mu)$ and that there exists

$N_0 \in \mathbb{N}$ such that

$$
\begin{aligned}
&\Big| \int\limits_{\mathbb{X}} \chi \circ R_l \, d\Big(\frac{1}{N} \sum_{j=0}^{N-1} \nu \circ T^{-j} \Big) - \int\limits_{\mathbb{X}} \chi \circ R_l \, d\Big(\frac{1}{N} \sum_{j=0}^{N-1} \tilde{\nu} \circ T^{-j} \Big) \Big| \\
&\leq d_{\mathcal{P}} \left(\frac{1}{N} \sum_{j=0}^{N-1} \nu \circ T^{-j}, \frac{1}{N} \sum_{j=0}^{N-1} \tilde{\nu} \circ T^{-j} \right) < \frac{\varepsilon}{4} \qquad (1.5)
\end{aligned}
$$

for all $N \geq N_0$ and for all $\nu, \tilde{\nu} \in \mathcal{P}(\mu)$. The first inequality follows from Remark 1.4 and from $|\chi| \leq 1$.

Because of Remark 1.4, $\mathcal{U}$ is compact in $\mathcal{P}(\mu)$ iff $\tilde{\mathcal{U}} := \{\frac{d\nu}{d\mu} : \nu \in \mathcal{U}\}$ is compact in $L^1(\mu)$. Hence $\tilde{\mathcal{U}}$ is uniformly integrable. There exists $\delta > 0$ such that

$$
\nu(A) = \int\limits_{A} \frac{d\nu}{d\mu} \, d\mu < \frac{\varepsilon}{8N_0} \qquad (1.6)
$$

for every $\nu \in \mathcal{U}$ and every $A \in \mathcal{A}$ for which $\mu(A) < \delta$ holds.

Since R_l is asymptotically T-invariant, there exists $l_1 \in \mathbb{N}$ such that $\mu(d_{\mathbb{M}}(R_l, R_l \circ T) > (c_\chi)^{-1} \frac{\varepsilon}{8N_0}) < \delta$, for all $l \geq l_1$, where c_χ is the Lipschitz-constant of χ. We infer

$$
\begin{aligned}
&\mu(|\chi \circ R_l \circ T^k - \chi \circ R_l \circ T^{k+1}| > \frac{\varepsilon}{8N_0}) = \mu(|\chi \circ R_l - \chi \circ R_l \circ T| > \frac{\varepsilon}{8N_0}) \qquad (1.7) \\
&\leq \mu(d_{\mathbb{M}}(R_l, R_l \circ T) > (c_\chi)^{-1} \frac{\varepsilon}{8N_0}) < \delta,
\end{aligned}
$$

where the first equality holds because T is measure preserving.

From Inequalities (1.6) and (1.7) and since $\|\chi\|_\infty \leq 1$ it follows that

$$
\begin{aligned}
&\int\limits_{\mathbb{X}} |\chi \circ R_l \circ T^k - \chi \circ R_l \circ T^{k+1}| \, d\nu \\
&\leq \frac{\varepsilon}{8N_0} + \int\limits_{\{|\chi \circ R_l \circ T^k - \chi \circ R_l \circ T^{k+1}| > \frac{\varepsilon}{8N_0}\}} |\chi \circ R_l \circ T^k - \chi \circ R_l \circ T^{k+1}| \, d\nu < \frac{3\varepsilon}{8N_0} \qquad (1.8)
\end{aligned}
$$

for all $l \geq l_1$ and all $\nu \in \mathcal{U}$.

We infer that for any $\nu \in \mathcal{U}$ and for all $l \geq l_1$

$$\begin{aligned}
&\Big| \int_{\mathbb{X}} \chi \circ R_l \, d\nu - \int_{\mathbb{X}} \chi \circ R_l \, d(N_0^{-1} \sum_{j=0}^{N_0-1} \nu \circ T^{-j}) \Big| \\
&\leq N_0^{-1} \sum_{j=0}^{N_0-1} \Big| \int_{\mathbb{X}} (\chi \circ R_l - \chi \circ R_l \circ T^j) \, d\nu \Big| \\
&\leq N_0^{-1} \sum_{j=0}^{N_0-1} \sum_{i=0}^{j-1} \int_{\mathbb{X}} |\chi \circ R_l \circ T^i - \chi \circ R_l \circ T^{i+1}| \, d\nu \\
&< N_0^{-1} \sum_{j=0}^{N_0-1} \sum_{i=0}^{j-1} \frac{3\varepsilon}{8N_0} \\
&\leq \frac{3\varepsilon}{8}
\end{aligned}$$

In the last step we use the triangle inequality and what we have shown above to get (1.4) uniformly in ν and $\tilde{\nu}$ for all $l \geq \max\{l_0, l_1\}$:

$$\begin{aligned}
&\Big| \int_{\mathbb{X}} \chi \circ R_l \, d\nu - \int_{\mathbb{X}} \chi \circ R_l \, d\tilde{\nu} \Big| \leq \Big| \int_{\mathbb{X}} \chi \circ R_l \, d\nu - \int_{\mathbb{X}} \chi \circ R_l \, d(\frac{1}{N} \sum_{j=0}^{N-1} \nu \circ T^{-j}) \Big| \\
&+ \Big| \int_{\mathbb{X}} \chi \circ R_l \, d(\frac{1}{N} \sum_{j=0}^{N-1} \nu \circ T^{-j}) - \int_{\mathbb{X}} \chi \circ R_l \, d(\frac{1}{N} \sum_{j=0}^{N-1} \tilde{\nu} \circ T^{-j}) \Big| + \Big| \int_{\mathbb{X}} \chi \circ R_l \, d(\frac{1}{N} \sum_{j=0}^{N-1} \tilde{\nu} \circ T^{-j}) - \int_{\mathbb{X}} \chi \circ R_l \, d\tilde{\nu} \Big| \\
&< \frac{3\varepsilon}{8} + \frac{\varepsilon}{4} + \frac{3\varepsilon}{8} = \varepsilon.
\end{aligned}$$

This proves Equation (1.4) and concludes the proof. □

In the next theorem we are going to use the above result to get equivalent notions of convergence for asymptotically T-invariant sequences. Later this will be useful to prove several theorems about hitting- and return-time statistics.

Theorem 1.16. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system, $(\mathbb{M}, d_{\mathbb{M}})$ a compact separable metric space endowed with the Borel σ-algebra $\mathcal{B}_{\mathbb{M}}$ and $R_l : \mathbb{X} \longrightarrow \mathbb{M}$ an asymptotically T-invariant sequence of measurable maps. The following are equivalent:*

(i) there exists $P \in \mathcal{P}(\mu)$ such that $R_l \overset{P}{\Longrightarrow} R$

(ii) for all $P \in \mathcal{P}(\mu)$ holds that $R_l \overset{P}{\Longrightarrow} R$

(iii) there exists a sequence $(P_l)_{l\geq 1}$ in $\mathcal{P}(\mu)$ with densities $(u_l)_{l\geq 1} \subseteq \mathfrak{K}$, for some compact $\mathfrak{K} \subseteq \mathcal{D}(\mu)$ such that $R_l \overset{P_l}{\Longrightarrow} R$

(iv) for every sequence $(P_l)_{l\geq 1}$ in $\mathcal{P}(\mu)$ with densities $(u_l)_{l\geq 1} \subseteq \mathfrak{K}$, for some compact $\mathfrak{K} \subseteq \mathcal{D}(\mu)$ such that $R_l \overset{P_l}{\Longrightarrow} R$

for some random element $R : \mathbb{X} \to \mathbb{M}$ and some probability measure $\mathbb{P}$ on $(\mathbb{X}, \mathcal{A})$.

Proof. We first show the equivalence of (i) and (ii). The implication from (ii) to (i) is clear, so let us assume (i) holds.
Let $P \in \mathcal{P}(\mu)$ be as in (i) and consider any other $Q \in \mathcal{P}(\mu)$. By Remark 1.2 $R_l \overset{P}{\Longrightarrow} R$ is equivalent to $\int_{\mathbb{X}} \tilde{\chi} \circ R_l \, dP \to \int_{\mathbb{X}} \tilde{\chi} \circ R \, d\mathbb{P}$ for all bounded Lipschitz-continuous functions $\tilde{\chi}$ and the same holds for $R_l \overset{Q}{\Longrightarrow} R$.
It therefore suffices to check that $|\int_{\mathbb{X}} \tilde{\chi} \circ R_l \, dP - \int_{\mathbb{X}} \tilde{\chi} \circ R_l \, dQ| \to 0$ as $l \to \infty$, to get that $R_l \overset{Q}{\Longrightarrow} R$.
But for every $\tilde{\varepsilon} > 0$ we can use the estimates from the proof of Theorem 1.14 by taking $\chi := \frac{\tilde{\chi}}{\|\tilde{\chi}\|_\infty}$, $\nu = P$, $\tilde{\nu} = Q$ and $\varepsilon = \frac{\tilde{\varepsilon}}{\|\tilde{\chi}\|_\infty}$ to show that for every $\tilde{\varepsilon} > 0$ there exists $l_0 \in \mathbb{N}$ such that $|\int_{\mathbb{X}} \tilde{\chi} \circ R_l \, dP - \int_{\mathbb{X}} \tilde{\chi} \circ R_l \, dQ| < \tilde{\varepsilon}$ for all $l \geq l_0$. This proves the equivalence of (i) and (ii).

Now we will show that (ii) implies (iv) and (iii) implies (i). This will conclude the proof since the implication of (iv) to (iii) is again clear.
Assume (ii) and let $(P_l)_{l\geq 1} \subseteq \mathcal{P}(\mu)$ be such that the densities u_l of P_l are contained in a compact set $\mathfrak{K}$. By Remark 1.4 this means that $(P_l)_{l\geq 1}$ is a subset of some compact set $\mathcal{U} \subseteq \mathcal{P}(\mu)$.
For every fixed $l_0 \in \mathbb{N}$ $R_l \overset{P_{l_0}}{\Longrightarrow} R$, i.e. $D_{\mathbb{M}}(P_{l_0} \circ R_l^{-1}, \mathbb{P} \circ R^{-1}) \to 0$ as $l \to \infty$. This, in combination with Theorem 1.14, gives us that $D_{\mathbb{M}}(P_l \circ R_l^{-1}, \mathbb{P} \circ R^{-1}) \to 0$, since

$$D_{\mathbb{M}}(P_l \circ R_l^{-1}, \mathbb{P} \circ R^{-1}) \leq D_{\mathbb{M}}(P_l \circ R_l^{-1}, P_{l_0} \circ R_l^{-1}) + D_{\mathbb{M}}(P_{l_0} \circ R_l^{-1}, \mathbb{P} \circ R^{-1})$$

and the first term on the right hand side tends to 0 uniformly in $\mathcal{U} \supseteq (P_l)_{l\geq 1}$.

The only implication we still need to show is $(iii) \Rightarrow (i)$.
Let $(P_l)_{l\geq 1}$ and $\mathfrak{K}$ be as in (iii). Then by Remark 1.4 $(P_l)_{l\geq 1} \subseteq \mathcal{U}$ for some compact $\mathcal{U} \subseteq \mathcal{P}(\mu)$. By Theorem 1.14 it follows that for every fixed l_0 the right hand side of the inequality

$$D_{\mathbb{M}}(P_{l_0} \circ R_l^{-1}, \mathbb{P} \circ R^{-1}) \leq D_{\mathbb{M}}(P_{l_0} \circ R_l^{-1}, P_l \circ R_l^{-1}) + D_{\mathbb{M}}(P_l \circ R_l^{-1}, \mathbb{P} \circ R^{-1})$$

converges to 0 and therefore $R_l \overset{P_{l_0}}{\Longrightarrow} R$. □

1.4 Hitting- and Return-time Statistics of a sequence of rare events

In the following chapters we are interested in return and hitting times of a point x into a given set $A \in \mathcal{A}$. More to this content can be found in [15]. We would like to know how many times one has to apply T if we start in a point $x \in \mathbb{X}$ to be in A. For this we define the function φ_A as follows:

Definition 1.17. For a dynamical system $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T)$ and a set $A \in \mathcal{A}$ we define $\varphi_A : \mathbb{X} \longrightarrow \overline{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$ with $\varphi_A(x) := \inf\{n \in \mathbb{N}, n \geq 1 : T^n x \in A\}$.
We call

(i) $\varphi_A : A \longrightarrow \overline{\mathbb{N}}$ the first return time and $T_A : A \cap \{x \in \mathbb{X} : \varphi_A(x) < \infty\} \longrightarrow A$ defined as $T_A x = T^{\varphi_A(x)} x$ the first return map.

(ii) $\varphi_A : \mathbb{X} \longrightarrow \overline{\mathbb{N}}$ the first hitting/ first entrance time and
$T_A : \{x \in \mathbb{X} : \varphi_A(x) < \infty\} \longrightarrow A$ defined as $T_A x = T^{\varphi_A(x)} x$ the first entrance map.

Note that φ_A is measurable and has some useful properties included below.

Remark 1.18. (i) For $A \in \mathcal{A}$ and every $x \in T^{-1}A^c$, $\varphi_A \circ T(x) = \varphi_A(x) - 1$ and more general $\varphi_A \circ T^j(x) = \varphi_A(x) - j$ for all $x \in \{\varphi_A \geq j+1\}$.

(ii) If $\mathcal{S}$ is ergodic and probability-preserving and $\lambda(A) > 0$, then $\varphi_A(x) < \infty$ for almost every $x \in \mathbb{X}$. (See Theorem 2.1.1, Corollary 2.1.1 and Proposition 2.2.2 in Chapter 2 in [5])

Furthermore we define a sequence A_l of rare events as follows.

Definition 1.19. Let $(\mathcal{A}, \lambda)$ be a measure space and let $(A_l)_{l\geq 1} \subset \mathcal{A}$ be a sequence of sets. We call $(A_l)_{l\geq 1}$ a sequence of rare events, if $\lim_{l\to\infty} \lambda(A_l) = 0$ and $\lambda(A_l) > 0$ for all $l \geq 1$.

Later we are going to need that for a sequence $(A_l)_{l\geq 1}$ of rare events, the sequence $(\lambda(A_l)\varphi_{A_l})_{l\geq 1}$ is asymptotically T-invariant in measure.

Remark 1.20. Assume we have an ergodic probability-preserving system $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \mu, T)$ and a sequence of rare events $(A_l)_{l\geq 1}$, then we get that $|\mu(A_l)\varphi_{A_l} - \mu(A_l)\varphi_{A_l} \circ T|$ converges in measure to 0. Fix $\varepsilon > 0$ arbitrary. Then there exists $l_0 \in \mathbb{N}$ such that for all $l \geq l_0$ we have that $\mu(A_l) < \varepsilon$. From Remark 1.18 we infer

$$\mu(|\mu(A_l)\varphi_{A_l} - \mu(A_l)\varphi_{A_l} \circ T| > \varepsilon) \leq \mu(|\varphi_{A_l} - \varphi_{A_l} \circ T| > 1) \leq \mu(\varphi_{A_l} - 1 \neq \varphi_{A_l} \circ T) \leq \mu(T^{-1}A_l) < \varepsilon$$

where we used in the last step that μ is invariant.
We conclude that $(\mu(A_l)\varphi_{A_l})_{l\geq 1}$ is asymptotically T-invariant.

In the following theorems we are going to talk about weak-convergence of probability measures and about whether or not the random elements $R_l := \mu(A_l)\varphi_{A_l}$ converge in distribution to some other random element R. We will look at the convergence with respect to the measure μ and with respect to the measures μ_{A_l}, where the first one corresponds to the hitting-time statistics and the second one to the return-time statistics. For more details on the different notions of convergence of measures and random variables see Definition 1.1 and Chapter 1 in [2].

The following Remark is an immediate consequence of the Generalized Kac formula (see Proposition A.2) which uses the correlation of the invariant measure of a system $\mathcal{S}$ and the corresponding induced system $\mathcal{S}|_A$ (see Chapter 11.3 in [5]).

Remark 1.21. From Proposition A.2 $\mu(\varphi_A = n) = \mu(A \cap \{\varphi_A \geq n\})$ follows by taking $f = \mathbb{1}_{\{\varphi_A = n\}}$.

$$\mu(\{\varphi_A = n\}) = \int_{\mathbb{X}} \mathbb{1}_{\{\varphi_A = n\}} d\mu \overset{A.2}{=} \int_A \sum_{j=0}^{\varphi_A - 1} \mathbb{1}_{\{\varphi_A = n\}} \circ T^j d\mu$$

$$= \int_A \sum_{j=0}^{\varphi_A - 1} \mathbb{1}_{\{\varphi_A \circ T^j = n\}} d\mu = \int_A \sum_{j=0}^{\varphi_A - 1} \mathbb{1}_{\{\varphi_A = n+j\}} d\mu = \mu(A \cap \{\varphi_A \geq n\})$$

The last equality holds since

$$\sum_{j=0}^{\varphi_A - 1} \mathbb{1}_{\{\varphi_A = n+j\}}(x) = \begin{cases} 1, & if\ \varphi_A \geq n \\ 0, & if\ \varphi_A < n \end{cases} .$$

With the following lemma we will be able to prove a Theorem about convergence properties of hitting- and return-time statistics. It will be used to prove that the hitting-time distributions converge iff the return-time distributions do and even shows a relation between the two limit distributions.

Lemma 1.22 (Return- versus Hitting-time distribution)**.** *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \mu, T)$ be a conservative measure preserving ergodic dynamical system and $A \in \mathcal{A}$ such that $\mu(A) \in (0, \infty)$. Let $t \geq 0$, then*

$$\mu(\mu(A)\varphi_A \leq t) \leq \int_0^t \mu_A(\mu(A)\varphi_A > s) ds \leq \mu(\mu(A)\varphi_A \leq t) + \mu(A). \tag{1.9}$$

Here μ_A is defined as $\mu_A(Y) := \frac{1}{\mu(A)}\mu(Y \cap A)$.

Proof. For the proof we use Remark 1.21:

$$\mu(\mu(A)\varphi_A \leq t) = \sum_{i=1}^{\lfloor t/\mu(A) \rfloor} \mu(\varphi_A = i) = \sum_{i=1}^{\lfloor t/\mu(A) \rfloor} \mu(\{\varphi_A \geq i\} \cap A)$$

$$= \int_0^{\lfloor t/\mu(A) \rfloor} \mu(\{\varphi_A > s\} \cap A) ds \leq \int_0^{t/\mu(A)} \mu(\{\varphi_A > s\} \cap A) ds.$$

With a change of variables $r = s\mu(A)$ we get that this is equal to $\int_0^t \mu_A(\mu(A)\varphi_A > r)dr$ and therefore show the first inequality.

If we now switch back to $s = \mu(A)r$ we get the second inequality:

$$\int_0^{\lfloor t/\mu(A) \rfloor} \mu(\{\varphi_A > s\} \cap A) ds + \int_{\lfloor t/\mu(A) \rfloor}^{t/\mu(A)} \mu(\{\varphi_A > s\} \cap A) ds$$

$$\leq \mu(\mu(A)\varphi_A \leq t) + \mu(A).$$

□

As mentioned before, we now state the first theorem about convergence of hitting- and return-times and show that if one converges then so does the other. Because of Kac-formula we know that $\int\limits_A \varphi_A d\mu_A = \frac{\mu(\mathbb{X})}{\mu(A)}$. To get convergence of the return-time distributions we thus have to normalize the return-time by the factor $\mu(A)$, i.e. we investigate the random variables $\mu(A)\varphi_A$.
The following theorem and its proof can be found in [7, 3 ff].

Theorem 1.23 (Hitting- vs. Return-time Statistics)**.** *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, T, \mu)$ be an ergodic probability preserving system, $(A_l)_{l\geq 1}$ a sequence of rare events in $\mathcal{A}$ and define $R_l := \mu(A_l)\varphi_{A_l}$. Then*

$$R_l \overset{\mu}{\Longrightarrow} R\,, \tag{1.10}$$

for some random variable R iff

$$R_l \overset{\mu_{A_l}}{\Longrightarrow} \tilde{R}\,, \tag{1.11}$$

for some random variable $\tilde{R}$.
In this case, the sub-distribution functions $F(t) := \Pr(R \leq t)$ and $\tilde{F}(t) := \Pr(\tilde{R} \leq t)$ are related by

$$F(t) = \int_0^t (1 - \tilde{F}(s))ds\,, \tag{1.12}$$

for all $t \geq 0$.

Proof. From Helly's selection Theorem it follows that there exist a subsequence $(l_k)_{k\geq 1}$ and distribution functions F, $\tilde{F}$ such that $\mu(R_{l_k} \leq t) \overset{k\to\infty}{\longrightarrow} F(t)$ for all continuity points t of F and $\mu_{A_{l_k}}(R_{l_k} \leq t) \overset{k\to\infty}{\longrightarrow} \tilde{F}(t)$ for all continuity points t of $\tilde{F}$.
Let $t \geq 0$ be a continuity point of F, then by Lemma 1.22

$$\mu(R_{l_k} \leq t) \leq \int\limits_0^t (1 - \mu_{A_{l_k}}(R_{l_k} \leq s))ds \leq \mu(R_{l_k} \leq t) + \mu(A_{l_k}).$$

Since $|1 - \mu_{A_{l_k}}(R_{l_K} \leq s)| \leq 1$ we can use the dominated convergence theorem and the fact that $\tilde{F}$ can only have countably many points of discontinuity, to infer that

$$F(t) \leq \int\limits_0^t (1 - \tilde{F}(s))ds \leq F(t).$$

So for each pair $(F, \tilde{F})$ of accumulation points equation (1.12) holds.
If we assume (1.10), then all accumulation points $\tilde{F}$ coincide, since each $\tilde{F}$ is right-continuous and therefore uniquely determined by equation (1.12). Now since all accumulation points coincide, distributional convergence as in (1.11) follows.
If we conversely assume (1.11), then all accumulation points F coincide, since F is uniquely determined by (1.12) and therefore (1.10) follows. $\square$

Now we can easily prove the following statement about an expontential limit distribution of hitting- and return-times.

Theorem 1.24. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic probability-preserving dynamical system, $(A_l)_{l\geq 1} \subset \mathcal{A}$ be a sequence of rare events and $\mathcal{E}$ an exponentially distributed random variable with expected value 1. TFAE:*

(i) $R_l \overset{\mu}{\Longrightarrow} \mathcal{E}$

(ii) $R_l \overset{\mu_{A_l}}{\Longrightarrow} \mathcal{E}$

(iii) There exists a random variable R with values in $[0,\infty]$ such that $R_l \overset{\mu}{\Longrightarrow} R$ and $R_l \overset{\mu_{A_l}}{\Longrightarrow} R$.

Proof. The equivalence follows directly from Theorem 1.2, Theorem 1.23 and the fact that, if $F : [0,\infty) \longrightarrow [0,1]$ is a distribution function which fulfills $F(t) = \int_0^t (1 - F(s))ds$, then $F(t) = 1 - e^{-t}$. $\square$

From now on $\mathcal{E}$ will always denote an exponentially distributed random variable with expected value 1.

We showed that if one of the time statistics converges to an exponential distribution then so does the other. In the next theorem we find sufficient conditions for convergence to an exponential distribution.

Theorem 1.25. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic probability preserving dynamical system and $(A_l)_{l\geq 1} \subset \mathcal{A}$ be a sequence of rare events. If there exists a sequence $(\tau_l)_{l\geq 1} \subseteq \mathbb{N}$ such that*

(i) $\mu(A_l)\tau_l \overset{l\to\infty}{\longrightarrow} 0$

(ii) $\mu_{A_l}(\varphi_{A_l} \leq \tau_l) \overset{l\to\infty}{\longrightarrow} 0$

(iii) $\hat{T}^{\tau_l}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l}) \in \mathfrak{K}$ *for all $l \geq 1$ and some compact set $\mathfrak{K} \subset L^1(\mu)$.*

Then $R_l \overset{\mu}{\Longrightarrow} \mathcal{E}$ and $R_l \overset{\mu_{A_l}}{\Longrightarrow} \mathcal{E}$.

Intuitively one could explain conditions (ii) and (iii) as follows: τ_l corresponds to the number of jumps before we look at R_l, i.e. we look at the convergence in distribution with respect to μ_{A_l} of the random variables $R_l \circ T^{\tau_l}$. (ii) guarantees that the mass of the points which return to A_l before the τ_l-iteration can be neglected and (iii) ensures us that the new starting distribution $\hat{T}^{\tau_l}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})$ is in a compact set so that we can use Theorem 1.16 to get that $R_l \circ T^{\tau_l} \overset{\mu_{A_l}}{\Longrightarrow} R$ implies $R_l \overset{\mu}{\Longrightarrow} R$.

Remark 1.26. In the following proof we need Theorem 1.16. The assumption there is that R_l maps to a compact target space. So we want to compactify $[0,\infty]$ as follows:
We are going to consider $[0,\infty]$ with the metric $d(x,y) := |e^{-x} - e^{-y}|$. This gives us a compact space and the trace of the corresponding Borel σ-algebra $\mathcal{B}_{[0,\infty]}$ on $[0,\infty)$ coincides with the usual Borel σ-algebra on $[0,\infty)$.

Proof of Theorem 1.25. Let $(F_l)_{l\geq 1}$ be the sequence of sub-distribution functions of R_l with respect to μ and $(\tilde{F}_l)_{l\geq 1}$ be the sequence of sub-distribution functions of R_l with respect to μ_{A_l}. We use that a sequence $(F_l)_{l\geq 1}$, for which every subsequence has a convergent subsequence which converges to F, itself converges to F. Furthermore we will show that every accumulation point of the sequences F_l and $\tilde{F}_l$ coincides. Then we get with Theorem 1.24 that $R_l \overset{\mu}{\Longrightarrow} \mathcal{E}$ and $R_l \overset{\mu_{A_l}}{\Longrightarrow} \mathcal{E}$.

By Helly's selection Theorem we can choose without loss of generality a subsequence $(R_{l_j})_{j\geq 1}$ such that $R_{l_j} \overset{\mu}{\Longrightarrow} R$ and $R_{l_j} \overset{\mu_{A_{l_j}}}{\Longrightarrow} \tilde{R}$. We want to show that $R \overset{d}{=} \tilde{R}$.

Let $\varepsilon > 0$, then we can compare φ_{A_l} and $\varphi_{A_l} \circ T^{\tau_l}$ on $\{\varphi_{A_l} > \tau_l\}$ by Remark 1.18 as follows:

$$\varphi_{A_l} - \varphi_{A_l} \circ T^{\tau_l} = \tau_l$$

Now we can multiply the above equation by $\mu(A_l)$ and get the following estimate

$$\begin{aligned}
&\mu_{A_l}(|R_l - R_l \circ T^{\tau_l}| > \varepsilon) \\
&= \mu_{A_l}(\{\varphi_{A_l} \leq \tau_l\} \cap \{|R_l - R_l \circ T^{\tau_l}| > \varepsilon\}) + \mu_{A_l}(\{\varphi_{A_l} > \tau_l\} \cap \{|R_l - R_l \circ T^{\tau_l}| > \varepsilon\}) \\
&\leq \mu_{A_l}(\varphi_{A_l} \leq \tau_l) + \mu_{A_l}(\{\varphi_{A_l} > \tau_l\} \cap \{\mu(A_l)\tau_l > \varepsilon\})
\end{aligned}$$

and by (i) and (ii) the right hand side converges to 0 as $l \to 0$.
With the above we have that $|R_{l_j} - R_{l_j} \circ T^{\tau_{l_j}}| \overset{\mu_{A_{l_j}}}{\Longrightarrow} 0$. Therefore $R_{l_j} \overset{\mu_{A_{l_j}}}{\Longrightarrow} \tilde{R}$ implies $R_{l_j} \circ T^{\tau_{l_j}} \overset{\mu_{A_{l_j}}}{\Longrightarrow} \tilde{R}$.
Now we can use that by (iii) $\mu_{A_{l_j}} \circ T^{-\tau_{l_j}}$ has density $\hat{T}^{\tau_{l_j}}(\frac{1}{\mu(A_{l_j})}\mathbb{1}_{A_{l_j}}) \in \mathfrak{K}$ and apply Theorem 1.16 to get that $R_{l_j} \overset{\mu}{\Longrightarrow} \tilde{R}$.
From Theorem 1.24 it follows that $\tilde{R} \overset{d}{=} \mathcal{E}$ and we can conclude $R_l \overset{\mu}{\Longrightarrow} \mathcal{E}$ and $R_l \overset{\mu_{A_l}}{\Longrightarrow} \mathcal{E}$. □

2 Folklore Systems and their Hitting- and Return-time Statistics

In this chapter we are going to look at a special family of dynamical systems called Folklore systems and want to show that for some special sequences of rare events $(A_l)_{l\geq 1}$ the hitting and return time statistics are exponentially distributed.
Most of the following can be found in in the lecture notes of the course on ergodic theory given by Roland Zweimüller in 2021/2022 at the University of Vienna. In this chapter $\mathbb{X}$ will always be a bounded interval in $\mathbb{R}$ which is a countable union $(\mathrm{mod}\,\lambda)$ of open intervals and λ will always denote the one-dimensional Lebesgue measure.
We start by defining what a Folklore system is.

Definition 2.1. (i) $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T, \xi)$ is called a <u>Folklore system</u> if $\mathcal{S}$ is a measurable dynamical system, $\mathbb{X} \subseteq \mathbb{R}$ is a compact interval and $\mathcal{A}$ is the induced Borel σ-algebra on $\mathbb{X}$. Moreover T is piecewise C^2, i.e. there exists a countable partition ξ $(\mathrm{mod}\,\lambda)$ of $\mathbb{X}$ into open intervals, such that for every $Z \in \xi$ the map $T|_Z : Z \longrightarrow TZ$ is invertible and $T|_Z$ has a C^2 extension to $\overline{Z}$.

Furthermore

(F1) T is uniformly expanding, i.e. $\inf_{x\in\mathbb{X}}|T'x|>1$

(F2) Adler's condition holds, i.e. $A_T := \sup_{x\in\mathbb{X}}|\frac{T''x}{(T'x)^2}|<\infty$

(F3) $\forall Z\in\xi$: $TZ=\mathbb{X}$.

(ii) For $n\in\mathbb{N}$ we define ξ_n to be the set of all cylinders of degree n. If $Z\in\xi_n$ then it is of the form $Z=[Z_0,...,Z_{n-1}]=\cap_{i=0}^{n-1}T^{-i}Z_i$, for $Z_i\in\xi=\xi_1$, $0\le i\le n-1$.

On the cylinder sets we can look at the inverse of T restricted to the cylinder set. This gives us the possibility to write down the transfer operator explicitly.

Remark 2.2. (i) Condition (F1) is equivalent to the condition that there exists $\rho\in(0,1)$ such that $\sup_{Z\in\xi}\sup_{x\in\mathbb{X}}|v_Z'|\le\rho<1$, where we define $v_Z:\mathbb{X}\longrightarrow Z$ as $(T|_Z)^{-1}$.

(ii) For $Z=[Z_0,...,Z_{j-1}]\in\cup_{n\ge1}\xi_n$ we define $v_Z:\mathbb{X}\longrightarrow Z$ as $(T^j|_Z)^{-1}$. Then $v_Z=v_{Z_0}\circ v_{Z_1}\circ...\circ v_{Z_{j-1}}$ and therefore

$$|v_Z'|=\Pi_{i=0}^{j-1}|v_{Z_i}'|\circ(v_{Z_{i+1}}\circ...\circ v_{Z_{j-1}})=\Pi_{i=0}^{j-1}|v_{Z_i}'|\circ v_{[Z_{i+1},...,Z_{j-1}]}\le\rho^j,$$

where we choose ρ as in (i). We infer that $\lambda(Z)=\int_{v_Z(\mathbb{X})}d\lambda=\int_{\mathbb{X}}|v_Z'|d\lambda\le\lambda(\mathbb{X})\rho^n$.

(iii) Let $f:\mathbb{X}\to\mathbb{R}$ be in $L_\infty(\mathbb{X})$ and $u\in L^1(\lambda)$. Then we compute $\hat{T}$ for the system $\mathcal{S}=(\mathbb{X},\mathcal{A},\lambda,T,\xi)$ as follows

$$\begin{aligned}\int_{\mathbb{X}}(f\circ T)u\,d\lambda &= \sum_{Z\in\xi}\int_Z(f\circ T)u\,d\lambda=\sum_{Z\in\xi}\int_{\mathbb{X}}f(u\circ v_Z)|v_Z'|\,d\lambda\\ &=\int_{\mathbb{X}}f(\sum_{Z\in\xi}(u\circ v_Z)|v_Z'|)d\lambda.\end{aligned}$$

Here we used the substitution for Lebesgue integrals in the second step and dominated convergence to interchange sum and integral in the last step.
With the above and Remark 1.8 we conclude that $\hat{T}u=\sum_{Z\in\xi}(u\circ v_Z)|v_Z'|$ is the transfer operator with respect to λ.

We start with some general properties of Folklore systems which we state without proof. It can be found in the lecture notes of the course on ergodic theory given by Roland Zweimüller in 2021/2022 at the University of Vienna.

Theorem 2.3. *Let $\mathcal{S}=(\mathbb{X},\mathcal{A},\lambda,T,\xi)$ be a Folklore system. Then the following holds:*

(i) $\mathcal{S}$ is ergodic

(ii) $\mathcal{S}$ is conservative

(iii) $\mathcal{S}$ *has an invariant probability density* $h = \hat{T}h \in \mathcal{C}^1(\mathbb{X})$ *with* $\mathcal{R}_{\mathbb{X}}(h) \leq A_T^\infty$.

The derivatives of the above defined functions v_Z have a nice property which we call finite regularity. It will be used later to show that $\bigcup_{n\geq 1}\{\hat{T}^n(\frac{1}{\mu(Z)}\mathbb{1}_Z) : Z \in \xi_n\}$ is contained in a compact set. First we define what regularity means.

Definition 2.4. Let $\mathbb{X} \subseteq \mathbb{R}$ be a bounded interval and $u : \mathbb{X} \longrightarrow (0,\infty)$ a $\mathbb{C}^1$-function. The (ir-)regularity of u on $\mathbb{X}$ is defined as

$$\mathcal{R}_{\mathbb{X}}(u) := \sup_{x\in\mathbb{X}}\frac{|u'(x)|}{u(x)} = \sup_{x\in\mathbb{X}}|(\log u)'(x)| \in [0,\infty] \text{ and } \mathcal{R}_{\mathbb{X}}(0) := 0.$$

For functions with regularity less than r, we get the following three estimates.

Lemma 2.5. *Let* $\mathbb{X} \subseteq \mathbb{R}$ *be a bounded interval and* $u : \mathbb{X} \longrightarrow \mathbb{R}$ *a* C^1*-function such that* $\mathcal{R}_{\mathbb{X}}(u) \leq r < \infty$. *Then*

(i) $u(x) \leq e^{r|x-y|}u(y)$, *for all* $x, y \in \mathbb{X}$

(ii) $\frac{1}{\lambda(\mathbb{X})}\int_{\mathbb{X}} ud\lambda \leq \sup_{x\in\mathbb{X}} u(x) \leq e^{r\lambda(\mathbb{X})}\inf_{x\in\mathbb{X}} u(x) \leq e^{r\lambda(\mathbb{X})}\frac{1}{\lambda(\mathbb{X})}\int_{\mathbb{X}} ud\lambda$

(iii) $|u(x) - u(y)| \leq re^{r\lambda(\mathbb{X})}\inf_{t\in\mathbb{X}} u(t)|x-y|$ *for all* $x, y \in \mathbb{X}$.

The third inequality tells us that u *is Lipschitz continuous.*

Proof. (i) From the Mean Value Theorem it follows, that for all $x, y \in \mathbb{X}$ there exists a $z \in \mathbb{X}$ such that

$$\frac{|\log(u(x)) - \log(u(y))|}{|x-y|} = |(\log u)'(z)| = |\frac{u'(z)}{u(z)}| \leq r.$$

Now inequality (i) follows by exponentiation.

(ii) The only inequality we need to show is $\sup_{x\in\mathbb{X}} u(x) \leq e^{r\lambda(\mathbb{X})}\inf_{x\in\mathbb{X}} u(x)$. From (i) we infer that for all $x, y \in \mathbb{X}$

$$u(x) \leq e^{r|x-y|}u(y) \leq e^{r\lambda(\mathbb{X})}u(y).$$

We now take the infimum over all $y \in \mathbb{X}$ and the supremum over all $x \in \mathbb{X}$ and get (ii).

(iii) For this equation we again use the Mean Value Theorem and equation (ii).

$$\frac{|u(x) - u(y)|}{|x-y|} = |u'(z)| = \frac{|u'(z)|}{u(z)}u(z) \leq ru(z) \leq r\sup_{t\in\mathbb{X}} u(t) \overset{(ii)}{\leq} re^{r\lambda(\mathbb{X})}\inf_{t\in\mathbb{X}} u(t)$$

□

As noted before we show that v_Z' has regularity less then a constant A_T^∞ for arbitrary cylinder sets Z.

Lemma 2.6. *Let* $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T, \xi)$ *be a Folklore system and* $Z = [Z_0, ..., Z_{n-1}] \in \xi_n$ *for some* $n \geq 1$. *Then* $\mathcal{R}_{\mathbb{X}}(|v_Z'|) \leq A_T^\infty$, *with* $A_T^\infty := \frac{A_T}{1-\rho}$.

Proof. We are going to use Remark 2.2, then

$$
\begin{aligned}
|\frac{v''_Z}{v'_Z}| &= |(\log(v'_Z))'| = |(\log(\Pi_{i=0}^{n-1} v'_{Z_i} \circ (v_{Z_{i+1}} \circ \ldots \circ v_{Z_{k-1}})))'| \\
&= |\sum_{i=0}^{n-1}(\log(v'_{Z_i} \circ v_{[Z_{i+1},\ldots,Z_{n-1}]}))'| \leq \sum_{i=0}^{n-1} |\frac{v''_{Z_i} \circ v_{[Z_{i+1},\ldots,Z_{n-1}]}}{v'_{Z_i} \circ v_{[Z_{i+1},\ldots,Z_{n-1}]}}||v'_{[Z_{i+1},\ldots,Z_{n-1}]}| \\
&= \sum_{i=0}^{n-1} |\frac{v''_{Z_i} \circ v_{[Z_{i+1},\ldots,Z_{n-1}]}}{(v'_{Z_i} \circ v_{[Z_{i+1},\ldots,Z_{n-1}]})^2}||v'_{Z_i} \circ v_{[Z_{i+1},\ldots,Z_{n-1}]} \cdot v'_{[Z_{i+1},\ldots,Z_{n-1}]}| \\
&\leq A_T \sum_{i=0}^{n-1} \rho^{n-i} \leq \frac{A_T}{1-\rho}.
\end{aligned}
$$

□

The following lemma gives us information about how much of a cylinder set of degree n can hit a set A after n iterations, relative to the size of the cylinder set and the size of A. This will be used to show (ii) of the assumptions in Theorem 1.25 while proving that hitting- and return-time distributions of cylinder sets converge to an exponential distribution.

Lemma 2.7. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T, \xi)$ be a Folklore system and $\mu << \lambda$ the corresponding invariant probability measure. Then for all $A \in \mathcal{A}$ and for all $Z \in \xi_n$, $n \geq 1$ the following inequality holds:*

$$\frac{\mu(Z \cap T^{-n}A)}{\mu(A)\mu(Z)} \leq e^{3A_T^\infty \lambda(\mathbb{X})}. \tag{2.1}$$

Proof. We are going to use (ii) from Lemma 2.5 and let v_Z be as in Remark 2.2. Let u be the density of μ with respect to λ. By (iii) of Theorem 2.3 we get that $\mathcal{R}_{\mathbb{X}}(u) \leq A_T^\infty$ and infer

$$
\begin{aligned}
\mu(T^{-n}A \cap Z) = \mu(v_Z(A)) &= \int_{v_Z(A)} u d\lambda = \int_A (u \circ v_Z)|v'_Z| d\lambda \\
&\leq \sup_{x \in \mathbb{X}} u(x) \sup_{x \in \mathbb{X}} |v'_Z(x)| \lambda(A) \overset{2.2.(\mathrm{ii})}{\leq} e^{2A_T^\infty \lambda(\mathbb{X})} \inf_{x \in \mathbb{X}} u(x) \inf_{x \in \mathbb{X}} |v'_Z(x)| \int_A d\lambda \\
&\leq e^{2A_T^\infty \lambda(\mathbb{X})} \inf_{x \in \mathbb{X}} |v'_Z(x)| \mu(A) \frac{\lambda(\mathbb{X})}{\lambda(\mathbb{X})} \leq e^{2A_T^\infty \lambda(\mathbb{X})} \mu(A) \frac{1}{\lambda(\mathbb{X})} \int_{\mathbb{X}} |v'_Z| d\lambda \\
&= e^{2A_T^\infty \lambda(\mathbb{X})} \mu(A) \frac{1}{\lambda(\mathbb{X})} \lambda(Z) \leq e^{2A_T^\infty \lambda(\mathbb{X})} \mu(A) \frac{1}{\lambda(\mathbb{X})} \lambda(Z) \frac{\sup_{x \in \mathbb{X}} u(x)}{\frac{1}{\lambda(\mathbb{X})}} \\
&\overset{2.2.(\mathrm{ii})}{\leq} e^{3A_T^\infty \lambda(\mathbb{X})} \mu(A) \inf_{x \in \mathbb{X}} u(X) \lambda(Z) \leq e^{3A_T^\infty \lambda(\mathbb{X})} \mu(A) \mu(Z).
\end{aligned}
$$

□

The next two remarks provide us with the missing arguments, such that we are able to prove our first two statements of hitting- and return-time limits of some special sequences of rare events for Folklore systems. We will need them to show that (iii) in Theorem 1.25 holds with the assumptions we make

on the rare events.
In the first remark we show that $\hat{T}^n(\frac{1}{\mu(Z)}\mathbb{1}_Z)$ has bounded regularity for a cylinder set Z of degree n.

Remark 2.8. Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T, \xi)$ be a Folklore system, $\mu << \lambda$ the corresponding invariant probability measure and take $Z \in \xi_n$ for some $n \geq 1$. Let $\hat{T}$ denote the transfer-operator with respect to the invariant measure μ, i.e. for the system $(\mathbb{X}, \mathcal{A}, \mu, T)$. Then $\mathcal{R}_{\mathbb{X}}(\hat{T}^n(\frac{1}{\mu(Z)}\mathbb{1}_Z) \leq 3A_T^{\infty}$.
Let us denote by ν_Z the measure on $\mathcal{A}$ which has density $\mathbb{1}_Z$ with respect to μ. Then $\nu_Z << \mu << \lambda$ and $\lambda << \mu$ (see Theorem 2.3 part (iii)). We can use the properties of Radon Nikodym derivatives to get that almost everywhere

$$\begin{aligned}\hat{T}^n\mathbb{1}_Z &= \frac{d(\nu_Z \circ T^{-n})}{d\mu} = \frac{d(\nu_Z \circ T^{-n})}{d\lambda}\frac{d\lambda}{d\mu} = \sum_{\tilde{Z}\in\xi_n}((\frac{d\nu_Z}{d\lambda} \circ v_{\tilde{Z}})|v'_{\tilde{Z}}|)\frac{1}{u} \\ &= \sum_{\tilde{Z}\in\xi_n}(((\mathbb{1}_Z u) \circ v_{\tilde{Z}})|v'_{\tilde{Z}}|)\frac{1}{u} = (u \circ v_Z)|v'_Z|\frac{1}{u}.\end{aligned}$$

Therefore exists a version of $\hat{T}^n(\frac{1}{\mu(Z)}\mathbb{1}_Z)$ which is C^1. Now we get that

$$\begin{aligned}&\sup_{x\in\mathbb{X}}\frac{|((u \circ v_Z)|v'_Z|\frac{1}{u})'(x)|}{(u \circ v_Z)(x)|v'_Z(x)|\frac{1}{u(x)}} \\ &\leq \sup_{x\in\mathbb{X}}\frac{\left|(u' \circ v_z)(x)|v'_z(x)|^2\frac{1}{u(x)} + |(u \circ v_Z)(x)||v''_Z(x)|\frac{1}{u(x)}\right| + (u \circ v_Z)(x)|v'_Z(x)|\frac{1}{u^2(x)}|u'(x)|}{(u \circ v_Z)(x)|v'_Z(x)|\frac{1}{u(x)}} \\ &\leq \sup_{x\in\mathbb{X}}\frac{|(u' \circ v_Z)(x)|}{(u \circ v_Z(x))}|v'_Z(x)| + \sup_{x\in\mathbb{X}}\frac{|v''_Z(x)|}{|v'_Z(x)|} + \sup_{x\in\mathbb{X}}\frac{|u'(x)|}{u(x)} \\ &\leq 3A_T^{\infty}\end{aligned}$$

and therefore $\mathcal{R}_{\mathbb{X}}(\hat{T}^n(\frac{1}{\mu(Z)}\mathbb{1}_Z) \leq 3A_T^{\infty}$.

In the next remark we are going to note that the densities $\hat{T}^n(\frac{1}{\mu(Z)}\mathbb{1}_Z)$ are contained in a compact set.

Remark 2.9. Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T, \xi)$ be a Folklore system, $\mu << \lambda$ the corresponding invariant probability measure and define $\mathcal{F}_{\mathbb{X}}(r) := \{f \in \mathcal{C}^1(\mathbb{X}) : f > 0,\ \mathcal{R}_{\mathbb{X}}(f) \leq r \text{ and } \int_{\mathbb{X}} f\, d\mu = 1\}$. We infer from part (ii) of Lemma 2.5 that

$$\begin{aligned}\sup_{f\in\mathcal{F}_{\mathbb{X}}(r)}\|f\|_{\infty} &\leq \sup_{f\in\mathcal{F}_{\mathbb{X}}(r)}\frac{e^{r\lambda(\mathbb{X})}}{\lambda(\mathbb{X})}\int_{\mathbb{X}} f\, d\lambda \\ &\leq \frac{1}{\inf_{x\in\mathbb{X}} u(x)}\sup_{f\in\mathcal{F}_{\mathbb{X}}(r)}\frac{e^{r\lambda(\mathbb{X})}}{\lambda(\mathbb{X})}\int_{\mathbb{X}} f\, d\mu = \frac{1}{\inf_{x\in\mathbb{X}} u(x)}\frac{e^{r\lambda(\mathbb{X})}}{\lambda(\mathbb{X})}.\end{aligned}$$

We know from Theorem 2.3 and from Lemma 2.5 that $\inf_{x\in\mathbb{X}}\frac{1}{u(x)} \in (0, \infty)$. Therefore $\mathcal{F}_{\mathbb{X}}(r)$ is uniformly bounded and by part (iii) of Lemma 2.5 equicontinuous since all $f \in \mathcal{F}_{\mathbb{X}}(r)$ have r as a Lipschitz-constant. From the Arzela-Ascoli Theorem it follows that $\mathcal{F}_{\mathbb{X}}(r)$ is relatively compact in $\mathcal{C}(\mathbb{X})$. Because

$\mu(\mathbb{X})$ is finite we conclude that $\mathcal{F}_{\mathbb{X}}(r)$ is relatively compact in $L^1(\mu)$.

By Remark 2.8 we have that $\{\hat{T}^n(\frac{1}{\mu(Z)}\mathbb{1}_Z) : Z \in \bigcup_{n\geq 1}\xi_n\} \subseteq \mathcal{F}_{\mathbb{X}}(3A_T^\infty)$ and are now able to prove the following theorem.

Theorem 2.10. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T, \xi)$ be a Folklore system, $\mu << \lambda$ the corresponding invariant probability measure and $(A_l)_{l\geq 1}$ a rare sequence of events, such that for each $l \geq 1$ the set A_l is of the form $\bigcup_{Z\in\xi_{A_l}} Z \mod \lambda$, where $\xi_{A_l} := \{Z \in \xi : Z \subseteq A_l\}$. Then for $R_l := \mu(A_l)\varphi_{A_l}$ it holds that*

(i) $R_l \overset{\mu}{\Longrightarrow} \mathcal{E}$

(ii) $R_l \overset{\mu_{A_l}}{\Longrightarrow} \mathcal{E}$.

Proof. We are going to use Theorem 1.25 and show that $(\tau_l)_{l\geq 1} = (1)_{l\geq 1}$ fulfills assumptions (i)-(iii). With Remark 1.20 and Remark 1.26 the other assumptions in Theorem 1.25 are fulfilled.

(i) $\tau_l\mu(A_l) = \mu(A_l) \longrightarrow 0$ as $l \to \infty$.

(ii) For the second assumption we are going to use Lemma 2.7 to get the following estimate

$$\begin{aligned}\mu_{A_l}(\varphi_{A_l} \leq \tau_l) &= \mu_{A_l}(\varphi_{A_l} \leq 1) = \frac{1}{\mu(A_l)}\mu(\bigcup_{Z\in\xi_{A_l}} Z \cap T^{-1}A_l) \\ &= \frac{1}{\mu(A_l)}\sum_{Z\in\xi_{A_l}} \mu(T^{-1}A_l \cap Z) \overset{2.7}{\leq} e^{3A_T^\infty\lambda(\mathbb{X})}\sum_{Z\in\xi_{A_l}}\mu(Z) = e^{3A_T^\infty\lambda(\mathbb{X})}\mu(A_l)\end{aligned}$$

where the right side converges to 0 as $l \to \infty$.

(iii) We are going to show that $\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l}) \in \mathcal{F}_{\mathbb{X}}(3A_T^\infty)$. For this to be true we only need to show that $\mathcal{R}_{\mathbb{X}}(\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})) \leq 3A_T^\infty$.

$$\begin{aligned}\sup_{x\in\mathbb{X}}\frac{|(\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l}))'(x)|}{\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})(x)} &\leq \sup_{x\in\mathbb{X}}\frac{\sum_{Z\in\xi_{A_l}}\frac{1}{\mu(A_l)}|(\hat{T}\mathbb{1}_Z)'(x)|}{\sum_{Z\in\xi_{A_l}}\frac{1}{\mu(A_l)}\hat{T}\mathbb{1}_Z(x)} \\ &= \sup_{x\in\mathbb{X}}\frac{\sum_{Z\in\xi_{A_l}}\frac{1}{\mu(A_l)}\frac{|(\hat{T}\mathbb{1}_Z)'(x)|}{\hat{T}\mathbb{1}_Z(x)}\hat{T}\mathbb{1}_Z(x)}{\sum_{Z\in\xi_{A_l}}\frac{1}{\mu(A_l)}\hat{T}\mathbb{1}_Z(x)} \overset{2.8}{\leq} 3A_T^\infty\end{aligned}$$

With Remark 2.9 we conclude part (iii) of the proof.

$\square$

Below we state another theorem which is similar to the above, only that we now look at the cylinder sets around a point x^* as our sequence of rare events.

Definition 2.11. For a null-preserving dynamical system $(\mathbb{X}, \mathcal{A}, \lambda, T)$ with a partition ξ of $\mathbb{X}$ (mod λ) into open sets and a point $x^* \in \mathbb{X}$ we denote by $\xi_l(x^*) \in \xi_l$ the cylinder set of degree l containing x^* (i.e. $x^* \in \xi_l(x^*)$). This set exists for a.e. point $x^* \in \mathbb{X}$.

Theorem 2.12. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T, \xi)$ be a Folklore system, $\mu << \lambda$ the corresponding invariant probability measure and suppose that $x^* \in \mathbb{X}$ is a non periodic point such that for all $l \geq 1$ $\xi_l(x^*)$ is well-defined. Let $(A_l)_{l\geq 1}$ be a rare sequence of events given by $A_l = \xi_l(x^*) \mod \lambda$. Then it holds for $R_l := \mu(A_l)\varphi_{A_l}$ that*

(i) $R_l \overset{\mu}{\Longrightarrow} \mathcal{E}$

(ii) $R_l \overset{\mu_{A_l}}{\Longrightarrow} \mathcal{E}$.

Proof. Again we are going to use Theorem 1.25 and show that $(\tau_l)_{l\geq 1} = (l)_{l\geq 1}$ fulfills assumptions (i)-(iii). Let u be the density of μ with respect to λ.

(i) $\mu(\xi_l(x^*))l \leq \sup_{x\in\mathbb{X}} u(x)\lambda(\xi_l(x^*))l \overset{2.2}{\leq} \sup_{x\in\mathbb{X}} u(x)\rho^l l$ and this converges to 0 as $l \to \infty$.

(ii) We need to show $\mu_{\xi_l(x^*)}(\varphi_{\xi_l(x^*)} \leq l) \to 0$.
Since x^* is not periodic we know that for any $M \in \mathbb{N}$ fixed, each one of the points $Tx^*, ..., T^M x^*$ is different to x^*. The distance between these points is positive, i.e. $\min_{0\leq j<k\leq M} |T^j x^* - T^k x^*| \geq \varepsilon$ for some $\varepsilon > 0$. Since $\lambda(\xi_l(x^*)) \to 0$ and T, T^2, ..., T^M are continuous on $\xi_M(x^*)$, we can choose l_0 large enough such that

$$\xi_l(x^*) \cap (\bigcup_{j=1}^{M} T^j \xi_l(x^*)) = \emptyset$$

for all $l \geq l_0$. Then

$$\begin{aligned}
&\mu_{\xi_l(x^*)}(\varphi_{\xi_l(x^*)} \leq l) = \mu_{\xi_l(x^*)}(M < \varphi_{\xi_l(x^*)} \leq l) \\
&= \frac{1}{\mu(\xi_l(x^*))} \sum_{j=M+1}^{l} \mu(\xi_l(x^*) \cap \{\varphi_{\xi_l(x^*)} = j\}) \\
&\leq \frac{1}{\mu(\xi_l(x^*))} \sum_{j=M+1}^{l} \mu(\xi_j(x^*) \cap T^{-j}(\xi_l(x^*)) \\
&\overset{2.7}{\leq} e^{3A_T^\infty \lambda(\mathbb{X})} \sum_{j=M+1}^{l} \mu(\xi_j(x^*)) \leq \sup_{x\in\mathbb{X}} u(x) \frac{\rho^{M+1} e^{3A_T^\infty \lambda(\mathbb{X})}}{1-\rho}
\end{aligned}$$

and therefore $\mu_{\xi_l(x^*)}(\varphi_{\xi_l(x^*)} \leq l) \to 0$. We used in the third step that $\xi_l(x^*) \subseteq \xi_j(x^*)$ for all $j \leq l$.

(iii) By Remark 2.9 it follows that $\hat{T}^l(\frac{1}{\mu(\xi_l(x^*))}\mathbb{1}_{\xi_l(x^*)}) \in \mathcal{F}_{\mathbb{X}}(3A_T^\infty)$ which is relatively compact.

□

3 Rychlik Systems and their Hitting- and Return-time Statistics

In this section we find a more general family of dynamical systems, for which we are able to derive theorems about exponential hitting- and return-time statistics. In this chapter ν will always denote a regular Borel-measure on some compact set $\mathbb{X} \subseteq \mathbb{R}$.
The first way we derive exponential hitting- and return-time statistics is by using spectral theory and finding a subspace U of $L^1(\nu)$ such that $T : U \to U$ is quasi-compact. It turns out that the subspace $BV(\nu) \subseteq L^1(\nu)$ of functions with bounded variation fulfills this requirement. For this one can take a look at [4, pp. 7–10].
The second way we do this is more elementary in the sense that we will not use advanced spectral theory, but instead use a highly non-trivial inequality derived in [13, pp. 69–74] to verify the three conditions in Theorem 1.25.
We conclude the chapter by giving an example of a Rychlik system which does not fulfill the assumptions of the theorems about exponential hitting- and return-time statistics but still admits exponential hitting- and return-time limit distributions. Furthermore the given system $(\mathbb{X}, \mathcal{A}, \lambda, T)$ in the example itself is Rychlik but the system with respect to the absolutely continuous invariant measure $\mu << \lambda$, $(\mathbb{X}, \mathcal{A}, \mu, T)$ is not Rychlik. We analyze further whether or not the properties of the system with respect to μ are similar to the properties a Rychlik system admits.
As noted before one can find most of the content of the following chapter in [13] and [4].

3.1 Rychlik Systems

We start by defining the framework we will need later.

Definition 3.1. (i) The variation of a function f on a set $\mathbb{X} \subseteq \mathbb{R}$ is defined as
$V_{\mathbb{X}} f := \sup\{\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| : n \in \mathbb{N}, (x_0, ..., x_n) \in \mathbb{X}^n, x_0 \leq ... \leq x_n\}$.

(ii) $BV(\nu) := \{f \in L^1(\nu) :$ there exists $\tilde{f}$, s.t. $\tilde{f} = f$ a.e. and $\|\tilde{f}\|_{BV(\nu)} < \infty\}$ is the subspace of bounded variation. $(BV(\nu), \|\cdot\|_{BV(\nu)})$ is a Banach space with the norm $\|.\|_{BV(\nu)} := \|\cdot\|_{L^1(\nu)} + V_{\mathbb{X}} \cdot$.

Remark 3.2. (i) Whenever we write $V_{\mathbb{X}} f$ for a function $f \in L^1(\nu)$ we take $V_{\mathbb{X}} f$ to be $\min\{V_{\mathbb{X}} \tilde{f} : \tilde{f}$ is a version of $f\}$.

(ii) For $u, v \in BV(\nu)$ the following inequalities hold:

- $\|u\|_\infty \leq \|u\|_{BV(\nu)}$
- $\|uv\|_{BV(\nu)} \leq 2\|u\|_{BV(\nu)}\|v\|_{BV(\nu)}$.

We define a collection of dynamical systems called the Rychlik systems. They were first introduced by Marek Rychlik and generalize the family of dynamical systems studied by F. Hofbauer and G. Keller in [10] greatly. Instead of the assumption that $\mathbb{X}$ can be split into a finite number of pieces on which

the map T is homeomorphic to some interval, one now demands that $\mathbb{X}$ can be split into a countable number of such pieces.
The Rychlik systems include the Folklore systems which we discussed in the second Chapter.

Definition 3.3. Let $\mathbb{X} \subseteq \mathbb{R}$ be a set. We define the set $\mathcal{I}_{\mathbb{X}} := \{I \cap \mathbb{X} : I \subseteq \text{ is an interval}\}$ to be the set of $\mathbb{X}$-intervals.

Definition 3.4. Let $\mathbb{X} \subseteq \mathbb{R}$ be compact, and let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \nu, T, \xi)$ be a measurable dynamical system, $\mathcal{A}$ the induced Borel σ-algebra on $\mathbb{X}$ and ν a regular Borel probability measure on $\mathcal{A}$. We assume that there exists a countable partition ξ of $\mathbb{X}$ (mod ν) into open $\mathbb{X}$-intervals such that $\mathbb{X} = \bigcup\limits_{Z\in\xi} \bar{Z}$ and $Y := \mathbb{X} \setminus (\bigcup\limits_{Z\in\xi} Z)$ consists exactly of the endpoints of $Z \in \xi$.
We call $\mathcal{S}$ a Rychlik system if the following holds:

(R1) For any $Z \in \xi$, $T|_Z$ admits an extension to a homeomorphism from $\bar{Z}$ to some interval $T\bar{Z} \subseteq \mathbb{X}$

(R2) There exists a function $g : \mathbb{X} \to [0, \infty)$ such that $g|_Y = 0$ and $V_{\mathbb{X}}\, g < \infty$ and such that for the operator $\hat{T} : L^1(\nu) \to L^1(\nu)$, defined as $\hat{T}u(x) := \sum\limits_{y\in T^{-1}(x)} g(y)u(y)$, the equality $\int\limits_{\mathbb{X}} u\, d\nu = \int\limits_{\mathbb{X}} \hat{T}u\, d\nu$ holds. This function g is called the weight function.

(R3) T is expanding, i.e. $\sup\limits_{x\in\mathbb{X}} g(x) < 1$.

Remark 3.5. (i) One can show that $\hat{T}$ in Definition 3.4 coincides with the transfer operator.

(ii) We again write ξ_n for the set of all cylinders of degree n. $Z \in \xi_n$ is of the form $Z = [Z_0, ..., Z_{n-1}] = \cap_{i=0}^{n-1} T^{-i} Z_i$, for $Z_i \in \xi$, $0 \le i \le n-1$. In [13, p. 72] it is shown that $\nu(Z) \le \|g\|_\infty^n$ for $Z \in \xi_n$.

(iii) If one has a Rychlik system $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \nu, T, \xi)$ it follows that for every $x \in \mathbb{X}$, $\nu(\{x\}) = 0$.
We look at two possible cases. Either $x \in \mathbb{X}$ is such that the cylinder sets $\xi_n(x)$ are well defined for all $n \ge 1$ or not.
If $x \in \mathbb{X}$ is such that $\xi_n(x)$ are well defined, then $\nu(\{x\}) \le \nu(\xi_n(x)) \le \|g\|_\infty^n \to 0$ as $n \to \infty$.
If $x \in \mathbb{X}$ such that not all cylinder sets are well defined, it follows from the definition of Rychlik systems that $\nu(\{x\}) \le \nu(\cup_{n\ge 0} T^{-n}(\mathbb{X} \setminus Y)) = 0$, where $Y = \bigcup\limits_{Z\in\xi} Z$, since T is null preserving (See Theorem 3.6).

(iv) Folklore systems $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda, T, \xi)$ are Rychlik systems with weight function $g = \frac{1}{|T'|}$ on Y and $g = 0$ on $\mathbb{X} \setminus Y$.
We need to show $(R2)$:
We start by showing that $V_{\mathbb{X}} g$ is finite. We use that for any C^1-function f the equality $V f = \int |f'|\, d\lambda$ holds. Since $g|_Z$ is C^1 for any $Z \in \xi$, we get

$$V_{\mathbb{X}} g \le \sum_{Z\in\xi} V_Z g|_Z + 2 \sum_{Z\in\xi} \sup_{x\in Z} g|_Z(x) = \sum_{Z\in\xi} \int\limits_Z \frac{|T''|}{|T'|^2}\, d\lambda + 2 \sum_{Z\in\xi} \sup_{x\in Z} \frac{1}{|T'(x)|}$$
$$\overset{(F2)}{\le} \lambda(\mathbb{X}) A_T + 2 \sum_{Z\in\xi} \sup_{x\in Z} |v_Z'(x)|.$$

We are done if we show that $\sum_{Z\in\xi}\sup_{x\in Z}|v_Z'(x)|<\infty$. This follows from Lemma 2.5 and Lemma 2.6:

$$\sum_{Z\in\xi}\sup_{x\in Z}|v_Z'(x)| \overset{\text{Lemma2.5}}{\leq} 2\frac{e^{A_T^\infty\lambda(\mathbb{X})}}{\lambda(\mathbb{X})}\sum_{Z\in\xi}\underbrace{\int_{\mathbb{X}}|v_Z'|\,d\lambda}_{=\lambda(Z)} = 2\exp^{A_T^\infty\lambda(\mathbb{X})}<\infty.$$

From Remark 2.2 we know that $\hat{T}u=\sum_{Z\in\xi}(u\circ v_Z)|v_Z'|$, but $\sum_{y\in T^{-1}(x)}g(y)u(y)=\sum_{Z\in\xi}(u\circ v_Z)|v_Z'|$ and therefore g is as required in $(R2)$.

We will state some general facts about Rychlik systems. The proof of the following theorem uses advanced functional analysis and spectral theory and we are going to omit it here. For a detailed derivation of the Theorem including a proof see [13].

Theorem 3.6. *Let $\mathcal{S}=(\mathbb{X},\mathcal{A},\nu,T,\xi)$ be a Rychlik system. Then the following holds:*

(i) T is null preserving.

(ii) There exists an invariant probability measure $\mu<<\nu$ with density $\frac{d\mu}{d\nu}\in BV(\nu)$.

(iii) There exist only finitely many eigenvalues $\chi_1,...,\chi_N$ on the unit circle.

(iv) $\hat{T}:BV(\nu)\to BV(\nu)$ admits a representation

$$\hat{T}=\sum_{j=1}^{N}\chi_jQ_j+R,$$

where $Q_j:BV(\nu)\to BV(\nu)$ is the projection onto the eigenspace corresponding to the eigenvalue χ_j. The linear map $R:BV(\nu)\to BV(\nu)$ is such that $\inf_{k\in\mathbb{N}}\|R^k\|<\theta$ for some $\theta\in(0,1)$. Moreover $Q_jR=RQ_j=0$ and $Q_jQ_i=0$ for any $i,j\in\{1,..,N\}$.

(v) If ν is mixing, $\sigma(\hat{T})\cap S^1=\{1\}$ and $\hat{T}=Q+R$ on $BV(\nu)$, where $Q=Q_1$ and R are as in (ii).

3.2 Hitting- and Return-time Statistics of Rychlik Systems

Before we are able to state and prove the theorem about exponential hitting- and return-time statistics, we show a property of Rychlik systems inferred from [13, pp. 74–78]. It is called the Decay of correlations.

Theorem 3.7. *If $\mathcal{S}=(\mathbb{X},\mathcal{A},\nu,T,\xi)$ is a Rychlik system, such that $\mu<<\nu$ is an invariant probability measure which is mixing, i.e. $\mu(A\cap T^{-n}B)\to\mu(A)\mu(B)$ as $n\to\infty$ for all A, $B\in\mathcal{A}$. Then there exists a constant $C>0$ and $\rho\in(0,1)$ such that*

$$\left|\int_{\mathbb{X}}u\circ T^nf\,d\mu-\int_{\mathbb{X}}u\,d\mu\int_{\mathbb{X}}f\,d\mu\right|\leq C\rho^n\|u\|_{L^1(\nu)}\|f\|_{BV(\nu)} \tag{3.1}$$

for all $u \in L_1(\nu)$, $f \in BV(\nu)$.

Proof. Since μ is mixing, T^n is ergodic for all $n \in \mathbb{N}$. From Theorem 1, Theorem 3 and Theorem 4 in [13] it follows that $\hat{T} = Q + R$ on $BV(\nu)$, where $\hat{T}$ is the transfer operator with respect to ν, Q is the projection on the eigenspace $span(\mathbb{1}_{\mathbb{X}})$ corresponding to the eigenvalue 1 of $\hat{T}$ and R fulfills that $\varrho(R) = \lim_{k\to\infty} \|R^k\|^{\frac{1}{k}} < \theta \in (0,1)$. Moreover it also follows that μ has a density h with respect to ν, such that $h \in BV(\nu)$.
For each $n \in \mathbb{N}$ we get

$$\begin{aligned}
&\left|\int_{\mathbb{X}} u \circ T^n f \, d\mu - \int_{\mathbb{X}} u \, d\mu \int_{\mathbb{X}} f \, d\mu\right| = \left|\int_{\mathbb{X}} u\hat{T}^n \overbrace{(fh)}^{\in BV(\nu)} \, d\nu - \int_{\mathbb{X}} fh \, d\nu \int_{\mathbb{X}} uh \, d\nu\right| \\
&\leq \|u\|_{L^1(\nu)} \|\hat{T}^n(fh) - \underbrace{\int_{\mathbb{X}} fh \, d\nu h}_{=Q(fh)} \|_\infty \leq \|u\|_{L^1(\nu)} \underbrace{\|\hat{T}^n(fh) - Q^n(fh)\|_{BV(\nu)}}_{=\|R^n(fh)\|_{BV(\nu)}} \\
&\leq C\rho^n \|u\|_{L^1(\nu)} \|f\|_{BV(\nu)}
\end{aligned}$$

where we used in the last inequality that there exists $n_0 \in \mathbb{N}$, such that
$\rho := \sqrt[n_0]{\|R^{n_0}\|} < 1$ and for $n = kn_0 + r \in \mathbb{N}$, $r \in \{0, ..., n_0 - 1\}$

$$\|R^n\| \leq \|R^{n_0}\|^k \|R^r\| \leq \rho^{n_0 k} \underbrace{\max_{0 \leq r_0 < n_0} \|R^{r_0}\| \frac{1}{\rho^{n_0}}}_{=:\tilde{C}} \rho^r = \tilde{C}\rho^n$$

holds and that

$$\tilde{C}\rho^n \|fh\|_{BV(\nu)} \leq C\rho^n \|f\|_{BV(\nu)}$$

for some $C > 0$. □

Hirati, Saussol and Vaienti give in their paper 'Statistics of return times: A general framework and new applications' sufficient conditions to get an exponential limit distribution. We will later use these conditions to show that Rychlik maps have exponential limit law. For this we state the following lemma without proof. A detailed proof can be found in [9, pp. 5–9].
We define $c(A) := \sup_{n\in\mathbb{N}} |\mu_A(\varphi_A > n) - \mu(\varphi_A > n)|$.

Lemma 3.8. *Let* $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \mu, T)$ *be a probability preserving dynamical system and* $A \in \mathcal{A}$. *Then*

$$c(A) \leq \inf_{N\in\mathbb{N}} (a_N(A) + b_N(A) + N\mu(A)),$$

where $a_N(A) := \mu_A(\varphi_A \leq N)$ *and* $b_N(A) := \sup_{B\in\mathcal{A}} |\mu_A(T^{-N}B) - \mu(B)|$.

Remark 3.9. We know from Theorem 1.24 that if we have a sequence of rare events A_l and the corresponding hitting- and return-time statistics admit the same law, they are exponentially distributed.

If we do not assume that the statistics converge, but only assume that the hitting-time statistic and the return-time statistic are close, i.e. $c(A_l) \to 0$ as $l \to \infty$, we still get with Helly's selection Theorem that the distributions converge to an exponential one. It therefore is sufficient to show that $c(A_l)$ converges to 0 to get exponential hitting- and return-time statistics.

We are now able to state and prove the following theorem for Rychlik maps, which can be found in [4, pp. 9–10].

Theorem 3.10. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \nu, T, \xi)$ be a Rychlik system with an invariant mixing probability measure $\mu << \nu$. Let $x^* \in \mathbb{X}$ be a non periodic point such that for all $l \geq 1$ the set $\xi_l(x^*)$ is well-defined and that $(A_l)_{l\geq 1}$ is a sequence of rare events, where each $A_l \subseteq \mathbb{X}$ is an open ball centered at x^*. Then for μ almost every x^* and for $R_l := \mu(A_l) \cdot \varphi_{A_l}$ it holds that*

(i) $R_l \overset{\mu}{\Longrightarrow} \mathcal{E}$

(ii) $R_l \overset{\mu_{A_l}}{\Longrightarrow} \mathcal{E}$.

Proof. We are going to use Lemma 3.8 and Remark 3.9. Let us first look at the quantity a_N given in Lemma 3.8.
As in the Proof of Theorem 2.12 we argue, that for fixed $M \in \mathbb{N}$ we can choose $l_0 \in \mathbb{N}$ large enough, such that $A_l \cap (\cup_{j=1}^M T^j A_l) = \emptyset$ for all $l \geq l_0$. Since x^* is not periodic we know that for any $M \in \mathbb{N}$ fixed, each one of the points $Tx^*, ..., T^M x^*$ is different to x^* and the distance between these points is positive, i.e. $\min\limits_{0\leq j<k\leq M} |T^j x^* - T^k x^*| \geq \varepsilon$ for some $\varepsilon > 0$. Since A_l are intervals centered in x^* and $\nu(A_l) \to 0$ the assertion follows because T, $T^2, ..., T^M$ are continuous on $\xi_M(x^*)$. We infer for $N \geq M$

$$\begin{aligned}
a_N(A_l) = \mu_{A_l}(\varphi_{A_l} \leq N) &= \sum_{k=M}^{N} \mu_{A_l}(\underbrace{\varphi_{A_l} = k}_{\subseteq T^{-k}A_l}) \leq \sum_{k=M}^{N} \frac{1}{\mu(A_l)} \int\limits_{\mathbb{X}} \mathbb{1}_{A_l} \mathbb{1}_{A_l} \circ T^k \, d\mu \\
&= \sum_{k=M}^{N} \left(\frac{1}{\mu(A_l)} \int\limits_{\mathbb{X}} \mathbb{1}_{A_l} (\mathbb{1}_{A_l} \circ T^k) \, d\mu - \int\limits_{\mathbb{X}} \mathbb{1}_{A_l} d\mu \int\limits_{X} \frac{\mathbb{1}_{A_l}}{\mu(A_l)} d\mu + \mu(A_l) \right) \\
&\overset{3.7}{\leq} \sum_{k=M}^{N} C\rho^k \frac{\nu(A_l)}{\mu(A_l)} \|\mathbb{1}_{A_l}\|_{BV(\nu)} + N\mu(A_l) \leq 3C \frac{\nu(A_l)}{\mu(A_l)} \frac{\rho^M}{1-\rho} + N\mu(A_l),
\end{aligned}$$

since A_l is an interval and therefore $\|\mathbb{1}_{A_l}\|_{BV(\nu)} \leq 3$.
For b_N we get the following estimate:

$$\begin{aligned}
b_N(A_l) &= \sup_{B\in\mathcal{A}} |\mu_{A_l}(T^{-N}B) - \mu(B)| \\
&= \sup_{B\in\mathcal{A}} \left| \int\limits_{\mathbb{X}} \frac{\mathbb{1}_{A_l}}{\mu(A_l)} \mathbb{1}_B \circ T^N \, d\mu - \frac{1}{\mu(A_l)} \int\limits_{\mathbb{X}} \mathbb{1}_{A_l} \, d\mu \int\limits_{\mathbb{X}} \mathbb{1}_B \, d\mu \right| \\
&\overset{3.7}{\leq} \sup_{B\in\mathcal{A}} C\rho^N \nu(B) \frac{3}{\mu(A_l)} \leq \frac{3C}{\mu(A_l)} \rho^N.
\end{aligned}$$

Alltogether we now get with Lemma 3.8 that

$$c(A_l) \leq \inf_{N\in\mathbb{N}} \left(\frac{\nu(A_l)}{\mu(A_l)} \frac{3C}{1-\rho}\rho^M + N\mu(A_l) + \frac{3C}{\mu(A_l)}\rho^N + N\mu(A_l) \right).$$

We can now choose $N = \left\lceil \frac{2\log(\mu(A_l))}{\log(\rho)} \right\rceil$ and get

$$\begin{aligned} c(A_l) &\leq \frac{\nu(A_l)}{\mu(A_l)} \frac{3C}{1-\rho}\rho^M + 2\left\lceil \frac{2\log(\mu(A_l))}{\log(\rho)} \right\rceil \mu(A_l) + \frac{3C}{\mu(A_l)}\rho^{\frac{2\log(\mu(A_l))}{\log(\rho)}} \\ &\leq \frac{\nu(A_l)}{\mu(A_l)} \frac{3C}{1-\rho}\rho^M + 2\left\lceil \frac{2\log(\mu(A_l))}{\log(\rho)} \right\rceil \mu(A_l) + 3C\mu(A_l). \end{aligned}$$

The right hand side can be arbitrary small since M is arbitrary and for $l \to \infty$ the first term converges for μ almost every x^* to 0, since $\frac{d\mu}{d\nu}(x^*) > 0$ μ almost everywhere and since $\frac{d\mu}{d\nu} \in BV(\nu)$ has a version which is continuous except at countably many points, so especially μ almost everywhere and we therefore get

$$\frac{\nu(A_l)}{\mu(A_l)} \leq \frac{\nu(A_l)}{\nu(A_l)} \underbrace{\frac{1}{\operatorname*{ess\,inf}_{x\in A_l} \frac{d\mu}{d\nu}}}_{<\delta} < \infty$$

for some $\delta > 0$ and $l \geq l_0$ for some l_0 big enough. (Here ess inf denotes the essential infimum.)
The other terms also converge to 0 and with this convergence of R_l to the exponential distribution follows with Remark 3.9.

□

In the proof above we used Theorem 3.7, which follows from [13]. As noted before one uses spectral theoretical properties of the transfer operator to derive these statistical properties of the system. This is often employed, but it uses rather advanced fuctional analysis. One may ask the question, whether or not one could simplify the arguments in the proof in the sense that one uses less (or none) spectral theory.

3.3 Hitting- and Return-time Statistics of Rychlik Systems using Theorem 1.25

Next we will find an alternative proof to a similar theorem as Theorem 3.10. The idea is, to use Theorem 1.25 in the same manner as we already did in the proofs of Theorem 2.10 and Theorem 2.12. The key argument to verify (ii) in Theorem 1.25 is the so called Lasota-Yorke inequality. In [13] the inequality

$$V_{\mathbb{X}}(P^n(f)) \leq F\|f\|_{L^1(\nu)} + \chi^n V_{\mathbb{X}}(f) \tag{3.2}$$

for a Rychlik system $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \nu, T, \xi)$, where P denotes the transfer operator, is proven. M. Rychlik showed that the inequality holds for $n \geq n^*$ and $f \in BV(\nu)$ for some $F > 0$, $\chi \in (0,1)$ and $n^* \in \mathbb{N}$. The proof only uses elementary arguments but is nontrivial. We are going to omit the proof and use this inequality to show the Lasota-Yorke inequality as we need it. A proof can be found in [13][71-73]. More information on the arguments used in the following proof can be found in[3, 152 ff.].

Theorem 3.11 (Lasota-Yorke inequality)**.** *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \nu, T, \xi)$ be an ergodic Rychlik system with invariant probability measure $\mu << \nu$ with density $h := \frac{d\mu}{d\nu}$ such that there exists $k \geq 1$ such that $\frac{1}{k} \leq h \leq k$ holds. Then there exist constants $C > 0$, $\rho \in (0,1)$ and $n_0 \in \mathbb{N}$ such that for the transfer operator $\hat{T} : L^1(\mu) \to L^1(\mu)$ with respect to the measure μ and for all $n \geq n_0$*

$$V_{\mathbb{X}}(\hat{T}^n f) \leq C\|f\|_{L^1(\mu)} + \rho^n V_{\mathbb{X}}(f) \tag{3.3}$$

holds for every $f \in BV(\mu)$.

Proof. We first notice that for $u \in L^1(\mu)$, the transfer operator $\hat{T}$ is given by $\frac{P(uh)}{h}$, where P is the transfer operator with respect to the measure ν. Inductively the iterates of $\hat{T}$ are given by $\hat{T}^n(u) = \frac{P^n(uh)}{h}$.
We now take n^*, F and χ as needed in Equation (3.2). For $n \geq n^*$ we get for $f \in BV$

$$\begin{aligned}
V_{\mathbb{X}}(\hat{T}^n f) &= V_{\mathbb{X}}\left(\frac{P^n(fh)}{h}\right) \\
&\leq 2V_{\mathbb{X}}(1/h)V_{\mathbb{X}}(P^n(fh)) + V_{\mathbb{X}}(1/h)\|P^n(fh)\|_{L^1(\nu)} + V_{\mathbb{X}}(P^n(fh))\|1/h\|_{L^1(\mu)} \\
&\overset{(3.2)}{\leq} (2V_{\mathbb{X}}(1/h)+1)(F\|fh\|_{L^1(\nu)} + \chi^n V_{\mathbb{X}}(fh)) + V_{\mathbb{X}}(1/h)\|f\|_{L^1(\mu)} \\
&\leq (2V_{\mathbb{X}}(1/h)+1)(F\|f\|_{L^1(\mu)} + \chi^n(2V_{\mathbb{X}}(f)V_{\mathbb{X}}(h) + V_{\mathbb{X}}(h)\|f\|_{L^1(\mu)} + V_{\mathbb{X}}(f)) \\
&\quad + V_{\mathbb{X}}(1/h)\|f\|_{L^1(\mu)}
\end{aligned}$$

where we used in the first and last inequality that for $u \in BV(\mu) \subseteq L^1(\mu)$ and $v \in BV(\nu) \subseteq L^1(\nu)$ we get

$$\begin{aligned}
V_{\mathbb{X}}(uv) &\leq V_{\mathbb{X}}(u)\|v\|_\infty + V_{\mathbb{X}}(v)\|u\|_\infty \\
&\leq 2V_{\mathbb{X}}(u)V_{\mathbb{X}}(v) + V_{\mathbb{X}}(v)\|u\|_{L^1(\mu)} + V_{\mathbb{X}}(u)\|v\|_{L^1(\nu)}.
\end{aligned}$$

$V_{\mathbb{X}}(fh) \leq V_{\mathbb{X}}(f)\|h\|_\infty + V_{\mathbb{X}}(h)\|f\|_\infty$ holds since for $x, y \in \mathbb{X}$ we have that

$$\begin{aligned}
|f(x)h(x) - f(y)h(y)| &\leq |f(x)h(x) - f(y)h(x)| + |f(y)h(x) - f(y)h(y)| \\
&\leq \|h\|_\infty |f(x) - f(y)| + \|f\|_\infty |h(x) - h(y)|.
\end{aligned}$$

The second inequality follows from $\|\cdot\|_\infty \leq V_{\mathbb{X}}(\cdot) + \|\cdot\|_{L^1(\cdot)}$.
We now show that $V_{\mathbb{X}}(1/h)$ is finite. For $x, y \in \mathbb{X}$ we have

$$\left|\frac{1}{h(x)} - \frac{1}{h(y)}\right| = \left|\frac{h(y) - h(x)}{h(x)h(y)}\right| \leq k^2|h(x) - h(y)|$$

and therefore $V_{\mathbb{X}}(1/h) \leq k^2 V_{\mathbb{X}}(h)$.

So we get from the inequalities above that there exists $c_1, c_2 > 0$ such that

$$V_{\mathbb{X}}(\hat{T}^n f) \leq c_1 \|f\|_{L^1(\mu)} + c_2 \chi^n V_{\mathbb{X}}(f)$$

and we are done.

$\square$

We are now able to prove an analogous theorem to Theorem 3.10 with the help of Theorem 1.25. The following differs from Theorem 3.10 in the assumptions. We will not assume the invariant measure μ to be mixing, but instead assume that the density $h := \frac{d\mu}{d\nu}$ is bounded away from 0.

Theorem 3.12. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \nu, T, \xi)$ be an ergodic Rychlik system and the invariant probability measure $\mu << \nu$ of $\mathcal{S}$ such that there exists $k > 1$ such that for the density $h := \frac{d\mu}{d\nu}$, $\frac{1}{k} \leq h \leq k$ holds. We assume that $x^* \in \mathbb{X}$ is a non periodic point and that for all $l \geq 1$ $\xi_l(x^*)$ is well-defined. Furthermore let $(A_l)_{l\geq 1}$ be a sequence of rare events, where each A_l is a non-degenerate interval such that $\operatorname{dist}(A_l, x^*) \to 0$ and $\lambda(A_l) \to 0$ as $l \to \infty$. Then for $R_l := \mu(A_l)\varphi_{A_l}$ it holds that*

(i) $R_l \overset{\mu}{\Longrightarrow} \mathcal{E}$

(ii) $R_l \overset{\mu_{A_l}}{\Longrightarrow} \mathcal{E}$.

Proof. As already mentioned we want to verify conditions (i)-(iii) from Theorem 1.25. We first want to find a fitting sequence $(\tau_l)_{l\geq 1}$.
We define $\tau_l := \max\left\{ \underbrace{\left\lfloor \frac{(-\log \mu(A_l) + \log \frac{2}{C}}{-\log \rho} \right\rfloor + 1}_{=:t_l}, n_0 \right\}$, where ρ, n_0 and C are taken as in Theorem 3.11. It follows that

$$n \geq t_l \Rightarrow n \geq \frac{-\log \mu(A_l) + \log \frac{2}{C}}{-\log \rho} \Leftrightarrow$$
$$n \geq \frac{\log(\mu(A_l)\frac{C}{2})}{\log \rho} \Leftrightarrow n \log \rho \leq \log \frac{C\mu(A_l)}{2}$$

and via exponentiation we get $\rho^n \leq \frac{C\mu(A_l)}{2}$ for all $n \geq t_l$.

(iii) The set $\mathcal{U}(c) := \{u \in BV(\mu) \cap \mathcal{D}(\mu) : \, u \geq 0, \, V_{\mathbb{X}}(u) \leq c\}$ is compact in $L^1(\mu)$ for all $c > 0$. (See the following Remark 3.13). Take $l \in \mathbb{N}$ and $n \geq t_l$ arbitrary, then

$$V_{\mathbb{X}}\left(\hat{T}^n(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})\right) \overset{3.11}{\leq} C\left\|\frac{1}{\mu(A_l)}\mathbb{1}_{A_l}\right\|_{L^1(\mu)} + \rho^n V_{\mathbb{X}}\left(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l}\right) \leq C + \rho^n \frac{2}{\mu(A_l)}$$
$$\leq C + \frac{C\mu(A_l)}{2}\frac{2}{\mu(A_l)} = 2C.$$

We therefore get that $\hat{T}^{\tau_l}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l}) \in \mathcal{U}(2C)$ which is compact.

(i) There exists $l_0 \in \mathbb{N}$, such that $\tau_l = t_l$ for all $l \geq l_0$. We get $0 \leq \mu(A_l)\tau_l \leq \mu(A_l)\frac{-\log \mu(A_l)}{-\log \rho} + \mu(A_l)\frac{\log \frac{2}{C}}{-\log \rho} + \mu(A_l) \to 0$ as $l \to \infty$.

We are left to show condition (ii) from Theorem 1.25.

(ii) We argue as in Theorem 2.12. Since x^* is not periodic we know that for any $M \in \mathbb{N}$ fixed, each one of the points $Tx^*, ..., T^M x^*$ is different to x^*. The distance between these points is positive, i.e. $\min\limits_{0 \leq j < k \leq M} |T^j x^* - T^k x^*| \geq \varepsilon$ for some $\varepsilon > 0$. Since $\operatorname{dist}(A_l, x^*) \to 0$ and $\lambda(A_l) \to 0$ as $l \to \infty$ and since T, $T^2, ..., T^M$ are continuous on $\xi_M(x^*)$, we can choose l_0 large enough such that $A_l \cap (\bigcup_{j=1}^M T^j A_l) = \emptyset$ for all $l \geq l_0$.

$$\begin{aligned}
\mu_{A_l}(\varphi_{A_l} \leq \tau_l) &= \sum_{k=M+1}^{\tau_l} \mu_{A_l}(\underbrace{\varphi_{A_l} = k}_{\subseteq T^{-k}A_l}) \leq \sum_{k=M+1}^{\tau_l} \frac{1}{\mu(A_l)} \int\limits_{\mathbb{X}} \mathbb{1}_{A_l} \circ T^k \mathbb{1}_{A_l} \, d\mu \\
&= \sum_{k=M+1}^{\tau_l} \frac{1}{\mu(A_l)} \int\limits_{\mathbb{X}} \mathbb{1}_{A_l} \hat{T}^k \mathbb{1}_{A_l} \, d\mu \leq \sum_{k=M+1}^{\tau_l} \|\hat{T}^k \mathbb{1}_{A_l}\|_\infty \\
&\leq \sum_{k=M+1}^{\tau_l} \|\hat{T}^k \mathbb{1}_{A_l}\|_{BV(\mu)} = \sum_{k=M+1}^{\tau_l} \int\limits_{\mathbb{X}} \hat{T}^k \mathbb{1}_{A_l} \, d\mu + V_{\mathbb{X}}(\hat{T}^k \mathbb{1}_{A_l}) \\
&\overset{3.11}{\leq} \tau_l \mu(A_l) + \sum_{k=M+1}^{\tau_l} C \|\mathbb{1}_{A_l}\|_{L^1(\mu)} + \rho^k V_{\mathbb{X}}(\mathbb{1}_{A_l}) \\
&\leq (1+C)\tau_l \mu(A_l) + \frac{3\rho^{M+1}}{1-\rho}
\end{aligned}$$

Since we can choose M arbitrary we get that $\mu_{A_l}(\varphi_{A_l} \leq \tau_l) \to 0$ for $l \to \infty$.

Now the statement follows with Theorem 1.25.

□

Remark 3.13. The sets $\mathcal{U}(c) := \{u \in \mathcal{D}(\mu) \cap BV(\mu) : u \geq 0, V_{\mathbb{X}} \leq c\}$ are compact in $L^1(\mu)$.
Take a sequence $(u_n)_{n \geq 1} \subseteq \mathcal{U}(c)$, then each u_n has a version which is a sum of two monotone functions $u_n^1 + u_n^2$. By Helly's selection Theorem there exists a pointwise convergent subsequence $(u_{n_k})_{k \geq 1}$ such that $u_{n_k}(x) \to u(x)$ for some function u.
Now it follows via the Dominated Convergence Theorem, that $\|u\|_{L^1(\mu)} = 1$ and therefore $u \in L^1(\mu)$.
With the Lemma of Fatou we infer

$$\int\limits_{\mathbb{X}} \liminf_{k \to \infty} (|u_{n_k}| + |u| - |u_{n_k} - u|) d\mu \leq \liminf_{k \to \infty} \int\limits_{\mathbb{X}} (|u_{n_k}| + |u| - |u_{n_k} - u|) d\mu.$$

From this it follows that

$$\int\limits_{\mathbb{X}} 2|u| d\mu \leq 2 \int\limits_{\mathbb{X}} |u| d\mu - \limsup_{k \to \infty} \int\limits_{\mathbb{X}} |u_{n_k} - u| d\mu$$

from which we conclude L^1 convergence.

3.4 Example

If one knows that a system $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \nu, T, \xi)$ is Rychlik, one does not know that the system $(\mathbb{X}, \mathcal{A}, \mu, T, \xi)$, where $\mu << \nu$ is the invariant probability measure, is Rychlik. (See the example given below.)

Example. Let $\mathbb{X} = [0, 1]$ and λ be the Lebesgue measure. If we look at the map $T : [0, 1] \to [0, 1]$ defined as

$$Tx := \begin{cases} n^2(x - \frac{1}{n+1}), & x \in [\frac{1}{n+1}, \frac{1}{n}), \quad \text{for } n \geq 2 \\ \frac{n(n+1)}{2}(x - \frac{n-1}{n}), & x \in [\frac{n-1}{n}, \frac{n}{n+1}), \quad \text{for } n \geq 2. \end{cases}$$

This is a Rychlik map with the partition $\xi = \bigcup\limits_{n \geq 2} \{(\frac{1}{n+1}, \frac{1}{n}), (\frac{n-1}{n}, \frac{n}{n+1})\}$.

The weight function g is given by $\frac{1}{|T'|}$ on $Z \in \xi$ and by 0 else. One easily infers that g has bounded variation:

$$V_{\mathbb{X}} g = \sum_{n \geq 2} \frac{2}{n(n+1)} + \frac{1}{n^2} < \infty.$$

Moreover $\sup\limits_{x \in \mathbb{X}} g(x) < 1$ and the system $\mathcal{S} = ([0, 1], \mathcal{B}_{[0,1]}, \lambda, T, \xi)$ is ergodic (see Remark 3.15).

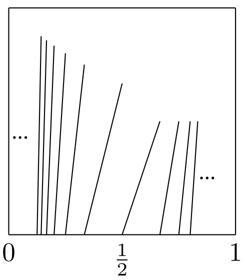

Figure 1: The interval map T

Next we calculate the invariant density h for this system. First we assume that an invariant density of the form $h = \sum\limits_{n=1}^{\infty} c_n \mathbb{1}_{[\frac{n-1}{n}, \frac{n}{n+1})}$ exists and infer what the weights c_n would have to be.
Let $\tilde{h} := \frac{h}{c_1}$ and $\tilde{c}_n := \frac{c_n}{c_1}$, then $\hat{T}\tilde{h} = \tilde{h}$ still holds and we get the following equalities:

$$1 = \tilde{h}(x) = \sum_{y \in T^{-1}x} g(y)\tilde{h}(y) = \sum_{n \geq 2} \left(\frac{1}{n^2} + \frac{2}{n(n+1)} \tilde{c}_n \right), \text{ for } x \in \left[0, \frac{1}{2}\right) \tag{3.4}$$

$$\tilde{c}_k = \tilde{h}(x) = \sum_{y \in T^{-1}x} g(y)\tilde{h}(y) = \sum_{n=k}^{\infty} \frac{1}{n^2}, \text{ for } x \in \left[\frac{k-1}{k}, \frac{k}{k+1}\right) \tag{3.5}$$

for $k \geq 2$.

From the relations above we know what the weights c_n must look like if there exists an invariant density which is piecewise constant on the sets $\left[\frac{n-1}{n}, \frac{n}{n+1}\right)$. However we do not yet know that such a density exists. There are two ways to show the existence. Either (i) we prove that equations (3.4) and (3.5) hold (then we know that h is the unique invariant probability density) or (ii) we show that the invariant probability density has to be piecewise continuous on the sets $\left[\frac{n-1}{n}, \frac{n}{n+1}\right)$ (and conclude that the weights have to be as calculated in (3.4) and (3.5)). We are going to discuss both methods in detail and start with the first one.

(i) The only equality we need to show is (3.4). After inserting the values of $\tilde{c}_n$ into (3.4) we get

$$\begin{aligned}\sum_{n=2}^{\infty}\left(\frac{1}{n^2}+\frac{2}{n(n+1)}\left(\sum_{k=n}^{\infty}\frac{1}{k^2}\right)\right) &= \frac{\pi^2}{6}-1+2\sum_{n=2}^{\infty}\left(\frac{1}{n}-\frac{1}{n+1}\right)\left(\sum_{k=n}^{\infty}\frac{1}{k^2}\right)\\ &= \frac{\pi^2}{6}-1+2\sum_{n=2}^{\infty}\left(\frac{1}{n}\left(\sum_{k=n}^{\infty}\frac{1}{k^2}\right)-\frac{1}{n+1}\left(\sum_{k=n+1}^{\infty}\frac{1}{k^2}\right)-\frac{1}{(n+1)n^2}\right)\\ &= \frac{\pi^2}{6}-1+\frac{\pi^2}{6}-1-2\sum_{n=2}^{\infty}\frac{1}{n^2(n+1)}\\ &= \frac{2\pi^2}{6}-1-2\sum_{n=1}^{\infty}\left(\frac{1}{n^2}-\frac{1}{n}+\frac{1}{n+1}\right)=2-1=1.\end{aligned}$$

In the first and the last step we used the partial fraction decomposition of $\frac{1}{n(n+1)}$ and $\frac{1}{n^2(n+1)}$ respectively. From the second to the third line and in the last step we used properties of a telescopic sum and that $\sum\limits_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$.
With this calculation we conclude that h is indeed the invariant probability density.

(ii) With the following method we want to skip the calculations in (i) by arguing why there has to exist an invariant density h which is piecewise continuous.
As noted before, the system $\mathcal{S}$ is Rychlik and ergodic. It therefore has to admit an invariant probability density $h \in BV(\mathbb{X})$. From the ergodicity we infer that

$$\left\|\frac{1}{j}\sum_{i=0}^{j-1}\hat{T}^i u-h\right\|_{L^1(\lambda)} \to 0 \tag{3.6}$$

as $j\to\infty$ for every $u\in L^1(\lambda)$. This especially holds for $u=\mathbb{1}_{[0,1]}$.
Let $f=\sum\limits_{n=1}^{\infty}k_n\mathbb{1}_{B_n}$ where $B_n := [\frac{n-1}{n},\frac{n}{n+1})$. Then $\hat{T}f=\sum\limits_{n=1}^{\infty}\tilde{k}_n\mathbb{1}_{B_n}$ for some $\tilde{k}_n$ and we conclude that $\frac{1}{j}\sum\limits_{i=0}^{j-1}\hat{T}^i\mathbb{1}_{[0,1]}$ is of the form $\sum\limits_{n=1}^{\infty}k_n^j\mathbb{1}_{B_n}$ for every $j\geq 0$ and for some $k_n^j\geq 0$.
For each $n\geq 1$ we know from (3.6) that

$$\|k_n^j\mathbb{1}_{B_n}-h\mathbb{1}_{B_n}\|_{L^1(\lambda)}\to 0$$

as $j \to \infty$. It follows that $(k_n^j \mathbb{1}_{B_n})_{j\geq 1} \subseteq L^1(\lambda)$ is Cauchy and since $\lambda(B_n) > 0$ we conclude that $(k_n^j)_{j\geq 1} \subseteq \mathbb{R}$ is Cauchy.
We infer that there exists $k_n \in \mathbb{R}$ such that $k_n^j \to k_n$ as $j \to \infty$ and get

$$\|k_n^j \mathbb{1}_{B_n} - k_n \mathbb{1}_{B_n}\|_{L^1(\lambda)} \to 0 \text{ as } j \to \infty.$$

We conclude that $h\mathbb{1}_{B_n} = k_n \mathbb{1}_{B_n}$ for every $n \geq 1$. Therefore h is piecewise constant and the weights have to be of the form given by (3.4) and (3.5).

With the above we now know that the invariant probability density of this system is given by

$$h = c\left(\mathbb{1}_{[0,\frac{1}{2})} + \sum_{n\geq 2}(\sum_{k\geq n}\frac{1}{k^2})\mathbb{1}_{[\frac{n-1}{n},\frac{n}{n+1})}\right),$$

where $c > 0$ is the normalizing constant.
The transfer operator of the new system with respect to the invariant measure μ, where $\frac{d\mu}{d\lambda} = h$, is given by $\hat{T}u(x) = \sum\limits_{y\in T^{-1}x} \frac{g(y)h(y)u(y)}{h(x)}$ and the new weight function by $\tilde{g}(x) = \frac{g(x)h(x)}{(h\circ T)(x)}$ (see Remark 1.9).
$\tilde{g}$ has unbounded variation:

$$V_{\mathbb{X}}\tilde{g} \geq \sum_{n\geq 2} \operatorname*{ess\,sup}_{x\in[\frac{1}{n+1},\frac{1}{n})} \frac{g(x)h(x)}{(h\circ T)(x)} \geq \sum_{n\geq 2}\frac{1}{\sum\limits_{k\geq n}\frac{n^2}{k^2}} = \infty \tag{3.7}$$

We conclude that the system with respect to μ is not Rychlik.

In the next paragraphs we will introduce the notion of inducing and use that some properties of the induced system $(Y, \mathcal{A}\cap Y, \nu|_{Y\cap\mathcal{A}}, T_Y)$, for some $Y \in \mathcal{A}$ transfer to the system $(\mathbb{X}, \mathcal{A}, \nu, T)$.
From this theory we infer that the system $\mathcal{S}$ in the example above is ergodic and that even though the invariant density $h := \frac{d\mu}{d\lambda}$ is not bounded away from 0, the Hitting and Return time statistics still admit an exponential law for a sequence of rare events $(A_l)_{l\geq 1}$ chosen as in Theorem 3.12.

Definition 3.14. Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \nu, T)$ be a null-preserving system, ν a σ-finite measure and $Y \in \mathcal{A}$ a recurrent set such that $\nu(Y) > 0$. We denote by $\mathcal{S}_Y = (Y, \mathcal{A}\cap Y, \nu|_{Y\cap\mathcal{A}}, T_Y)$ the induced system on Y. We call $T_Y : Y \to Y$ the induced map on Y or the first return map of Y, which is defined by

$$T_Y x := T^{\varphi_Y(x)}x.$$

Remark 3.15. Let $\mathcal{S} = ([0,1], \mathcal{B}_{[0,1]}, \lambda, T, \xi)$ be the system from the example above. As noted before we prove that this system is ergodic.
Look at the induced system on $Y = [0, \frac{1}{2}]$. The map T_Y is given by

$$T_Y x := \begin{cases} n^2(x - \frac{1}{n+1}), & x \in I_0^n, \quad \text{for } n \geq 2 \\ \frac{(k+1)(k+2)n^2}{2}(x - \frac{1}{n+1} - \frac{k}{(k+1)n^2}), & x \in I_k^n, \quad \text{for } 1 \leq k < n, \end{cases}$$

where $I_k^n := [\frac{1}{n+1} + \frac{k}{(k+1)n^2}, \frac{1}{n+1} + \frac{k+1}{(k+2)n^2})$ for $n \in \mathbb{N}$ and $0 \le k < n$.

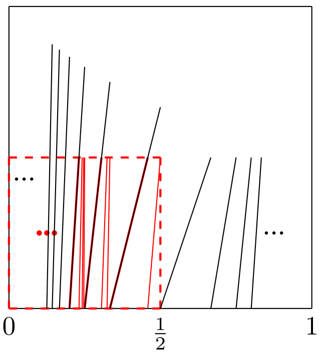

Figure 2: The induced map $T_Y : [0, \frac{1}{2}] \to [0, \frac{1}{2}]$

One sees easily that the system $\mathcal{S}_Y$ is Folklore and therefore ergodic. We chose the set Y such that $\mathbb{X} \subseteq \bigcup_{n\ge 1} T^{-n}Y \mod \lambda$ and conclude that $\mathcal{S}$ is ergodic.
Assume otherwise, then there exists $A \in \mathcal{B}_{[0,1]}$ such that $\lambda(A) \in (0,1)$, $T^{-1}A = A$ and $T^{-1}A^c = A^c$. It follows that $Y \cap A$ and $Y \cap A^c$ are invariant with respect to T_Y. We will show that $0 \notin \{\lambda(Y \cap A^c), \lambda(Y \cap A)\}$ which is a contradiction to the ergodicity of $\mathcal{S}_Y$.
Assume w.l.o.g that $\lambda(Y \cap A) = 0$. We infer $Y \subseteq A^c \mod \lambda$ and conclude our argument by noticing that

$$\mathbb{X} \subseteq \bigcup_{n\ge 1} T^{-n}Y \subseteq \bigcup_{n\ge 1} T^{-n}A^c = A^c.$$

In Theorem 3.12 we assumed that the invariant density of the system is bounded away from 0 and inferred that the hitting- and return-time statistics admit exponential law. In the example above the invariant probability density $h = \frac{d\mu}{d\lambda}$ gets arbitrarily close to 0 near 1 and we cannot use Theorem 3.12. Below we show that the limit distribution of the hitting- and return-time distributions is the exponential distribution. We will look at the case $A_l := (1-\varepsilon_l, 1)$ separately, since this is where the invariant density becomes arbitrarily small. For sets A_l with $\operatorname{dist}(A_l, 1) > \delta$ for all $l \ge 1$ we will use that the hitting- and return-time statistics for the induced system on some Y chosen appropriately admit exponential law and infer that the hitting- and return-time statistics for the original system are exponentially distributed.

Remark 3.16. We look again at the system $\mathcal{S} = ([0,1], \mathcal{B}_{[0,1]}, \lambda, T, \xi)$ of the example above and μ is the corresponding invariant probability measure. Let $A_l := (1-\varepsilon_l, 1)$ for some $0 < \varepsilon_l < \frac{1}{2}$ such that $\varepsilon_l \searrow 0$ as $l \to \infty$ and let $(A_l)_{l\ge 1} \subseteq \mathcal{B}_{[0,1]}$ be the sequence of rare events we are interested in.
By using Theorem 1.25 we will show that for $R_l := \mu(A_l)\varphi_{A_l}$

(i) $R_l \overset{\mu}{\Longrightarrow} \mathcal{E}$

(ii) $R_l \overset{\mu_{A_l}}{\Longrightarrow} \mathcal{E}$

holds.
We already showed that $\tilde{\mathcal{S}} = ([0,1], \mathcal{B}_{[0,1]}, \mu, T, \xi)$ is an ergodic probability-preserving system and we take $\tau_l = 1$. Assertions (i) and (ii) follow immediately since $(A_l)_{l\geq 1}$ is a sequence of rare events and since $T^{-1}A_l \subseteq [0, \frac{1}{2}]$.
For (iii) we calculate

$$\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})(x) = \sum_{y\in T^{-1}x} \frac{g(y)h(y)}{h(x)} \frac{1}{\mu(A_l)}\mathbb{1}_{A_l}(x)$$
$$= \begin{cases} \sum\limits_{k\geq n_l+1} \frac{2}{k(k+1)} \left(\sum\limits_{j\geq k} \frac{1}{j^2}\right) \frac{c}{\mu(A_l)}, & x \in [0, \frac{n_l(n_l+1)}{2}\delta_l) \\ \sum\limits_{k\geq n_l} \frac{2}{k(k+1)} \left(\sum\limits_{j\geq k} \frac{1}{j^2}\right) \frac{c}{\mu(A_l)}, & x \in [\frac{n_l(n_l+1)}{2}\delta_l, \frac{1}{2}) \\ 0, & x \in [\frac{1}{2}, 1], \end{cases} \tag{3.8}$$

where $n_l \geq 2$ is sucht that $\frac{n_l-1}{n_l} \leq \varepsilon_l < \frac{n_l}{n_l+1}$ and $\delta_l := \varepsilon_l - \frac{n_l-1}{n_l}$. With (3.8) we can show that $\|\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})\|_{BV(\mu)} \leq K$ for some $K > 0$ and all $l \geq 1$. It suffices to show that $V_{\mathbb{X}}(\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})) \leq K-1$ since $\|\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})\|_{L^1(\mu)} = 1$. Since $\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})$ is constant on a partition of $[0,1]$ into three intervals it suffices to show $\|\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l}\|_\infty < \infty$.
The transfer operator $\hat{T}$ preserves the integral and $\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})(x) \leq \hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})(y)$ for a.e. $x \in [0, \frac{n_l(n_l+1)}{2}\delta_l)$ and a.e. $y \in [\frac{n_l(n_l+1)}{2}\delta_l, \frac{1}{2})$ and we infer that

$$\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})(x) \leq 2.$$

The factor $\mu(A_l)$ is given by

$$\sum_{k\geq n_l+1} \left(\sum_{j\geq k} \frac{1}{j^2}\right) \frac{1}{k(k+1)} + \left(\sum_{j\geq n_l} \frac{1}{j^2}\right) \left(\frac{1}{n_l(n_l+1)} - \delta_l\right).$$

For a.e. $x \in [\frac{n_l(n_l+1)}{2}\delta_l, \frac{1}{2})$ we therefore get

$$\begin{aligned} &\frac{1}{c}\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})(x) \\ &= \sum_{k\geq n_l} \frac{2}{k(k+1)} \left(\sum_{j\geq k} \frac{1}{j^2}\right) \left(\sum_{k\geq n_l+1} \left(\sum_{j\geq k} \frac{1}{j^2}\right) \frac{1}{k(k+1)} + \left(\sum_{j\geq n_l} \frac{1}{j^2}\right) \left(\frac{1}{n_l(n_l+1)} - \delta_l\right)\right)^{-1} \\ &= 2 + 2\delta_l \left(\sum_{j\geq n_l} \frac{1}{j^2}\right) \left(\sum_{k\geq n_l+1} \left(\sum_{j\geq k} \frac{1}{j^2}\right) \frac{1}{k(k+1)} + \left(\sum_{j\geq n_l} \frac{1}{j^2}\right) \left(\frac{1}{n_l(n_l+1)} - \delta_l\right)\right)^{-1} \\ &\leq 4. \end{aligned}$$

In the last inequality we used $\delta_l \leq \frac{1}{n_l(n_l+1)}$, $n_l \geq 2$ and that

$$
\begin{aligned}
&2\frac{1}{n_l(n_l+1)}\left(\sum_{j\geq n_l}\frac{1}{j^2}\right) \leq 2\sum_{k\geq n_l+1}\left(\sum_{j\geq k}\frac{1}{j^2}\right)\frac{1}{k(k+1)} \\
\iff &2\frac{1}{n_l(n_l+1)}\left(\sum_{j\geq n_l}\frac{1}{j^2}\right) \leq \frac{2}{n_l^3(n_l+1)} + \frac{1}{n_l+1}\sum_{j\geq n_l+1}\frac{1}{j^2} \leq 2\frac{1}{n_l+1}\sum_{j\geq n_l+1}\frac{1}{j^2} \\
\Longleftarrow &2\frac{1}{n_l^3} \leq \sum_{j\geq n_l+1}\frac{1}{j^2}
\end{aligned} \tag{3.9}
$$

holds. One verifies (3.9) easily for $n_l \geq 3$ since the estimates $\frac{1}{n_l^3} \leq \frac{1}{(n_l+1)^2}$ and $\frac{1}{n_l^3} \leq \frac{1}{(n_l+2)^2}$ hold. For $n_l = 2$ Equation (3.9) also holds true.
With the above we get that $V_{\mathbb{X}}(\hat{T}(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l}) \leq 6$ and Remark 3.13 concludes the argument.
Now let $(A_l)_{l\geq 1} \subseteq \mathcal{B}_{[0,1]}$ be a sequence of rare events as assumed in Theorem 3.12 with the additional property that

$$\inf_{l\geq 1} \operatorname{dist}(A_l, 1) > \delta \text{ for some } \delta > 0. \tag{3.10}$$

We use that the induced system on a suitably chosen Y has exponential hitting- and return-time statistics to infer the same for $\mathcal{S}$. Assumption (3.10) ensures that there exists $N \in \mathbb{N}$ such that $A_l \subseteq [0, \frac{N}{N+1}] =: Y$ for all $l \geq 1$.

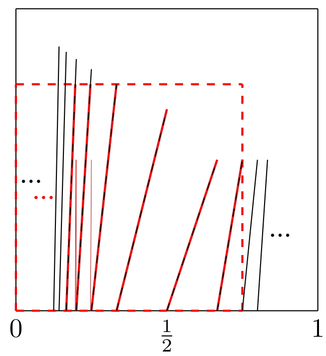

Figure 3: The induced map T_Y for $N = 3$

$$
T_Y x = \begin{cases} \frac{n(n+1)}{2}(x - \frac{n-1}{n}), & x \in [\frac{n-1}{n}, \frac{n}{n+1}], \quad \text{for } 2 \leq n \leq N \\ n^2(x - \frac{1}{n+1}), & x \in I^n, \quad \text{for } n \geq 2 \\ \frac{(k+1)(k+2)n^2}{2}(x - \frac{1}{n+1} - \frac{k}{(k+1)n^2}), & x \in I_k^n, \quad \text{for } N \leq k < n, \end{cases}
$$

where $I_n := [\frac{1}{n+1}, \frac{1}{n}]$ for $n \leq N$, $I_n := [\frac{1}{n+1}, \frac{1}{n+1} + \frac{N}{(N+1)n^2}]$ for $n > N$ and

$I_k^n := [\frac{1}{n+1} + \frac{k}{(k+1)n^2}, \frac{1}{n+1} + \frac{k+1}{(k+2)n^2})$ for $n \in \mathbb{N}$ and $N \leq k < n$. Since

$$\sum_{n\geq 2} \frac{1}{n^2}\left(1 + \sum_{k=N}^{n-1} \frac{2}{(k+1)(k+2)}\right) + \sum_{n=2}^{N} \frac{2}{n(n+1)} < \infty$$

the system $\mathcal{S}_Y$ is Rychlik.

We infer from the theory about induced systems, see [1, 41 ff.] and the appendix, that the invariant measure of $\mathcal{S}_Y$ is given by $\mu_Y := \frac{1}{\mu(Y)}\mu|_{Y\cap\mathcal{A}}$ where μ is the invariant measure of the system $\mathcal{S}$ and thus bounded and bounded away from 0. We conclude the argument by using Theorem 3.12 and that if the limit hitting- and return-time distributions of the induced system are exponentially distributed, then so are the one of the original system (See Theorem A.3).

Even though we know that the system $\tilde{\mathcal{S}} = (\mathbb{X}, \mathcal{A}, \mu, T, \xi)$ from the example above is not Rychlik, we do not know if it still possesses similar properties, for example if the transfer operator $\hat{T}$ maps $BV(\mu)$ onto itself or if it still fulfills some kind of Lasota-Yorke inequality. These questions are subject of the next remark.

Remark 3.17. We will show that the systems $(\mathbb{X}, \mathcal{A}, \mu, T^2, \xi_2)$ and $(\mathbb{X}, \mathcal{A}, \mu, T^3, \xi_3)$ are Rychlik and infer that

$$V_{\mathbb{X}}(\hat{T}u) \leq C\|u\|_{L^1(\mu)} + \chi^n V_{\mathbb{X}}(u)$$

holds for all $u \in BV(\mu)$ and $n \geq n^*$ for some $n^* \in \mathbb{N}$.

The first observation we make is that the n-th iteration of the transfer operator is given by

$$(\hat{T}^n u)(x) = \sum_{y\in T^{-n}x} \left(\frac{g(g\circ T)\cdots(g\circ T^{n-1})h}{h\circ T}u\right)(y). \tag{3.11}$$

It suffices to show $\sup\limits_{x\in\mathbb{X}} \frac{g(g\circ T)h}{h\circ T}(x) < 1$ (respectively $\sup\limits_{x\in\mathbb{X}} \frac{g(g\circ T)(g\circ T^2)h}{h\circ T}(x) < 1$) and $V_{\mathbb{X}}\left(\frac{g(g\circ T)h}{h\circ T}\right) < \infty$ (respectively $V_{\mathbb{X}}\left(\frac{g(g\circ T)(g\circ T^2)h}{h\circ T}\right) < \infty$) to conclude that $(\mathbb{X}, \mathcal{A}, \mu, T^2, \xi_2)$ (respectively $(\mathbb{X}, \mathcal{A}, \mu, T^3, \xi_3)$) is Rychlik.

$$\sup_{x\in\mathbb{X}} \frac{g(g\circ T)h}{h\circ T}(x) \leq \sup_{n\geq 2}\left\{\sup_{2\leq j\leq n} \frac{\frac{2}{n^2 j(j+1)}}{\sum\limits_{k\geq j}\frac{1}{k^2}}, \sup_{j\geq 2} \frac{2}{n(n+1)j^2}\sum_{k\geq j}\frac{1}{k^2}\right\} < 1.$$

The above combined with the fact that $\sup\limits_{x\in\mathbb{X}} g(x) < 1$ implies that $\sup\limits_{x\in\mathbb{X}} \frac{g(g\circ T)(g\circ T^2)h}{h\circ T}(x) < 1$. Next we show that the variation is bounded.

$$V_{\mathbb{X}}\left(\frac{g(g\circ T)h}{h\circ T}\right)=\sum_{n\geq 2}\overbrace{\frac{1}{n^2}}^{gh}\left(\sum_{k\geq 2}\overbrace{\frac{1}{k^2}}^{\frac{g\circ T}{h\circ T}}+\sum_{k=2}^{n}\overbrace{\frac{\frac{2}{k(k+1)}}{\sum\limits_{i\geq k}\frac{1}{i^2}}}^{\frac{g\circ T}{h\circ T}}\right)+\sum_{n\geq 2}\overbrace{\frac{2}{n(n+1)}}^{g}\left(\overbrace{\sum_{i\geq n}\frac{1}{n^2}}^{h}\right)\left(\sum_{k\geq 2}\overbrace{\frac{1}{k^2}}^{\frac{g\circ T}{h\circ T}}\right)$$

$$\leq\left(\frac{\pi^2}{6}\right)^2+\sum_{n\geq 2}\frac{1}{n^{\frac{3}{2}}}\sum_{k=2}^{n}\frac{2}{k^{\frac{3}{2}}(k+1)}\left(\sum_{i\geq k}\frac{1}{i^2}\right)^{-1}+\left(\frac{\pi^2}{6}\right)^2$$

With the above estimate it suffices to show that

$$\sum_{k\geq 2}\underbrace{\frac{1}{k^{\frac{3}{2}}(k+1)}\left(\sum_{i\geq k}\frac{1}{i^2}\right)^{-1}}_{=:a_k}<\infty. \tag{3.12}$$

Since we are interested in the finiteness of the above series we only need to find a non-negative sequence b_k such that $\sum\limits_{k\geq 2} b_k<\infty$ and $0<\lim\limits_{k\to\infty}\frac{b_k}{a_k}=c<\infty$. Let us choose $b_k:=\frac{1}{k^{\frac{3}{2}}}$ then

$$\lim_{k\to\infty}\frac{b_k}{a_k}=(k+1)\left(\sum_{i\geq k}\frac{1}{i^2}\right)\geq\lim_{k\to\infty}\sum_{i\geq k}\frac{k}{i(i+1)}=1.$$

$\lim\limits_{k\to\infty}\frac{b_k}{a_k}\leq 1$ can be obtained in a similar way. We infer finiteness of (3.12) from the finiteness of $\sum\limits_{k\geq 2}\frac{1}{k^{\frac{3}{2}}}$. The variation of $\frac{g(g\circ T)(g\circ T^2)h}{h\circ T}$ is given by

$$\begin{aligned}&V_{\mathbb{X}}\left(\frac{g(g\circ T)(g\circ T^2)h}{h\circ T}\right)=\\&\sum_{n\geq 2}\frac{1}{n^2}\left(\sum_{k\geq 2}\frac{1}{k^2}\left(\sum_{j\geq 2}\frac{1}{j^2}+\sum_{j=2}^{k}\frac{2}{j(j+1)}\right)+\sum_{k=2}^{n}\frac{2}{k(k+1)}\left(\sum_{i\geq k}\frac{1}{i^2}\right)^{-1}\left(\sum_{j\geq 2}\frac{1}{j^2}\right)\right)\\&+\sum_{n\geq 2}\frac{2}{n(n+1)}\left(\sum_{i\geq n}\frac{1}{i^2}\right)\left(\sum_{k\geq 2}\frac{1}{k^2}\left(\sum_{j\geq 2}\frac{1}{j^2}+\sum_{j=2}^{k}\frac{2}{j(j+1)}\right)\right).\end{aligned}$$

Using the observations from above this is again finite and we conclude that both systems are Rychlik and inequality (3.2) holds for $P\in\{\hat{T}^2,\hat{T}^3\}$. One can easily derive

$$V_{\mathbb{X}}\left(\hat{T}^n u\right)\leq\rho^n V_{\mathbb{X}}(u)+c\|u\|_{L^1(\mu)},$$

for some $c>0$, $\rho\in(0,1)$ and $n\geq n^*$, for some $n^*\in\mathbb{N}$ and all $u\in BV(\mu)$.

4 Hitting- and Return- time Statistics for Piecewise Expanding Systems in higher Dimensions

In this section we are going to introduce another family of dynamical systems and again look at the hitting- and return-time statistics and which assumptions are sufficient such that they admit exponential limit distributions. In the preceeding chapters we only looked at systems defined on $\mathbb{X} \subseteq \mathbb{R}$, whereas in this chapter we try to find similar results for dynamical systems where $\mathbb{X} \subseteq \mathbb{R}^n$ is compact. We will look at a family of systems introduced by B. Saussol in [14, pp. 226–227]. The approach in this paper is similar to the one from F. Hofbauer and G. Keller in [10], in the sense that one also uses the theorem of Ionescu-Tulcea and Marinescu to get a spectral decomposition of the transfer operator on some function space $\tilde{L} \subseteq L^1(\nu)$. Whereas in the one dimensional case, one could restrict the transfer operator to the well known $BV(\mathbb{X})$ function space to get the needed properties for the transfer operator, this is not as easy in higher dimensions.

4.1 Oscillation of a function

At the beginning of this section, we construct a function space by looking at the oscillation of a function $f : R^n \to \mathbb{R}$ and whether it is finite. Again we denote by $L^1(\lambda^n)$ the space of L^1-integrable functions on $\mathbb{X} \subseteq \mathbb{R}^n$ with respect to the n-dimensional Lebesgue measure λ^n and by $L^1(\mathbb{R}^n)$ we denote the L^1-integrable functions $f : \mathbb{R}^n \to \mathbb{R}$ with respect to λ^n.

In the following definition we define the oscillation of a Lebesgue-integrable function $f \in L^1(\mathbb{R}^n)$. If we then take some $f \in L^1(\lambda^n)$ defined on $\mathbb{X} \subseteq \mathbb{R}^n$ and talk about the oscillation of f, we mean the oscillation of $\tilde{f} \in L^1(\mathbb{R}^n)$ where

$$\tilde{f}(x) := \begin{cases} f(x), & x \in \mathbb{X} \\ 0, & x \in \mathbb{R}^n \setminus \mathbb{X}. \end{cases}$$

Definition 4.1. For $0 < \alpha \leq 1$ and $f \in L^1(\mathbb{R}^n)$ we define for a Borel-set $A \subseteq \mathbb{R}^n$, the oscillation of f on A by $\operatorname{osc}(f, A) := \operatorname*{ess\,sup}_{x \in A} f(x) - \operatorname*{ess\,inf}_{x \in A} f(x)$, where ess inf and ess sup are the essential infimum and the essential supremum respectively.

The oscillation function fulfills some inequalities which can be found in [14, pp. 232–233]. We are going to state them without proof.

Proposition 4.2. *Let u, u_j, $v \in L^\infty(\mathbb{R}^n)$ for $j \in \mathbb{N}$, $v > 0$, $\varepsilon > 0$, $x \in \mathbb{R}^n$ and $A \subseteq \mathbb{R}^n$ be a Borel-set. Then the following properties hold:*

(i) $\operatorname{osc}(\sum_{j=1}^{\infty} u_j, B_\varepsilon(x)) \leq \sum_{j=1}^{\infty} \operatorname{osc}(u_j, B_\varepsilon(x))$

(ii) $\operatorname{osc}(u\mathbb{1}_A, B_\varepsilon(x)) \leq \operatorname{osc}(u, A \cap B_\varepsilon(x))\mathbb{1}_A(x) + 2 \operatorname*{ess\,sup}_{y \in A \cap B_\varepsilon(x)} |u(y)|\mathbb{1}_{B_\varepsilon(A) \cap B_\varepsilon(A^c)}(x)$

(iii) $\operatorname{osc}(uv, A) \leq \operatorname{osc}(u, A) \operatorname*{ess\,sup}_{y \in A} v(y) + \operatorname{osc}(v, A) \operatorname*{ess\,inf}_{y \in A} |u(y)|$

(iv) For $a, b, c > 0$ *such that* $a + b \leq c$, *it holds that*
$\operatorname*{ess\,sup}_{y \in B_a(x)} u(y) \leq \frac{1}{\lambda^n(B_b(x))} \int\limits_{B_b(x)} (f + \operatorname{osc}(f, B_c(\cdot)))d\lambda^n.$

One can show that the function $x \mapsto \operatorname{osc}(f, B_\varepsilon(x))$ is measurable for $\varepsilon > 0$, since it is lower semi continuous, which makes it possible to define

$$|f|_\alpha := \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int\limits_{\mathbb{R}^n} \operatorname{osc}(f, B_\varepsilon(x))d\lambda^n(x)$$

for $\varepsilon_0 > 0$.
One should remark that the value of $|\cdot|_\alpha$ depends on the choice of ε_0, but whether it is finite or infinite is independent of the choice of ε_0, which is why we neglect ε_0 in the notation.
With the definitions above we are now able to define the space of functions with finite oscillation, which will replace the space of functions with bounded variation, as follows:

$$\operatorname{Osc}_\alpha(\mathbb{X}) := \{f \in L^1(\lambda^n) : |f|_\alpha < \infty\}.$$

The space $\operatorname{Osc}_\alpha(\mathbb{X})$ with the norm $\|\cdot\|_\alpha := \|\cdot\|_{L^1(\lambda^n)} + |\cdot|_\alpha$ is a Banach space. For this space the sets $\operatorname{Osc}_{\alpha,c} := \{f \in L^1(\lambda^n) : \|f\|_\alpha \leq c\}$ are compact in $L^1(\lambda^n)$ for $c > 0$.

Remark 4.3. In this remark we are going to show that the sets $\operatorname{Osc}_{\alpha,c}(\mathbb{X})$ are relatively compact in $L^1(\lambda^n)$. As noted before these sets are even compact, but for our purpose it suffices to know that they are relatively compact. In [11, pp. 466–467] one can take a look at Lemma 1.12 which concludes the proof of compactness of $\operatorname{Osc}_{\alpha,c}(\mathbb{X})$.
Since we assume throughout this chapter that $\mathbb{X} \subseteq \mathbb{R}^n$ is compact, we find finite partitions $\pi_\varepsilon = \{X_1^\varepsilon, ..., X_{N_\varepsilon}^\varepsilon\}$ of $\mathbb{X}$, for $0 < \varepsilon \leq \varepsilon_0$, with the property that for each $X \in \pi_\varepsilon$ there exists $x \in \mathbb{X}$ such that $X \subseteq B_{\frac{\varepsilon}{2}}(x)$. If we look at the operator $U_{\pi_\varepsilon} : \operatorname{Osc}_\alpha(\mathbb{X}) \to L^1(\lambda^n)$ defined as

$$U_{\pi_\varepsilon} u := \sum_{i=1}^{N_\varepsilon} \frac{1}{\lambda^n(X_i^\varepsilon)} \left(\int\limits_{X_i^\varepsilon} u \, d\lambda^n \right) \mathbb{1}_{X_i^\varepsilon},$$

then $\|U_{\pi_\varepsilon} u\|_{L^1(\lambda^n)} \leq \sum_{i=1}^{N_\varepsilon} \frac{1}{\lambda^n(X_i^\varepsilon)} \int\limits_{\mathbb{X}} \left(\int\limits_{X_i^\varepsilon} |u| \, d\lambda^n \right) \mathbb{1}_{X_i^\varepsilon} \, d\lambda^n = \|u\|_{L^1(\lambda^n)}$ and the operator is bounded.

We estimate

$$\begin{aligned}
\|u - U_{\pi_\varepsilon} u\|_{L^1(\lambda^n)} &= \int\limits_{\mathbb{X}} |u - \sum_{i=1}^{N_\varepsilon} \frac{1}{\lambda^n(X_i^\varepsilon)} (\int\limits_{X_i^\varepsilon} u \, d\lambda^n) \mathbb{1}_{X_i^\varepsilon}| \, d\lambda^n \\
&\leq \int\limits_{\mathbb{X}} |\sum_{i=1}^{N_\varepsilon} (\operatorname*{ess\,sup}_{x \in X_i^\varepsilon} u(x) - \operatorname*{ess\,inf}_{x \in X_i^\varepsilon} u(x)) \mathbb{1}_{X_i^\varepsilon}| \, d\lambda^n \\
&\leq \int\limits_{\mathbb{X}} \operatorname{osc}(u, B_\varepsilon(\cdot)) \, d\lambda^n \leq \varepsilon^\alpha |u|_\alpha,
\end{aligned}$$

where the second to last inequality holds since all $X \in \pi_\varepsilon$ are enclosed by $\frac{\varepsilon}{2}$-balls.
If we fix $\delta > 0$ we can choose $\varepsilon > 0$ small such that $\|u - U_{\pi_\varepsilon} u\|_{L^1(\lambda^n)} < \frac{\delta}{3}$ for all $u \in \operatorname{Osc}_{\alpha,c}(\mathbb{X})$.
Since the operator U_{π_ε} is bounded and the set $\operatorname{Osc}_{\alpha,c}(\mathbb{X})$ is bounded and the image of $\operatorname{Osc}_\alpha(\mathbb{X})$ is finite dimensional, we know that $U_{\pi_\varepsilon}(\operatorname{Osc}_{\alpha,c}(\mathbb{X}))$ is relatively compact in $L^1(\lambda^n)$. This implies that there exist $\{u_1, ..., u_k\} \subseteq \operatorname{Osc}_{\alpha,c}(\mathbb{X})$ such that
$\inf\limits_{i \in \{1,..,k\}} \|U_{\pi_\varepsilon} u - U_{\pi_\varepsilon} u_i\|_{L^1(\lambda^n)} < \frac{\delta}{3}$.
Alltogether we now get that for every $\delta > 0$ there exist $0 < \varepsilon \leq \varepsilon_0$ and $\{u_1, ..., u_k\}$ chosen as above, such that $\forall u \in \operatorname{Osc}_{\alpha,c}(\mathbb{X})$

$$\begin{aligned}
&\min_{i \in \{1,...k\}} \|u - u_i\|_{L^1(\lambda^n)} \\
&\leq \min_{i \in \{1,...k\}} \|u - U_{\pi_\varepsilon} u\|_{L^1(\lambda^n)} + \|U_{\pi_\varepsilon} u - U_{\pi_\varepsilon} u_i\|_{L^1(\lambda^n)} + \|U_{\pi_\varepsilon} u_i - u_i\|_{L^1(\lambda^n)} \\
&< \delta.
\end{aligned}$$

Since $L^1(\lambda^n)$ is a complete metric space, this gives us that $\operatorname{Osc}_{\alpha,c}(\mathbb{X})$ is relatively compact. For further information one can take a look at [6, 259 ff.]

The following proposition implies that functions in $\operatorname{Osc}_\alpha(\mathbb{X})$ are in $L^\infty(\lambda^n)$. A proof can be found in [14, pp. 233–234].

Proposition 4.4. *Let u, $v \in \operatorname{Osc}_\alpha(\mathbb{X})$, then the following two estimates hold:*

(i) $\|u\|_\infty \leq \frac{\max\{1, \varepsilon_0^\alpha\}}{\lambda^n(B_{\varepsilon_0}(0))} \|u\|_\alpha$

(ii) $\|uv\|_\alpha \leq \frac{2 \max\{1, \varepsilon_0^\alpha\}}{\lambda^n(B_{\varepsilon_0}(0))} \|u\|_\alpha \|v\|_\alpha$.

4.2 Piecewise Expanding Systems

We now define the family of dynamical systems. By $\| \cdot \|$ we denote the euclidean norm on $\mathbb{R}^n$.

Definition 4.5. Let $\mathbb{X} \subseteq \mathbb{R}^n$ be a compact set such that the closure of the interior of $\mathbb{X}$ coincides with $\mathbb{X}$ and there exists a countable partition $\xi = \{Z_i : i \in \mathbb{N}\} \, (\operatorname{mod} \lambda^n)$ of $\mathbb{X}$, where each $Z_i \subseteq \mathbb{X}$ is open. Let $(\mathbb{X}, \mathcal{A}, \lambda^n, T)$ be a measurable dynamical system, $\mathcal{A}$ the induced Borel σ-algebra on $\mathbb{X}$ and λ^n the

Lebesgue measure on $\mathcal{A}$. We call a system $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda^n, T, \xi)$ piecewise expanding if the following holds for some $0 < \alpha \leq 1$ and for some $\varepsilon_0 > 0$ small enough:

(PE1) For every set $Z_i \in \xi$ there exists an open set $V_i \subseteq \mathbb{R}^n$ such that $\bar{Z}_i \subseteq V_i$ and such that $T|_{Z_i}$ admits an extension to a C^1-diffeomorphism $T_i : V_i \to T_i(V_i)$. Moreover $T_i(V_i) \supseteq B_{\varepsilon_0}(T(Z_i))$ holds.

(PE2) There exists $c > 0$ such that for all $i \in \mathbb{N}$, $0 < \varepsilon \leq \varepsilon_0$, $z \in T_i(V_i)$ and $x, y \in B_\varepsilon(z) \cap T_i(V_i)$ the following holds:

$$|\det D_x T_i^{-1} - \det D_y T_i^{-1}| \leq c\varepsilon^\alpha |\det D_z T_i^{-1}|. \tag{4.1}$$

(PE3) There exists $0 < s < 1$ such that for all $x, y \in T_i(V_i)$ such that $\|x - y\| \leq \varepsilon_0$ the following holds:

$$\|T_i^{-1}x - T_i^{-1}y\| \leq s\|x - y\|. \tag{4.2}$$

(PE4) Define $G(\varepsilon, \varepsilon_0) := \sup_{x \in \mathbb{X}} G(x, \varepsilon, \varepsilon_0)$ where

$$G(x, \varepsilon, \varepsilon_0) := \sum_{i \in \mathbb{N}} \frac{\lambda(T_i^{-1}(B_\varepsilon(\partial T(Z_i))) \cap B_{(1-s)\varepsilon_0}(x))}{\lambda(B_{(1-s)\varepsilon_0(x)}(x))}. \tag{4.3}$$

We then define $\eta(\delta) := s^\alpha + 2 \sup_{0<\varepsilon\leq\delta} \frac{G(\varepsilon,\varepsilon_0)}{\varepsilon^\alpha} \delta^\alpha$, and finally demand that $\eta := \sup_{0<\delta\leq\varepsilon_0} \eta(\delta) < 1$.

In [14] B. Saussol derives some properties for such system. We will include the most important ones here. For further details and proofs take a look at [14, pp. 233–238].

Theorem 4.6. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda^n, T, \xi)$ be a piecewise expanding system, then the following holds:*

(i) The transfer operator $\hat{T} : L^1(\lambda^n) \to L^1(\lambda^n)$ has only finitely many eigenvalues $\chi_1, ..., \chi_N$ on the unit circle.

(ii) $\hat{T}$ admits a representation as

$$\hat{T} = \sum_{j=1}^{N} \chi_j Q_j + R,$$

where Q_j is the projection onto the eigenspace corresponding to the eigenvalue χ_j and R is some linear operator such that $R(\mathrm{Osc}_\alpha(\mathbb{X})) \subseteq \mathrm{Osc}_\alpha(\mathbb{X})$. For these operators $RQ_j = Q_j R = 0$ and $Q_j Q_i = 0$ hold for any $i, j \in \{1, ..., N\}$.

(iii) $Q_j(L^1(\lambda^n)) \subseteq \mathrm{Osc}_\alpha(\mathbb{X})$

(iv) 1 is an eigenvalue and there exists an invariant measure $\mu << \lambda^n$ such that $\frac{d\mu}{d\lambda^n} \in \mathrm{Osc}_\alpha(\mathbb{X})$.

We now start with our derivation for a result about exponential hitting- and return-time statistics of such systems. In the proof we again want to use Theorem 1.25.

First we show a Lasota-Yorke-type inequality, which we will then use to infer (ii) from Theorem 1.25. An inequality similar to the one for Rychlik systems (Equation (3.2)) can be found in [14, pp. 235–237]. We will include it without proof, which consists of elementary facts but is non-trivial.

Lemma 4.7. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda^n, T, \xi)$ be a piecewise expanding system and P be the corresponding transfer operator. Then there exists $0 < \chi < 1$ and $F > 0$ such that for $f \in \mathrm{Osc}_\alpha(\mathbb{X})$ and $j \in \mathbb{N}$*

$$|P^j f|_\alpha \leq \chi^j |f|_\alpha + F\|f\|_{L^1(\lambda^n)}. \tag{4.4}$$

The above inequality can only be used for the transfer operator with respect to the Lebesgue measure. We want to use it for the transfer operator $\hat{T}$ with respect to the invariant measure $\mu << \lambda^n$. That this is possible is subject of the next theorem.

Theorem 4.8. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda^n, T, \xi)$ be an ergodic piecewise expanding system and let the invariant probability measure $\mu << \lambda^n$ with density $h := \frac{d\mu}{d\lambda^n}$ be such that there exists $k \in \mathbb{N}$ such that $\frac{1}{k} \leq h \leq k$. Then there exist constants $C > 0$, $\rho \in (0,1)$ such that for the transfer operator $\hat{T} : L^1(\mu) \to L^1(\mu)$ with respect to the measure μ we get*

$$|\hat{T}^j f|_\alpha \leq \rho^j |f|_\alpha + C\|f\|_{L^1(\mu)}, \tag{4.5}$$

for $f \in \mathrm{Osc}_\alpha(\mathbb{X})$ and $j \geq j_0$, for some $j_0 \in \mathbb{N}$ big enough.

Proof. Let P be the transfer operator of the system $\mathcal{S}$ corresponding to λ^n. Then $\hat{T}^j f = \frac{P^j(fh)}{h}$ and we get that there exists a constant c_1 chosen according to Proposition 4.4(i) and F and χ chosen as in Lemma 4.7.

$$\begin{aligned}
|\hat{T}^j f|_\alpha = \left|\frac{P^j(fh)}{h}\right|_\alpha &\overset{4.2(iii)}{\leq} \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha}\Bigg(\int_{\mathbb{R}^n} \mathrm{osc}(P^j(fh), B_\varepsilon(\cdot)) \operatorname*{ess\,sup}_{y\in B_\varepsilon(\cdot)} \frac{1}{h(y)} d\lambda^n \\
&+ \int_{\mathbb{R}^n} \mathrm{osc}\left(\frac{1}{h}, B_\varepsilon(\cdot)\right) \operatorname*{ess\,inf}_{y\in B_\varepsilon(\cdot)} |P^j(fh)(y)| d\lambda^n\Bigg) \\
&\overset{4.4(i)}{\leq} k|P^j(fh)|_\alpha + c_1\left|\frac{1}{h}\right|_\alpha (|P^j(fh)|_\alpha + \|P^j(fh)\|_{L^1(\lambda^n)}) \\
&\overset{4.7}{\leq} (k + c_1\left|\frac{1}{h}\right|_\alpha)(\chi^j|fh|_\alpha + F\|fh\|_{L^1(\lambda^n)}) + c_1\left|\frac{1}{h}\right|_\alpha \|f\|_{L^1(\mu)} \\
&\overset{4.2(iii)}{\leq} (k + c_1\left|\frac{1}{h}\right|_\alpha)\chi^j\Bigg(\sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int_{\mathbb{R}^n} \mathrm{osc}(f, B_\varepsilon(\cdot)) \operatorname*{ess\,sup}_{y\in B_\varepsilon(\cdot)} h(y) d\lambda^n \\
&+ \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int_{\mathbb{R}^n} \mathrm{osc}(h, B_\varepsilon(\cdot)) \operatorname*{ess\,inf}_{y\in B_\varepsilon(\cdot)} |f(y)| d\lambda^n\Bigg) + (k + c_1\left|\frac{1}{h}\right|_\alpha)(1+F)\|f\|_{L^1(\mu)} \\
&\overset{4.4(i)}{\leq} (k + c_1\left|\frac{1}{h}\right|_\alpha)\left(\chi^j(k + c_1|h|_\alpha)|f|_\alpha + (1 + F + c_1|h|_\alpha k)\|f\|_{L^1(\mu)}\right).
\end{aligned}$$

In [14, pp. 237–238] it is shown that $h \in \mathrm{Osc}_\alpha(\mathbb{X})$, if we show that also $\frac{1}{h} \in \mathrm{Osc}_\alpha(\mathbb{X})$ the result follows.

$$\begin{aligned}
\left|\frac{1}{h}\right|_\alpha &= \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int_{\mathbb{R}^n} \operatorname*{ess\,sup}_{y\in B_\varepsilon(\cdot)} \frac{1}{h(y)} - \operatorname*{ess\,inf}_{y\in B_\varepsilon(\cdot)} \frac{1}{h(y)} d\lambda^n \\
&= \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int_{\mathbb{R}^n} \frac{1}{\operatorname*{ess\,inf}_{y\in B_\varepsilon(\cdot)} h(y)} - \frac{1}{\operatorname*{ess\,sup}_{y\in B_\varepsilon(\cdot)} h(y)} d\lambda^n \\
&\leq \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int_{\mathbb{R}^n} \frac{\operatorname*{ess\,sup}_{y\in B_\varepsilon(\cdot)} h(y) - \operatorname*{ess\,inf}_{y\in B_\varepsilon(\cdot)} h(y)}{\operatorname*{ess\,inf}_{y\in B_\varepsilon(\cdot)} h(y) \operatorname*{ess\,sup}_{y\in B_\varepsilon(\cdot)} h(y)} d\lambda^n \\
&\leq k^2 |h|_\alpha < \infty.
\end{aligned}$$

It follows that $\frac{1}{h} \in \mathrm{Osc}_\alpha(\mathbb{X})$ and we conclude the proof.

□

We are now going to prove a theorem about the exponential hitting- and return-time statistics of piecewise expanding systems for a rare sequence of δ_l-balls. The proof will be similar to the one of Theorem 3.12 for Rychlik systems.

Theorem 4.9. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda^n, T, \xi)$ be an ergodic piecewise expanding system with invariant probability measure $\mu << \lambda^n$ with density $h := \frac{d\mu}{d\lambda}$ such that $\frac{1}{k} \leq h \leq k$ for some $k \in \mathbb{N}$. Furthermore let $x^* \in \mathbb{X}$ be non periodic and such that $\xi_n(x^*)$ is well defined for all $n \in \mathbb{N}$. If we look at a sequence $(A_l)_{l\geq 1}$ of rare events, such that each $A_l \subseteq \mathbb{X}$ is a δ_l-ball and $\mathrm{dist}(A_l, x^*) \to 0$ as $l \to \infty$, then for $R_l := \mu(A_l)\varphi_{A_l}$ it holds that*

(i) $R_l \overset{\mu}{\Longrightarrow} \mathcal{E}$

(ii) $R_l \overset{\mu_{A_l}}{\Longrightarrow} \mathcal{E}$.

Proof. We are going to use Theorem 1.25 and verify that (i)-(iii) hold. We start by choosing a fitting sequence τ_l.
We define $\delta := \sup_{l\in\mathbb{N}} \delta_l < \infty$ and

$$K := \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)} \max\left\{ \varepsilon_0^{(1-\alpha)} \sum_{j=1}^{\lfloor \frac{n+1}{2} \rfloor} \binom{n}{2j-1} 2\varepsilon_0^{(2j-2)} \delta^{n-(2j-1)}, \frac{2^n \varepsilon_0^n}{\varepsilon_0^\alpha} \right\}.$$

We will choose $\tau_l := \max\left\{ \underbrace{\left\lfloor \frac{-\log\mu(A_l) + \log\frac{K}{C}}{-\log\rho^j} \right\rfloor + 1}_{=:t_l},\ j_0 \right\}$, where ρ, j_0 and C are as in Theorem 4.8.

Then it follows that

$$j \geq t_l \Rightarrow j \geq \frac{-\log \mu(A_l) + \log \frac{K}{C}}{-\log \rho} \Leftrightarrow$$
$$j \geq \frac{\log(\mu(A_l)\frac{C}{K})}{\log \rho} \Leftrightarrow j \log \rho \leq \log \frac{C\mu(A_l)}{K}$$

and via exponentiation we get

$$\rho^j \leq \frac{C\mu(A_l)}{K} \text{ for all } j \geq t_l. \tag{4.6}$$

(iii) As mentioned before the sets $\overline{\mathrm{Osc}_{\alpha,c}}(\mathbb{X})$ are compact in $L^1(\lambda^n)$. It therefore suffices to show that $|\hat{T}^{\tau_l}\frac{1}{\mu(A_l)}\mathbb{1}_{A_l}|_\alpha \leq c$ for some $c > 0$ and all $l \in \mathbb{N}$.
We first need to show that $|\mathbb{1}_{A_l}|_\alpha \leq K$.

$$\begin{aligned}
|\mathbb{1}_{A_l}|_\alpha &= \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int_{\mathbb{R}^n} \operatorname*{ess\,sup}_{y\in B_\varepsilon(x)} \mathbb{1}_{A_l} - \operatorname*{ess\,inf}_{y\in B_\varepsilon(x)} \mathbb{1}_{A_l}\, d\lambda^n(x) \\
&= \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \left((\delta_l+\varepsilon)^n \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)} - (\max\{0,\delta_l-\varepsilon\})^n \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)} \right) \\
&\leq \max\left\{ \sup_{0<\varepsilon\leq\delta_l} \frac{\pi^{\frac{n}{2}}}{\varepsilon^\alpha\Gamma(\frac{n}{2}+1)}((\delta_l+\varepsilon)^n - (\delta_l-\varepsilon)^n), \sup_{\delta_l<\varepsilon\leq\varepsilon_0} \frac{\pi^{\frac{n}{2}}}{\varepsilon^\alpha\Gamma(\frac{n}{2}+1)}(\delta_l+\varepsilon)^n \right\} \\
&\leq \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)} \max\left\{ \varepsilon_0^{(1-\alpha)} \sum_{j=1}^{\lfloor\frac{n+1}{2}\rfloor} \binom{n}{2j-1} 2\varepsilon_0^{(2j-2)}\delta^{n-(2j-1)}, \frac{(2\varepsilon_0)^n}{\varepsilon_0^\alpha} \right\}.
\end{aligned}$$

At the beginning we used that the volume of a n-dimensional ball with radius r has size $\frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}r^n$.
In the last step we used the Binomial theorem and that $\varepsilon \leq \varepsilon_0$:

$$(\delta_l+\varepsilon)^n - (\delta_l-\varepsilon)^n = \varepsilon(\sum_{j=1}^n \binom{n}{j}\delta_l^{n-j}(\varepsilon^{j-1} + (-\varepsilon)^{j-1}).$$

By our choice of τ_l we have that

$$\begin{aligned}
|\hat{T}^j(\frac{1}{\mu(A_l)}\mathbb{1}_{A_l})|_\alpha &\overset{Thm4.8}{\leq} \rho^j|\frac{1}{\mu(A_l)}\mathbb{1}_{A_l}|_\alpha + C\|\frac{1}{\mu(A_l)}\mathbb{1}_{A_l}\|_{L^1(\mu)} \\
&\leq \frac{C\mu(A_l)}{K}\frac{1}{\mu(A_l)}K + C = 2C
\end{aligned}$$

is true for all $j \geq \tau_l$. In the last inequality, we used that $|\mathbb{1}_{A_l}|_\alpha \leq K$ and (4.6).

(i) There exists $l_0 \in \mathbb{N}$ such that for all $l \geq l_0$ it holds that $\tau_l = t_l$. We get

$$0 \leq \mu(A_l)\tau_l \leq \mu(A_l)\frac{-\log \mu(A_l)}{-\log \rho} + \mu(A_l)\frac{\log \frac{K}{C}}{-\log \rho} + \mu(A_l)$$

and the right hand side converges to 0 as $l \to \infty$.

(ii) We argue similarly as in Theorem 3.12. x^* is not periodic, therefore for fixed $M \in \mathbb{N}$ each of the points $x^*, ..., T^M x^*$ is different to each other and they have at least distance $\eta > 0$ to each other. Since $\operatorname{dist}(A_l, x^*) \to 0$ as $l \to \infty$ and T, $T^2, ..., T^M$ are continuous on $\xi_M(x^*)$ we choose l_0 large enough such that $A_l \cap (\cup_{j=1}^M T^j A_l) = \emptyset$ for all $l \geq l_0$. With this we get

$$\begin{aligned}
\mu_{A_l}(\varphi_{A_l} \leq \tau_l) &= \sum_{j=M+1}^{\tau_l} \mu_{A_l}(\underbrace{\varphi_{A_l} = j}_{\subseteq T^{-j}A_l}) \leq \sum_{j=M+1}^{\tau_l} \frac{1}{\mu(A_l)} \int\limits_{\mathbb{X}} (\mathbb{1}_{A_l} \circ T^j)\mathbb{1}_{A_l}\, d\mu \\
&= \sum_{j=M+1}^{\tau_l} \frac{1}{\mu(A_l)} \int\limits_{\mathbb{X}} \mathbb{1}_{A_l} \hat{T}^j \mathbb{1}_{A_l}\, d\mu \leq \sum_{j=M+1}^{\tau_l} \|\hat{T}^j \mathbb{1}_{A_l}\|_\infty \\
&\leq \sum_{j=M+1}^{\tau_l} c_1 \|\hat{T}^j \mathbb{1}_{A_l}\|_\alpha = \sum_{j=M+1}^{\tau_l} c_1 (\int\limits_{\mathbb{X}} \hat{T}^j \mathbb{1}_{A_l}\, d\lambda^n + |\hat{T}^j \mathbb{1}_{A_l}|_\alpha) \\
&\overset{4.8}{\leq} c_1(\tau_l k \mu(A_l) + \sum_{j=M+1}^{\tau_l} C\|\mathbb{1}_{A_l}\|_{L^1(\mu)} + \rho^j |\mathbb{1}_{A_l}|_\alpha) \\
&\leq c_1(k + C)\tau_l \mu(A_l) + \frac{K\rho^{M+1}}{1-\rho}
\end{aligned}$$

for all $l \geq l_0$. Here we used that $\frac{1}{k} \leq \frac{d\mu}{d\lambda^n} \leq k$. Since M is arbitrary the left hand side converges to 0 as $l \to \infty$ and we conclude the proof.

□

It is also possible to get exponential hitting- and return-time statistics if we look at a rare sequence $(A_l)_{l\geq 1}$ of n-dimensional intervals. In this case we have to assume that the intervals are contained in δ_l-balls which get close to a non periodic point x^* as $l \to \infty$ and such that $\delta_l \to 0$ as $l \to \infty$. We need that the intervals are close to x^* as whole so we can use the continuity of $T, .., T^M$ around x^*, which is why we make the assumption that they are contained in δ_l-balls.

Theorem 4.10. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \lambda^n, T, \xi)$ be an ergodic piecewise expanding system with invariant probability measure $\mu << \lambda^n$ with density $h := \frac{d\mu}{d\lambda}$ such that $\frac{1}{k} \leq h \leq k$ for some $k \in \mathbb{N}$. Furthermore let $x^* \in \mathbb{X}$ be non periodic and such that $\xi_n(x^*)$ is well defined for all $n \in \mathbb{N}$. If we look at a sequence $(A_l)_{l\geq 1}$ of rare events, such that each $A_l \in \mathcal{A}$ is a non-degenerate n-dimensional interval which is contained in a δ_l-ball B_{δ_l} such that $\delta_l \to 0$ and $\operatorname{dist}(B_{\delta_l}, x^*) \to 0$ as $l \to \infty$, then for $R_l := \mu(A_l)\varphi_{A_l}$ it holds that*

(i) $R_l \overset{\mu}{\Longrightarrow} \mathcal{E}$

(ii) $R_l \overset{\mu_{A_l}}{\Longrightarrow} \mathcal{E}$.

The proof of this theorem is to great extent similar to the one of Theorem 4.9, which is why we will omit parts of it. We still need to show, that $|\mathbb{1}_{A_l}|_\alpha < K$ for some $K \geq 0$ and all $l \in \mathbb{N}$.

Proof. We start by defining $\delta := \sup_{l\in\mathbb{N}} \delta_l < \infty$ and
$K := 4\varepsilon_0^{1-\alpha} n 2^{n-1}(\varepsilon_0 + \delta)^{n-1}$ and choosing $\tau_l := \max\left\{\left\lfloor \frac{-\log \mu(A_l) + \log \frac{K}{C}}{-\log \rho} \right\rfloor + 1\,, j_0\right\}$, where ρ, j_0 and C

are as in Theorem 4.8. As mentioned before we are going to omit most of the proof. We will only show that $|\mathbb{1}_{A_l}|_\alpha \leq K$. Assume that $\overline{A_l} = \bigtimes_{i=1}^{n} [a_i, b_i]$ and define

$A_l^{\varepsilon,m} := \left(\bigtimes_{i=1}^{m-1} [a_i - \varepsilon, b_i + \varepsilon] \right) \times ([a_m - \varepsilon, a_m + \varepsilon]) \cup [b_m - \varepsilon, b_m + \varepsilon]) \times \left(\bigtimes_{i=m+1}^{n} [a_i - \varepsilon, b_i + \varepsilon] \right)$. We infer

$$
\begin{aligned}
|\mathbb{1}_{A_l}|_\alpha &= \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int_{\mathbb{R}^n} \operatorname*{ess\,sup}_{y\in B_\varepsilon(x)} \mathbb{1}_{A_l} - \operatorname*{ess\,inf}_{y\in B_\varepsilon(x)} \mathbb{1}_{A_l} \, d\lambda^n(x) \\
&\leq \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int_{\mathbb{R}^n} \sum_{m=1}^{n} \mathbb{1}_{A_l^{\varepsilon,m}} \, d\lambda^n \\
&\leq \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \sum_{m=1}^{n} (4\varepsilon) \prod_{i=1,\, i\neq m}^{n} (b_i + 2\varepsilon - a_i) \\
&\leq 4\varepsilon_0^{1-\alpha} n 2^{n-1} (\varepsilon_0 + \delta)^{n-1}.
\end{aligned}
$$

The first inequality holds since for every $x \in \partial A_l$ and for all $\varepsilon > 0$ we get that $B_\varepsilon(x) \subseteq \bigcup_{m=1}^{n} A_l^{\varepsilon,m}$ and the oscillation of the indicator function $\mathbb{1}_{A_l}$ can be greater than zero only on this set. In the last inequality we used that $A_l \subseteq B_{\delta_l}(y) \subseteq B_\delta(y)$ for some $y \in \mathbb{X}$.

$\square$

A Induced Systems

In this appendix we are going to include some theory about induced systems. We recall Definition 3.14 and assume that $(\mathbb{X}, \mathcal{A}, \nu, T)$ is a probability-preserving dynamical system and $Y \in \mathcal{A}$ a recurrent set with $\nu(Y) > 0$. For a set $A \in \mathcal{A} \cap Y$ we denote by φ_A^Y the first return time for the induced system on Y.

Our goal is to prove that if the hitting time distributions for the induced system $\mathcal{S}_Y$ converge to some distribution R, one can deduce that the hitting time statistic of the original system $\mathcal{S}$ is given by R if the system $\mathcal{S}$ is ergodic.. We used this in the discussion of the example in Chapter 3, Remark 3.16.

Before this we include the proof of Proposition A.2, which we omitted before. It is the immediate consequence of the relation of the invariant probability measure of the induced system $\mathcal{S}_Y$ and the invariant probability measure of the original system $\mathcal{S}$, if $\mathcal{S}$ is conservative and ergodic.

Theorem A.1. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \nu, T)$ be a null-preserving system and $Y \in \mathcal{A}$ a recurrent set such that $\nu(Y) > 0$. If the induced system $\mathcal{S}_Y$ has an invariant probability measure $\mu^Y << \nu_Y$ then $\mathcal{S}$ has*

$$\mu(A) := \sum_{n\geq 0} \mu^Y(Y \cap \{\varphi_Y > n\} \cap T^{-n}A) = \sum_{n\geq 0} \mu^Y|_{\{\varphi_Y > n\}\cap Y} \circ T^{-n}(A),$$

for $A \in \mathcal{A}$ as an invariant measure $\mu << \nu$.

Proof. Take $A \in \mathcal{A}$, we want to show that $\mu(A) = \mu(T^{-1}A)$.

$$\begin{aligned}
\mu(T^{-1}A) &= \sum_{n\geq 0} \mu^Y(Y \cap \{\varphi_Y > n\} \cap T^{-(n+1)}A) \\
&= \sum_{n\geq 0} \mu^Y(Y \cap \{\varphi_Y = n+1\} \cap T^{-(n+1)}A) + \sum_{n\geq 0} \mu^Y(Y \cap \{\varphi_Y > n+1\} \cap T^{-(n+1)}A) \\
&= \sum_{n\geq 0} \mu^Y(\{\varphi_Y = n+1\} \cap T_Y^{-1}A) + \mu(A) - \mu^Y(Y \cap A) \\
&= \mu^Y(T_Y^{-1}A) + \mu(A) - \mu^Y(A) = \mu(A).
\end{aligned}$$

We used the invariance of μ^Y in the last step. $\square$

Now the Generalized Kac-formula easily follows.

Proposition A.2 (Generalized Kac formula). *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic conservative measure-preserving dynamical system. Then for all $A \in \mathcal{A}$ for which $0 < \mu(A) < \infty$ holds and for all measurable $f : \mathbb{X} \longrightarrow [0, \infty]$*

$$\int_A \sum_{j=0}^{\varphi_A - 1} f \circ T^j d\mu = \int_{\mathbb{X}} f \, d\mu \tag{A.1}$$

holds.

Proof. We look at the induced system $\mathcal{S}_A$ and know that $\mu|_A$ is a finite invariant measure for the induced system. From Theorem A.1 we know that $\tilde{\mu} = \sum_{k\geq 0} \mu|_{A\cap\{\varphi_A > k\}\cap A} \circ T^{-k}$ is an invariant measure of the original system, which coincides on $A \cap \mathcal{A}$ with $\mu|_A$. We will now show that $\tilde{\mu}$ is σ-finite and infer that $\mu = \tilde{\mu}$. The sets $X_n := \{\varphi_A = n\}$ are such that $\mathbb{X} = \bigcup_{n\geq 1} X_n \mod \mu$ since $\mathcal{S}$ is conservative. Moreover we get that for $n \in \mathbb{N}$

$$\begin{aligned}
\sum_{k\geq 0} \mu|_{A\cap\{\varphi_A > k\}} \circ T^{-k}(X_n) &\leq \sum_{k\geq 0} \mu|_{A\cap\{\varphi_A > k\}}(\{\varphi_A < k\} \cup \{\varphi_A = k+n\}) \\
&\leq \sum_{k\geq 0} \mu|_{A\cap\{\varphi_A > k\}}(\{\varphi_A = k+n\}) \leq \mu|_A(A) < \infty.
\end{aligned}$$

$\mathcal{S}$ is ergodic and we infer that there is only one (up to scalar multiples) invariant measure and since $\mu = \tilde{\mu}$ on $A \cap \mathcal{A}$ we get $\mu = \tilde{\mu}$ on $\mathcal{A}$.
Now take some $f \geq 0$ measurable.

$$\begin{aligned}
\int_{\mathbb{X}} f \, d\mu &= \int_{\mathbb{X}} f \, d \sum_{k\geq 0} \mu|_{A\cap\{\varphi_A > k\}} \circ T^{-k} = \sum_{k\geq 0} \int_{\mathbb{X}} f \circ T^k \, d\mu|_{A\cap\{\varphi_A > k\}} \\
&= \int_A \sum_{k\geq 0} (f \circ T^k) \mathbb{1}_{\{\varphi_A > k\}} \, d\mu = \int_A \sum_{k=0}^{\varphi_A - 1} f \circ T^k \, d\mu
\end{aligned}$$

where we used the monotone convergence Theorem to interchange the sum and integral. $\square$

The next theorem is about the connection between the hitting-time statistic for the induced system and the one for the original system. It can be found in [8, pp. 167–168].

Theorem A.3. *Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic probability-preserving system and $Y \in \mathcal{A}$ such that $\mu(Y) > 0$. We denote $\mu_Y := \frac{1}{\mu(Y)}\mu|_Y$ Take a sequence of rare events $(A_l)_{l\geq 1} \subseteq \mathcal{A} \cap Y$ such that*

$$\mu_Y(A_l)\varphi^Y_{A_l} \overset{\mu_Y}{\Longrightarrow} R.$$

Then also

$$\mu(A_l)\varphi_{A_l} \overset{\mu}{\Longrightarrow} R.$$

Remark A.4. We only need the implication that when an induced system admits expontential hitting-time statistics then so does the original system, but an even stronger version of the above theorem holds with the same assumptions. One gets that one of these hitting-time distributions converges to R iff the other one does so as well.

Proof. From Theorem 1.16 we infer that $\mu(A_l)\varphi_{A_l} \overset{\mu}{\Longrightarrow} R$ iff $\mu(A_l)\varphi_{A_l} \overset{\mu_Y}{\Longrightarrow} R$ and with Remark 1.2 it suffices to look at the convergence

$$\mu_Y(\mu(A_l)\varphi_{A_l} \leq t) \overset{l\to\infty}{\longrightarrow} F(t),$$

for all continuity points $t \in [0, \infty)$ of F, where F is the subdistribution function of R.
Take some continuity point $t > 0$ and some $\varepsilon > 0$. Then we find $\delta > 0$ such that

$$F(t) - \frac{\varepsilon}{4} < F(e^{-\delta}t) \leq F(t) \leq F(e^{\delta}t) < F(t) + \frac{\varepsilon}{4}.$$

We assume w.l.o.g that $e^{-\delta}t$ and $e^{\delta}t$ are continuity points of F (monotone functions have at most countable many points of disontinuity).
From the ergodicity of the system $\mathcal{S}$ we infer that $\mathcal{S}_Y$ is ergodic (see Remark A.5) and we know that

$$\frac{1}{n}\sum_{k=0}^{n-1} \varphi_Y \circ T_Y^k \overset{n\to\infty}{\longrightarrow} \int\limits_Y \varphi_Y d\mu_Y \text{ a.e. on } Y.$$

From the Generalized Kac-formula A.2 (the system $\mathcal{S}$ is conservative since it is probability-preserving) we infer with $f = \mathbb{1}_{\mathbb{X}}$ that

$$\int\limits_Y \varphi_Y d\mu_Y = \frac{1}{\mu(Y)}\int\limits_Y \sum_{k=0}^{\varphi_Y - 1} \mathbb{1}_{\mathbb{X}} \circ T^k d\mu = \frac{1}{\mu(Y)}.$$

We define sets

$$E_N := \{\sum_{k=0}^{n-1} \varphi_Y \circ T_Y^k \geq e^{-\delta}\frac{n}{\mu(Y)} \text{ for all } n \geq N\} \tag{A.2}$$

which are increasing and such that $\mu_Y(E_N^c) \overset{N\to\infty}{\longrightarrow} 0$ with the above. We fix $N \in \mathbb{N}$ such that $\mu_Y(E_N^c) < \frac{\varepsilon}{4}$ and define sets

$$F_l^N := \{\varphi_{A_l}^Y \geq N\} \text{ for } l \geq 1. \tag{A.3}$$

Then $F_l^N = Y \cap \bigcap_{k=1}^{N-1} T_Y^{-k} A_l^c$ and we infer that $\mu_Y((F_l^N)^c) \leq \sum_{k=1}^{N-1} \mu_Y(T_Y^{-k} A_l) = N\mu_Y(A_l)$ since T_Y preserves μ_Y. The right hand side converges to 0 as $l \to \infty$ since $(A_l)_{l\geq 1}$ is asymptotically rare. Now we can choose $l_0 \in \mathbb{N}$ such that $\mu_Y((F_l^N)^c) < \frac{\varepsilon}{4}$ and $\mu_Y(\mu_Y(A_l)\varphi_{A_l}^Y \leq te^\delta) < F(te^\delta) + \frac{\varepsilon}{4}$ for all $l \geq l_0$. The relation between φ_{A_l} and $\varphi_{A_l}^Y$ is given by

$$\varphi_{A_l} = \sum_{k=0}^{\varphi_{A_l}^Y - 1} \varphi_Y \circ T_Y^k \text{ on } Y.$$

On $F_l^N \cap E_N$ we thus get

$$\varphi_{A_l} \geq \frac{e^{-\delta}}{\mu(Y)} \varphi_{A_l}^Y. \tag{A.4}$$

We assumed that $\mu_Y(\mu_Y(A_l)\varphi_{A_l}^Y \leq t) \to F(t)$ as $l \to \infty$ and infer

$$\begin{aligned} \mu_Y(\mu(A_l)\varphi_{A_l} \leq t) &\leq \mu_Y\left(\frac{\mu(A_l)}{\mu(Y)}\varphi_{A_l}^Y \leq te^\delta\right) + \mu_Y((F_l^N)^c) + \mu_Y(E_N^c) \\ &< F(te^\delta) + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < F(t) + \varepsilon. \end{aligned}$$

By defining $E_N := \{\sum_{k=0}^{n-1} \varphi_Y \circ T_Y^k \leq e^\delta \frac{n}{\mu(Y)}$ for all $n \geq N\}$ one obtains the estimate $F(t) - \varepsilon < \mu_Y(\mu(A_l)\varphi_{A_l} \leq t)$ in a similar way and since $\varepsilon > 0$ was arbitrary we conclude the proof. □

Remark A.5. Let $\mathcal{S} = (\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic probability-preserving system and $Y \in \mathcal{A}$ such that $\mu(Y) > 0$. Then $\mathcal{S}_Y$ is ergodic. Assume otherwise, then there exists $A \in \mathcal{A} \cap Y$ with $0 < \mu_Y(A) < 1$ such that

$$T_Y^{-1} A = A. \tag{A.5}$$

Define $\tilde{A} := \bigcup_{n\geq 0} T^{-n} A$. We know from (A.5) that

$$A \subseteq \bigcup_{n\geq 1} T^{-n} A$$

and therefore $T^{-1}\tilde{A} = \tilde{A}$. From the ergodicity of $\mathcal{S}$ it follows that $\tilde{A} = \mathbb{X} \mod \mu$ and we infer that

$$Y = \bigcup_{n\geq 0} T^{-n} A \cap Y = \bigcup_{n\geq 0} T_Y^{-n} A = A \mod \mu$$

which is a contradiction to the assumption $\mu_Y(A) < 1$.

References

[1] Jon Aaronson. *An Introduction to Infinite Ergodic Theory*. American Mathematical Society, 1997.

[2] Patrick Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, Inc., 1999.

[3] A. Boyarski and P. Gora. *Laws of Chaos*. Birkhäuser, 1997.

[4] H. Bruin et al. "Return time statistics via inducing". In: *Ergodic Theory and Dynamical Systems* 23(4) (2003), pp. 991–1013.

[5] K. Dajani and C. Kalle. *A First Course in Ergodic Theory*. CRC Press, 2021.

[6] N. Dunford and J. Schwartz. *Linear Operators, Part I: General Theory*. Wiley Interscience, 1988.

[7] N. Haydn, Y. Lacroix, and S. Vaienti. "Hitting and return times in ergodic dynamical systems". In: (2004).

[8] N. Haydn, N. Winterberg, and R. Zweimüller. "Return-time statistics, Hitting-time statistics and inducing". In: *Ergodic Theory, Open Dynamics, and Coherent Structures* (2014), pp. 217–227.

[9] M. Hirata, B. Saussol, and S. Vaienti. "Statistics of return times: A general framework and new applications". In: (1998).

[10] F. Hofbauer and G. Keller. "Ergodic properties of invariant measures for piecewise monotonic transformations". In: (1982).

[11] G. Keller. "Generalized Bounded Variation and Applications to Piecewise Monotonic Transformations". In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 69 (1985), pp. 461–478.

[12] David Lecomte. "Weak Topologies". In: (2006). URL: `https://perso.crans.org/lecomte/Math/WeakTopologies.pdf`.

[13] Marek Rychlik. "Bounded variation and invariant measures". In: *Studia Mathematica* 76 (1983), pp. 69–80. URL: `http://eudml.org/doc/218512`.

[14] B. Saussol. "Absolutely continuous invariant measures for multidimensional expanding maps". In: *Israel Journal of Mathematics* 116 (2000), pp. 223–248.

[15] Roland Zweimüller. "Hitting Times and Positions in Rare Events". In: (2022). URL: `https://arxiv.org/pdf/1810.10381.pdf`.

[16] Roland Zweimüller. "Mixing Limit Theorems for Ergodic Transformations". In: *Journal of Theoretical Probability* 20 (2007), pp. 1059–1071.