# MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Exponential laws for rare events in dynamical systems"

verfasst von / submitted by

Sebastian Fischer, BSc

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

Master of Science (MSc.)

Wien, 2016 / Vienna 2016

| | |
|---|---|
| Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet: | A 066821 |
| Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet: | Masterstudium Mathematik UG2002 |
| Betreut von / Supervisor: | Assoz.-Prof. Privatdoz. Mag. Dr. Roland Zweimüller |

# Contents

# Abstract

A measurable discrete dynamical system is a mathematical structure consisting of a measurable space $(\mathbb{X}, \mathcal{A}, \lambda)$ and a map $T : \mathbb{X} \to \mathbb{X}$ whose iterations describe the development in time. Such systems are often motivated from a physical problem like gas particles in a sealed container. Although $T$ is a deterministic function, in the system itself we can find plenty of randomness if we do not specify a definite starting point $x_0 \in \mathbb{X}$ but model our uncertainty with a probability density on $\mathbb{X}$. Despite the short-term chaotic behaviour in complicated systems we can frequently make statements about the long-time evolution. To prove such often very strong results we usually want two things.

First there should exist an invariant probability measure denoted by $\mu$ in this work such that $\mu = \mu \circ T$. For a general system it is not always clear if such a measure exists and even if so, it might be hard to find one. We avoid such difficulties by restricting our analysis to such systems where we already know the desired $\mu$. For practical reasons, we want $\mu$ to be absolutely continuous to the reference measure $\lambda$ about which we eventually forget by looking at $(\mathbb{X}, \mathcal{A}, \mu)$ in the first place. Often one can think of $\lambda$ to be the Lebesgue measure.

The second big assumption is the ergodicity of $T$. If a set $A \in \mathcal{A}$ is invariant, meaning $A = T^{-1}A$ holds, we want it to be trivial, i.e. $\mu(A) \in \{0, 1\}$. This might sound counter-intuitive since it forces the system to be even more chaotic, but it turns out that we need this level of chaos to avoid certain dependence structures which impede general statements.

In this paper, we examine so-called rare events - meaning a sequence $(A_l)_{l \geq 0} \in \mathcal{A}$ such that $\mu(A_l) \to 0$ holds. Considering their first hitting time $\varphi_{A_l}$ we observe that this random variable often converges to some variant of the exponential distribution if we use the right scaling. Furthermore we will show that under the right conditions, we see Poisson asymptotics, i.e. convergence to a point process with exponentially distributed waiting times.

As so often in higher mathematics, a certain problem permits very different methods of approach. In chapter 1 we prove some results for the convergence of the pro-

cess following the work of my supervisor Roland Zweimüller. Here we will start with showing a connection between first hitting and first return time and eventually we will also use this connection to establish a efficient way to prove the existence of an exponential law. Furthermore we will extend this ideas to point processes, where we look at consecutive visits to a small set. It turns out that many ideas from the one-dimensional case can be adapted and transferred to the level of processes.

In chapter 2, we completely change our point of view to talk about the hitting time. We identify the $A_l$'s with shrinking $\varepsilon$-balls around a certain point $x^*$. In each step, we look at some specific function of the distance to $x^*$ in such a way that if the maximum of those first $n$ measurements exceeds $\varepsilon$, we have hit the set $A_l$ up to time $n$. So we moved the problem to the study of extremal behaviour of functions. This idea is mostly motivated by the papers [3,4,7,5,8] of Jorge Milhazes Freitas.

Finally I would like to thank Professor Zweimüller for the supervision of my work. He answered my questions and gave me hints in countless meetings, and it was him in die end who created my curiosity about dynamical systems.

Sebastian Fischer
Wien, September 2, 2016

# Chapter 1

# Exponential laws and Poisson asymptotics

This chapter follows some papers of Roland Zweimüller. Let $S = (\mathbb{X}, \mathcal{A}, \lambda, T)$ be an ergodic dynamical system. We want it (or rather the map $T$) to be 0-preserving (i.e. $\lambda \circ T^{-1} \ll \lambda$) and we assume there exists an invariant probability measure $\mu = \mu \circ T^{-1}$, which in addition is absolutely continuous with respect to $\lambda$. In this setting we look at so called rare events.

**Definition 1.0.1.** We call a sequence $(A_l)_{l\geq 0}$ in an 0-preserving system $S = (\mathbb{X}, \mathcal{A}, \lambda, T)$ a **sequence of asymptotically rare events** if all $A_l$ are measurable, $\lambda(A_l)$ is positive and at the same time converging to 0.

For those sequences we look at the **first hitting time** of the sets, defined by

$$\varphi_{A_l}(x) := \min\{n \geq 1 : T^n x \in A_l\}.$$

If we restrict $\varphi_{A_l}$ to the set $A_l$, we get the **first return time**. Unsurprisingly, those two random variables are deeply connected, as we will see in section 1.2. The main difference is that while the first hitting time is defined on $(\mathbb{X}, \mathcal{A}, \mu)$, the first return time lives on $(A_l, \mathcal{A} \cap A_l, \mu_{A_l})$, where

$$\mu_{A_l}(B) := \frac{\mu(A_l \cap B)}{\mu(A_l)}$$

denotes the normalised trace of $\mu$. Consequently the measures change with the index $l$.

Another important property of a dynamical system is called conservativity, which demands that any set with positive measure is recurrent, i.e. almost every point in it will return. By Poincaré's recurrence theorem, a finite invariant measure already ensures conservativity. Note that in the case of a conservative ergodic system, the hitting time is finite almost everywhere.

## 1.1 Interval maps - A first example

We start with introducing a class of examples, so called **Folklore interval maps**. So let $\mathbb{X}$ be the unit interval and $\lambda$ the Lebesgue measure. We assume there is a countable partition $\xi$ which divides the interval into subintervals called cylinder sets. For every cylinder $Z$, we want the branch $T|_Z : Z \to TZ$ to be injective with inverse $v_z : TZ \to Z$. An easy example is the so called *doubling map*, given by $Tx = 2x \bmod 1$.

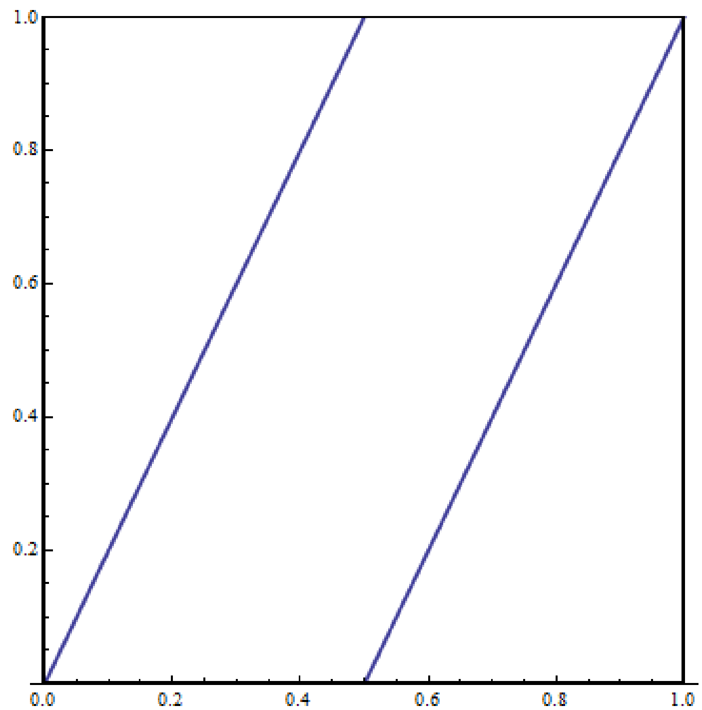


Figure 1.1: The doubling map

The doubling map is a special case of a bigger class of examples, which have very nice properties. We call an interval map a **Folklore** one, if the following assumptions are fulfilled:

F1) For all cylinders $Z \in \xi$ the branch $T|_Z$ is twice continuously differentiable.

F2) We want uniform expansion, that is the derivative of $T$ should have absolute value greater some constant $\rho > 1$.

F3) We want Adler's condition to be true, this is

$$A_T := \sup_Z \left| \frac{T''}{(T')^2} \right| < \infty.$$

F4) $T$ has full branches, i.e. $TZ = \mathbb{X}$.

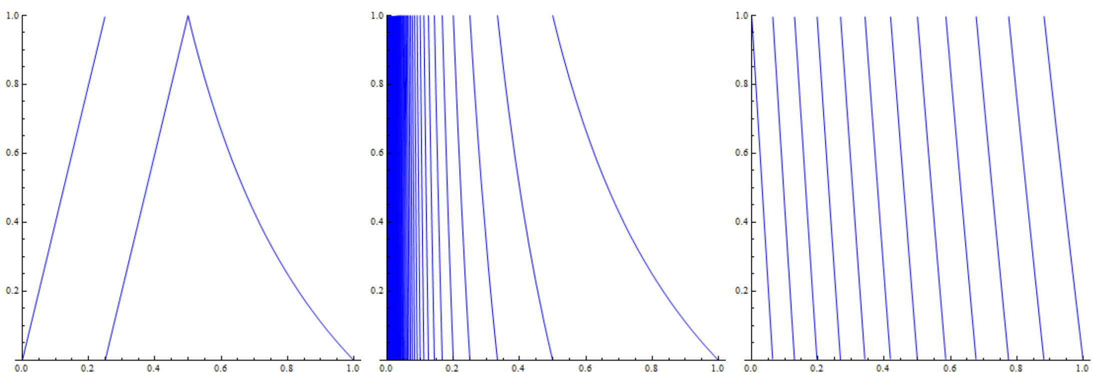

Figure 1.2: Different Folklore maps

Those assumptions have far-reaching consequences. In [18], the following theorem is shown:

**Theorem 1.1.1.** *A system $S = (\mathbb{X}, \mathcal{A}, \lambda, T, \xi)$ with a map $T$ that fulfils $F1)$-$F4)$ and $\lambda(\mathbb{X}) = 1$ has the following properties*

*a) $S$ is ergodic.*

*b) $S$ is recurrent with a unique invariant probability density $h$ with*

$$0 < \inf h \leq \sup h < \infty$$

*which is additionally Lipschitz continuous.*

*c) For all Lipschitz densities u, the mapped version $\hat{T}^n u$, where*

$$\hat{T}u := \frac{(du \odot \mu) \circ T^{-1}}{\mu}$$

*denotes the transfer operator, converge to h uniformly and exponentially fast, i.e. there is some $q \in ]0,1[$ such that*

$$||\hat{T}^n u - h||_\infty \leq K \cdot q^n$$

*where K only depends on the regularity of u.*

In [16], another important property of our Folklore maps is shown, the so-called exponential decay of correlations. To understand this, we need the concept of bounded variation.

**Definition 1.1.2.** Let

$$\mathfrak{P} = \{P = \{x_0, \ldots, x_{n_P}\} | P \text{ is partition of } [0,1]\}$$

be the set of all partitions of $[0,1]$. The total variation of a real-valued function $f$ on $[0,1]$ is the quantity

$$V(f) := \sup_{P \in \mathfrak{P}} \sum_{k=0}^{n_P - 1} |f(x_k) - f(x_{k+1})|.$$

We can use this to define the norm

$$||f||_{BV} := ||f||_{\mathcal{L}_1(\mu)} + V(f),$$

which makes $\mathcal{BV} := \{f \in \mathcal{L}_1(\mu) : V(f) < \infty\}$ a Banach space. Using this norm, we can finally state what decay of correlations means.

**Theorem 1.1.3.** *Let $S = (\mathbb{X}, \mathcal{A}, \lambda, T, \xi)$ be a Folklore system. For any two observables $\phi \in \mathcal{L}_1(\mu)$ and $\psi \in BV$ there exists some constant K and $q \in ]0,1[$ such that*

$$\left| \int_{\mathbb{X}} \phi \circ T^n \psi \; d\mu - \int_{\mathbb{X}} \phi \; d\mu \int_{\mathbb{X}} \psi \; d\mu \right| \leq K ||\phi||_{\mathcal{L}_1(\mu)} ||\psi||_{BV} \cdot q^n$$

*holds.*

## 1.2 Duality of hitting and return time

The idea is now to look for normalising constants $\gamma_l$ such that $\gamma_l \varphi_{A_l}$ converges in distribution to some non-trivial random variable $R$. Convergence in distribution for these real-valued random variables means

$$\mu(\gamma_l \varphi_{A_l} \leq t) \xrightarrow{l \to \infty} F(t)$$

for all continuity points $t$ of $F(t)$, where $F$ represents the distribution function of $R$. We denote this situation with respect to the measure with

$$\gamma_l \varphi_{A_l} \overset{\mu}{\Longrightarrow} R.$$

In [17] we see that this kind of convergence does in fact come from the system itself and does not depend that strongly on the measure we choose. This fact is captured by the following proposition.

**Proposition 1.2.1** (Strong distributional convergence)**.** *Let $S = (\mathbb{X}, \mathcal{A}, \lambda, T)$ be a conservative ergodic system on a $\sigma$-finite space and $\mu$ some probability measure which is absolutely continuous to $\lambda$. For a sequence of asymptotically rare events $(A_l)_{l \geq 0}$ and normalising constants $\gamma_l$ and any non-negative random variable $R$*

$$\gamma_l \varphi_{A_l} \overset{\mu}{\Longrightarrow} R \quad \textit{implies} \quad \gamma_l \varphi_{A_l} \overset{\mathbb{P}}{\Longrightarrow} R \quad \textit{for all} \quad \mathbb{P} \ll \lambda.$$

*We use the notation $\gamma_l \varphi_{A_l} \overset{\mathcal{L}(\lambda)}{\Longrightarrow} R$ in this case.*

A proof for this can be found in [9]. Regarding the question of the right scaling we have Kac' lemma, which tells us that the expectation of the first return is given by the inverse of the measure of the set, that is

$$\int_{A_l} \varphi_{A_l} d\mu_{A_l} = \frac{1}{\mu(A_l)}.$$

Actually this gives the right hint for the choice of $\gamma_l$. The product $\mu(A_l)\varphi_{A_l}$ is a random variable with expectation 1 on $(A_l, \mathcal{A} \cap A_l, \mu_{A_l})$ and it turns out that this normalisation guarantees useful limit random variables.

Now as one can suspect, the hitting and return time are deeply connected since the return time is simply the restriction of the hitting time viewed through another measure. Actually in [9] the following theorem is shown.

**Theorem 1.2.2.** *Let $S = (\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system and $(A_l)_{l\geq 0}$ a sequence of asymptotically rare events. Then*

$$\mu(A_l)\varphi_{A_l} \overset{\mu}{\Longrightarrow} R \text{ for some random variable } \varphi \text{ in } [0,\infty] \tag{1.1}$$

*if and only if*

$$\mu(A_l)\varphi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \tilde{R} \text{ for some random variable } \tilde{\varphi} \text{ in } [0,\infty]. \tag{1.2}$$

*In this case the sub-probability functions $F$ and $\tilde{F}$ of $\varphi$ and $\tilde{\varphi}$ satisfy the equality*

$$\int_0^t \big(1 - \tilde{F}(s)\big)ds = F(t).$$

The functional equation above determines the distribution functions uniquely. In the special case where $F$ is standard exponentially distributed, solving a ODE yields that $\tilde{F}$ is exponential as well. Following [17], we need the next lemma to prove the theorem. It provides an estimate for the law of the hitting time compared with the law of the return time. We state a more general version from [22], in fact we only need the special case for the proof. Nevertheless, the main statement will be useful later. To formulate the lemma, we need the following definition.

**Definition 1.2.3.** Let $S = (\mathbb{X}, \mathcal{A}, \mu, T)$ be a dynamical system and $A \in \mathcal{A}$. Define

$$T_A x := T^{\varphi_A(x)}x,$$

then $T_A : \mathbb{X} \to A$ is called the **first entrance map** and $T_A : A \to A$ the **first return map**.

With this, we can state the following

**Lemma 1.2.4.** *Let $S = (\mathbb{X}, \mathcal{A}, \mu, T)$ be a conservative ergodic measure preserving system and $A$ and $B$ measurable sets with $B \subseteq A$ and $\mu(B) > 0$. Then we have for every $n \geq 1$*

$$\mu\big(\{\varphi_A = n\} \cap T_A^{-1}B\big) = \mu\big(A \cap \{\varphi_A \geq n\} \cap T_A^{-1}B\big)$$

*and for every $t > 0$*

$$\begin{aligned} \mu\Big(\{\mu(A)\varphi_A \leq t\} \cap T_A^{-1}B\Big) &\leq \int_0^t \mu_A\Big(\{\mu(A)\varphi_A > s\} \cap T_A^{-1}B\Big)ds \\ &\leq \mu\Big(\{\mu(A)\varphi_A \leq t\} \cap T_A^{-1}B\Big) + \mu(A). \end{aligned} \tag{1.3}$$

*In particular, for $B = A$, we get*

$$\mu\Big(\mu(A)\varphi_A \leq t\Big) \leq \int_0^t \mu_A\Big(\mu(A)\varphi_A > s\Big)ds \leq \mu\Big(\mu(A)\varphi_A \leq t\Big) + \mu(A).$$

*Proof.* Denote with $A(n) := A^c \cap \{\varphi_A = n\}$ for $n \geq 1$ and $A(0) = A$. Then we get

$$\begin{aligned}
T^{-1}\big(A(n)\big) &= T^{-1}\big(A^c \cap \{\varphi_A = n\}\big) \\
&= T^{-1}A^c \cap \big\{x \in \mathbb{X} : T^{n+1}x \in A, T^j x \notin A \text{ for } j \in \{2, \dots, n\}\big\} \\
&= \big\{x \in \mathbb{X} : T^{n+1}x \in A, T^j x \notin A \text{ for } j \in \{1, \dots, n\}\big\} \\
&= \{\varphi_A = n+1\} \\
&= \big(A \cap \{\varphi_A = n\}\big) \uplus A(n+1). \qquad (1.4)
\end{aligned}$$

Let $\hat{T}$ be the transfer operator with respect to $\mu$, that is

$$\int f \cdot \hat{T}u \, d\mu = \int (f \circ T) \cdot u \, d\mu \text{ for } f \in \mathcal{L}_\infty(\mu) \text{ and } u \in \mathcal{L}_1(\mu).$$

By this definition, we get for indicator functions $\mathbb{1}_A = \hat{T}\mathbb{1}_{T^{-1}A}$ almost everywhere. Hence by (1.4) it follows that

$$\mathbb{1}_{A(n)} = \hat{T}\mathbb{1}_{A \cap \{\varphi_A = n+1\}} + \hat{T}\mathbb{1}_{A(n+1)}$$

holds almost everywhere for $n \geq 0$. Now we can repeat this idea to get

$$\begin{aligned}
\hat{T}^n \mathbb{1}_{\{\varphi_A = n\}} &= \hat{T}^n \mathbb{1}_{A \cap \{\varphi_A = n\}} + \hat{T}^n \mathbb{1}_{A(n)} \\
&= \hat{T}^n \mathbb{1}_{A \cap \{\varphi_A = n\}} + \hat{T}^{n+1} \mathbb{1}_{A \cap \{\varphi_A = n+1\}} + \hat{T}^{n+1} \mathbb{1}_{A(n+1)} \\
&= \hat{T}^n \mathbb{1}_{A \cap \{\varphi_A = n\}} + \hat{T}^{n+1} \mathbb{1}_{A \cap \{\varphi_A = n+1\}} + \hat{T}^{n+2} \mathbb{1}_{A \cap \{\varphi_A = n+2\}} + \hat{T}^{n+2} \mathbb{1}_{A(n+2)} \\
&\;\;\vdots \\
&= \sum_{k \geq n} \hat{T}^k \mathbb{1}_{A \cap \{\varphi_A = k\}}
\end{aligned}$$

where we used that $\mu(A(n))$ goes to zero for $n \to \infty$, so the remaining term vanishes. But this shows the first assumption:

$$\begin{aligned}
\mu\big(\{\varphi_A = n\} \cap T_A^{-1}B\big) &= \int_{T^{-n}B} \mathbb{1}_{\{\varphi_A = n\}} d\mu = \int_B \mathbb{1}_{\{\varphi_A = n\}} \circ T^n d\mu \\
&= \int_B \hat{T}^n \mathbb{1}_{\{\varphi_A = n\}} d\mu = \sum_{k \geq n} \int_B \hat{T}^k \mathbb{1}_{A \cap \{\varphi_A = k\}} d\mu \\
&= \sum_{k \geq n} \mu\big(A \cap \{\varphi_A = k\} \cap T^{-k}B\big) = \mu\big(A \cap \{\varphi_A \geq n\} \cap T_A^{-1}B\big).
\end{aligned}$$

Now take $t > 0$. Since $\varphi_A$ only takes integer values, we obtain the equality

$$\mu\big(\{\mu(A)\varphi_A \leq t\} \cap T_A^{-1}B\big) = \sum_{n=1}^{\lfloor t/\mu(A) \rfloor} \mu\big(\{\varphi_A = n\} \cap T_A^{-1}B\big)$$
$$= \sum_{n=1}^{\lfloor t/\mu(A) \rfloor} \mu\big(A \cap \{\varphi_A \geq n\} \cap T_A^{-1}B\big) = \int_0^{\lfloor t/\mu(A) \rfloor} \mu(A \cap \{\varphi_A \geq r\})dr.$$

But then we change variables to see

$$\begin{aligned}\int_0^{\lfloor t/\mu(A) \rfloor} \mu\big(A \cap \{\varphi_A \geq r\}\big)dr &\leq \int_0^{t/\mu(A)} \mu\big(A \cap \{\varphi_A \geq r\}\big)dr \\ &= \int_0^t \mu\big(A \cap \{\mu(A)\varphi_A > s\}\big)\frac{1}{\mu(A)}ds \\ &= \int_0^t \mu_A\big(A \cap \{\mu(A)\varphi_A > s\}\big)ds\end{aligned}$$

We get the result since the error we make with the inequality above can be estimated by

$$0 \leq \int_{\lfloor t/\mu(A) \rfloor}^{t/\mu(A)} \mu(A \cap \{\varphi_A \geq r\})dr \leq \int_{\lfloor t/\mu(A) \rfloor}^{t/\mu(A)} \mu(A)dr \leq 1 \cdot \mu(A).$$

□

With this result, we can almost proof theorem 1.2.2. As in [17], we do this in two steps. The following proposition represents the core of the later proof. In fact, again we will state a slightly more general version from [22], which we will use again later.

**Proposition 1.2.5.** *Let $(\tilde{F}_l)_{l\geq 0}$ be a sequence of probability distribution functions on $[0,\infty[$. Define $\tilde{F}_l(\infty) := \lim_{t\to\infty} \tilde{F}_l(t)$ and assume that $\tilde{F}_l(\infty) \to \mathfrak{f} \in [0,1]$ as $l \to \infty$. Then*

*a) If $F$ is some sub-probability function on $[0,\infty[$ (i.e. $\lim_{t\to\infty} F(t) \leq 1$) such that*

$$\int_0^t \big(\tilde{F}_l(\infty) - \tilde{F}_l(s)\big)ds \xrightarrow{l\to\infty} F(t) \tag{1.5}$$

*holds for a dense set in $(0,\infty)$ (e.g. the continuity points of $F$), then $F$ is continuous and $\tilde{F}_l \Rightarrow \tilde{F}$ for a sub-probability distribution $\tilde{F}$ which is uniquely defined by*

$$\int_0^t \big(1 - \tilde{F}(s)\big)ds = F(t) \tag{1.6}$$

*for all $t \geq 0$. Moreover, $\tilde{F}(\infty) = \mathfrak{f}$.*

*b) If conversely $\tilde{F}_l \Rightarrow \tilde{F}$ holds for some sub-probability function $\tilde{F}$ on $[0,\infty[$, then (1.5) is true for the uniquely defined continuous non-decreasing function $F$ satisfying (1.6).*

*Proof.* a) Any subsequence of $(F_l)_{l\geq 1}$ has weak accumulation points by Helly's selection theorem (e.g. [11]). Let $\tilde{F}$ be the limit of the subsequence $(\tilde{F}_{l_j})_{j\geq 1}$. Since any sub-probability function is non-negative and bounded by 1, we get by dominated convergence for every $t \geq 0$

$$\int_0^t \big(\tilde{F}_l(\infty) - \tilde{F}_{l_j}(s)\big)ds \to \int_0^t \big(\mathfrak{f} - \tilde{F}(s)\big)ds.$$

Hence we get by continuity of the integral and monotonicity of $F$

$$\int_0^t \big(\mathfrak{f} - \tilde{F}(s)\big)ds = F(t).$$

In particular, $F$ is continuous. Since $\tilde{F}$ is right-continuous, it is uniquely determined by this equation. Therefore all possible weak accumulation points of $(\tilde{F}_l)_{l\geq 1}$ equal $\tilde{F}$, so we have proven weak convergence of the sequence. Furthermore, the equation above guarantees that $F$ is continuous. Since $F$ is bounded, $\mathfrak{f} - \tilde{F}(s)$ goes to 0 for $s \to \infty$, hence $\tilde{F}(\infty) = \mathfrak{f}$.
b) Again we know by dominated convergence that $\tilde{F}_l \Rightarrow \tilde{F}$ implies for all $t \geq 0$

$$\int_0^t \big(\tilde{F}_l(\infty) - \tilde{F}_l(s)\big)ds \to \int_0^t \big(\mathfrak{f} - \tilde{F}(s)\big)ds =: F(t).$$

The function $F$ is non-decreasing and continuous by it's definition. □

Now the remaining proof of the theorem is quite short. Note that we use the special case of the foregoing lemma where the $\tilde{F}_l$ and $F_l$ are distribution functions, in particular $\tilde{F}_l(\infty) = \mathfrak{f} = 1$.

*Proof of theorem 1.2.2.* We define two sequences of probability functions on $[0,\infty[$ via

$$F_l(t) := \mu\big(\mu(A_l)\varphi_{A_l} \leq t\big) \text{ and } \tilde{F}_l(t) := \mu_{A_l}\big(\mu(A_l)\varphi_{A_l} \leq t\big)$$

for $t \geq 0$. Assume that (1.1) is true and fix a continuity point $t > 0$ of $F$. By assumption $F_l(t)$ converges to $F(t)$. Using lemma 1.2.4 together with $\mu(A_l) \to 0$, we get

$$\int_0^t \big(1 - \tilde{F}_l(s)\big)ds \to F(t).$$

Finally part $a)$ of the foregoing lemma 1.2.5 proves the result. If we assume that (1.2) is true, then the second part of lemma 1.2.5 guarantees that

$$\int_0^t \big(1 - \tilde{F}_l(s)\big)ds \to \int_0^t \big(1 - \tilde{F}(s)\big)ds =: F(t)$$

for $t \geq 0$ is a continuous non-decreasing function. Again we use lemma 1.2.4 to obtain $F_l \Rightarrow F$. Since the $F_l$ are distribution functions, $F$ is at least a sub-probability distribution.

□

## 1.3 Exponential limits

In any ergodic probability preserving system, a variety of distributions can by found if we choose the rare events in the right way. In [12] and [13], the following result is shown.

**Theorem 1.3.1.** *Let $(\mathbb{X}, \mathcal{A}, \mu, T)$ be an aperiodic ergodic probability preserving system. Then*

- *for any sub-probability distribution function $F$ on $[0, \infty[$ which is absolutely continuous, concave and fulfils $F(t) \leq t$ for $t \geq 0$, we can find a sequence of rare events such that the normalised first hitting time distribution converges weakly to it.*

- *for any probability distribution function $F$ on $[0, \infty[$ which fulfils*

$$\int_0^\infty (1 - F(s))ds \leq 1,$$

*we can find a sequence of rare events such that the normalised first return time distribution converges weakly to it.*

We will not prove this result here. Nevertheless, it demonstrates the variety of limiting distributions that is contained in any reasonable chaotic system. For now, we take a closer look under which conditions we see an exponential distribution in the limit. It will turn out that this is a rather common situation.

**Theorem 1.3.2.** *Let $(\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system and $(A_l)_{l \geq 1}$ a sequence of asymptotically rare events. If there is a sequence of non-negative* *integers $(m_l)_{l\geq 0}$ and a collection $\mathcal{U} \subset \mathcal{D}(\mu)$ of probability densities which is relatively compact in $\mathcal{L}_1(\mu)$ such that*

$$m_l \mu(A_l) \longrightarrow 0 \text{ for } l \to \infty \tag{1.7}$$

*and*

$$\mu_{A_l}(\varphi_{A_l} \leq m_l) \longrightarrow 0 \text{ for } l \to \infty \tag{1.8}$$

*and finally*

$$\hat{T}^{m_l}\big(\mu(A_l)^{-1}\mathbb{1}_{A_l}\big) \in \mathcal{U} \text{ for } l \geq 1. \tag{1.9}$$

*If all three assumptions are true, we see an exponential law for the first return time (and thus for the first hitting time), i.e.*

$$\mu(A_l)\varphi_{A_l} \overset{\mathcal{L}(\mu)}{\Longrightarrow} \mathcal{E} \text{ for } l \to \infty.$$

In section 1.4.2 we will see that these three conditions even imply convergence on the level of processes. Since this will be a generalisation of the statement above, we look at the proof later. But first, we try to use this theorem for our class of Folklore maps.

### Exponential law for the Folklore maps

Now we try to apply theorem 1.3.2 to our class of Folklore maps. We follow [19]. We will show that we see an exponential law for certain sets, namely cylinders of rank $l$.

**Definition 1.3.3.** Let $S = (\mathbb{X}, \mathcal{A}, \lambda, T, \xi)$ be a Folklore map. We call sets of the form

$$[Z_0, Z_1, \ldots, Z_{l-1}] := Z_0 \cap T^{-1}Z_1 \cap \cdots \cap T^{-l+1}Z_{l-1}$$

with $Z_i \in \xi$ **cylinders of rank** $l$. For a fixed $x \in \mathbb{X}$, we denote with $\xi_l(x)$ the cylinder of rank $l$ that contains $x$.

As figure 1.3 illustrates, the Lebesgue measure of order $l$ cylinders goes to 0 for $l \to \infty$. In fact, there is some $\delta \in ]0,1[$ given by the Lebesgue measure of the largest cylinder set, such that

$$\lambda([Z_0, Z_1, \ldots, Z_{l-1}]) \leq \delta^l \lambda(\mathbb{X}) = \delta^l. \tag{1.10}$$

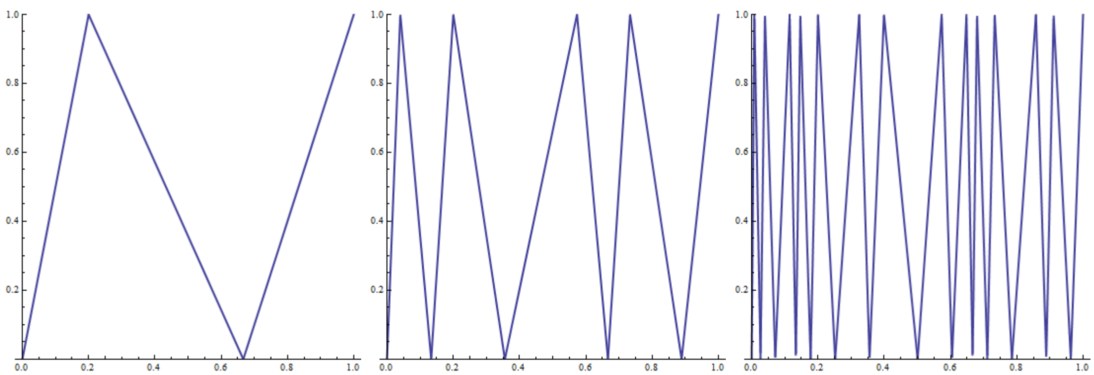

Figure 1.3: Cylinders shrink with increasing rank

We choose a point $x$ in such a way that it is non-periodic and a continuity point of every iteration of $T$ and take a closer look at $A_l := \xi_l(x)$. The choice $m_l = l$ seems to be the natural one since $T^l A_l = \mathbb{X}$ holds. Now condition (1.7) is fulfilled since

$$m_l\mu(A_l) = l\mu(A_l) = l\int \mathbb{1}_{A_l}h\ d\lambda \leq l\lambda(A_l)\sup_{x\in\mathbb{X}} h(x) \leq l\delta^l \sup_{x\in\mathbb{X}} h(x) \to 0.$$

For the third condition, we need another inequality which is connected to (1.10). If we reduce our system to some rank $l$ cylinder $A_l$, then $T^{-j}A_l$ is a rank $j$ cylinder for this new system and we can use (1.10) to show

$$\mu(A_l \cap T^{-j}A_l) \leq \lambda(A_j \cap T^{-j}A_j)\sup_{x\in\mathbb{X}} h(x) \leq \delta^j\lambda(A_j)\sup_{x\in\mathbb{X}} h(x) \leq \delta^j\mu(A_j)\frac{\sup_{x\in\mathbb{X}} h(x)}{\inf_{x\in\mathbb{X}} h(x)}.$$

Now we get (1.8) with a simple continuity argument. Since $x$ is not periodic, we have for all $M \in \mathbb{N}$

$$\max_{1\leq j\leq M} |x, T^j x| =: \varepsilon_M > 0.$$

Now fix $M$ and take $\gamma$ so small such that $T$ and the first $M$ iterations are continuous on $B_\gamma(x)$ and such that

$$|x-y| < \gamma \quad \Rightarrow \quad |T^j x - T^j y| < \frac{\varepsilon_M}{2}$$

holds. Now take $l_0 = l_0(M)$ so large such that we get $A_{l_0} \subseteq B_\gamma(x)$. Thus

$$|x - T^j y| \geq |x - T^j x| - |T^j x - T^j y| \geq \varepsilon_M - \frac{\varepsilon_M}{2} = \frac{\varepsilon_M}{2} > \gamma.$$

Consequently, for $l > l_0$, we get no returns up to time $M$ to $A_l$. But then we can conclude

$$\begin{aligned}
\mu_{A_l}(\varphi_{A_l} \leq l) &= \mu(A_l)^{-1}\mu\Bigg(\bigcup_{j=M+1}^{l} A_l \cap \{\varphi_{A_l} = j\}\Bigg) \\
&\leq \mu(A_l)^{-1}\mu\Bigg(\bigcup_{j=M+1}^{l} A_l \cap T^{-j}A_l\Bigg) \\
&\leq \mu(A_l)^{-1} \sum_{j=M+1}^{l} \mu(A_l \cap T^{-j}A_l) \\
&\leq \mu(A_l)^{-1} \sum_{j=M+1}^{l} \frac{\sup_{x\in\mathbb{X}} h(x)}{\inf_{x\in\mathbb{X}} h(x)}\mu(A_l) \cdot \delta^j \\
&\leq \frac{\sup_{x\in\mathbb{X}} h(x)}{\inf_{x\in\mathbb{X}} h(x)} \cdot \frac{\delta^M}{1-\delta}.
\end{aligned}$$

So for every $\varepsilon > 0$ we can find $M = M(\varepsilon) \in \mathbb{N}$ such that

$$\frac{\sup_{x\in\mathbb{X}} h(x)}{\inf_{x\in\mathbb{X}} h(x)} \cdot \frac{\delta^M}{1-\delta} < \varepsilon$$

holds and for $l > l_0 = l_0\big(M(\varepsilon)\big)$

$$\mu_{A_l}(\varphi_{A_l} \leq l) < \varepsilon$$

is true. To show (1.9), we need two results from [18], by proposition 1.2.4. and lemma 1.6.5., the mapped densities are given by

$$\hat{T}^l\big(\mathbb{1}_{A_l}\mu(A_l)^{-1}\big) = \mu(A_l)^{-1}\Big|\big(T^{-l}|_{A_l}\big)'\Big|$$

and all those densities have bounded regularity with the same universal bound. But this implies that they are all Lipschitz with a common Lipschitz constant, which implies equicontinuity. Furthermore, since we only look at propability densities, a uniform Lipschitz constant together with common integral 1 gives uniform boundedness. So we can apply Arzela-Ascoli to show that this set of uniform Lipschitz-bounded probability densities is precompact. Hence we meet all three requirements and by theorem 1.3.2 we see an exponential law.

## 1.4 Rare event processes

Instead of looking at one visit to $A_l$, we define, inspired by [22] and [20], a process which counts the time between two consecutive visits. To formulate this definition, recall the first return map $T_A x := T^{\varphi_A(x)}$. Then we can consider $\tilde{S} = (A, \mathcal{A} \cap A, \mu_A, T_A)$ as a new system, and it turns out that many properties of the original one stay preserved. In fact, if $S = (\mathbb{X}, \mathcal{A}, \mu, T)$ is probability preserving or ergodic, then $\tilde{S}$ is so as well.

The idea is now to look at the sequence of consecutive return and hitting times. Note that after the first visit, the following hitting and return times coincide since we are already in the set.

**Definition 1.4.1.** Let $S = (\mathbb{X}, \mathcal{A}, \mu, T)$ be a ergodic probability preserving dynamical system and $A \in \mathcal{A}$. Define

$$\Phi_A := (\varphi_A, \varphi_A \circ T_A, \varphi_A \circ T_A^2, \dots),$$

then $\Phi_A : \mathbb{X} \to [0, \infty)^{\mathbb{N}_0}$ on $(\mathbb{X}, \mathcal{A}, \mu)$ is called **hitting time process** and $\Phi_A : A \to [0, \infty)^{\mathbb{N}_0}$ on $(A, \mathcal{A} \cap A, \mu_A)$ is called **return time process**. We write $\Phi_A^{[d]}$ for the $d$-dimensional marginal of the process.

Note that the return time process is stationary, since $\mu_A \circ T_A = \mu_A$ holds. The sequence space

$$\mathfrak{S} = [0, \infty]^{\mathbb{N}_0} = \left\{ s = (s_j)_{j \geq 0} : s_j \in [0, \infty] \right\}$$

is a compact Polish space with the metric

$$d_{\mathfrak{S}}(s, \tilde{s}) = \sum_{j \geq 0} 2^{-j} d_{[0,\infty]}(s_j, \tilde{s}_j)$$

(note that $[0, \infty]$ is homeomorphic to $[0, 1]$, so we can find some metric $d_{[0,\infty]}$ there). As before, we try to find out what happens for $l \to \infty$ under the right normalisation. Since the random elements we consider now are not longer real-valued, we need to redefine what weak convergence means.

**Definition 1.4.2.** A sequence of random variables $(R_n)_{n \geq 1}$ with values on a metric space $\mathfrak{S}$ converges weakly to some random variable $R$n denoted by $R_n \overset{\mu}{\Longrightarrow} R$, if and only if

$$\int \psi \circ R_n \, d\mu \xrightarrow{n \to \infty} \int \psi \circ R \, d\mu$$

for all bounded Lipschitz-continuous functions $\psi : \mathfrak{S} \to \mathbb{R}$.

In [2], example 2.4 shows the following equivalence to weak convergence, which turns out to be practical or various calculations.

**Proposition 1.4.3.** *For a sequence of random variables $(R_l)_{l\geq 0}$ on the space $\mathfrak{S} = [0,\infty]^{\mathbb{N}_0}$, weak convergence to some random variable $R$ with distribution function $F$ is equivalent to convergence of all finite dimensional marginals in the sense of the first section, that is if $F_l$ is the distribution function of $R_l$, then we have*

$$F_l^{[d]}(t_0,\ldots,t_{d-1}) \longrightarrow F^{[d]}(t_0,\ldots,t_{d-1})$$

*for all $d \geq 1$ and all continuity points $(t_0,\ldots,t_{d-1})$ of $F^{[d]}$.*

In section 1.4.2 we will concentrate on the multidimensional generalisation of the exponential laws we have seen in the foregoing part. Our strategy also resembles closely the one before. In section 1.4.1 we try to establish a connection between the possible limits, which we then use later to find a sufficient condition for the exponential limit. Before we do this, we state the following observation from [22], which tells us more about the limit of the return time process.

**Proposition 1.4.4.** *Let $S = (\mathbb{X},\mathcal{A},\mu,T)$ be an ergodic probability preserving dynamical system and $(A_l)_{l\geq 1}$ a sequence of asymptotically rare events. If the return time process converges weakly to some process $\tilde{\Phi} = \left(\tilde{\varphi}^{(j)}\right)_{j\geq 0}$ in $[0,\infty]$, i.e.*

$$\mu(A_l)\tilde{\Phi}_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \tilde{\Phi},$$

*then the limit is again stationary and $\mathrm{E}\left[\tilde{\varphi}^{(0)}\right] \leq 1$.*

*Proof.* Let us start with the expectation. If we have weak convergence for the whole process, we get it especially for the first coordinate. But then by Skorokhod's representation theorem (e.g. [11]), we know that there exists a sequence of random variables $(R_l)_{l\geq 0}$ and some other random variable $R$ all defined on the same probability space such that $R_l \overset{d}{=} \mu(A_l)\tilde{\varphi}_{A_l}$, $R \overset{d}{=} \tilde{\varphi}^{(0)}$ and $R_l \longrightarrow R$ almost surely. Hence we can apply Fatou's lemma to get

$$1 = \lim_{l\to\infty} \mathrm{E}\left[\mu(A_l)\tilde{\varphi}_{A_l}\right] = \lim_{l\to\infty} \mathrm{E}\left[R_l\right] \geq \mathrm{E}\left[\lim_{l\to\infty} R_l\right] = \mathrm{E}[R] = \mathrm{E}\left[\tilde{\varphi}^{(0)}\right].$$

To show stationarity of $\tilde{\Phi}$, we look at it's finite dimensional distribution functions. But $\mu(A)\tilde{\Phi}_A$ is stationary whenever $\mu(A) > 0$, thus it is straightforward for the weak limit as well. $\square$

### 1.4.1 Connection between the two processes

The first step to deal with the limits of hitting time and return time processes is to prove a similar statement to theorem 1.2.2. As in [22], we state

**Theorem 1.4.5.** *Let $(\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system and $(A_l)_{l\geq 1}$ a sequence of asymptotically rare events. Then*

$$\mu(A_l)\Phi_{A_l} \overset{\mu}{\Longrightarrow} \Phi \text{ for some random sequence } \Phi \text{ in } [0,\infty] \tag{1.11}$$

*if and only if*

$$\mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \tilde{\Phi} \text{ for some random sequence } \tilde{\Phi} \text{ in } [0,\infty]. \tag{1.12}$$

*In this case the sub-probability functions $F^{[d]}$ and $\tilde{F}^{[d]}$ of $\Phi^{[d]}$ and $\tilde{\Phi}^{[d]}$ satisfy for any $d \geq 0$ and $t_j \geq 0$*

$$\int_0^{t_0} \big(\tilde{F}^{[d]}(t_1,\dots,t_d) - \tilde{F}^{[d+1]}(s,t_1,\dots,t_d)\big)ds = F^{[d+1]}(t_0,t_1,\dots,t_d) \tag{1.13}$$

*with the convention $\tilde{F}^{[0]} := 1$. Through this equality, the families $\{F^{[d]}\}_{d\geq 1}$ and $\{\tilde{F}^{[d]}\}_{d\geq 1}$ are uniquely determined.*

In the remaining section we will try to prove this theorem. For the beginning, we show the following observation from [22].

**Lemma 1.4.6.** *In an ergodic probability preserving system $S = (\mathbb{X}, \mathcal{A}, \mu, T)$ we can find for any sequence $(A_l)_{l\geq 1}$ of asymptotically rare events a subsequence $(A_{l_j})_{j\geq 1}$ and random elements $\Phi$ and $\tilde{\Phi}$ on $[0,\infty]^{\mathbb{N}}$ such that*

$$\mu(A_{l_j})\Phi_{A_{l_j}} \overset{\mu}{\Longrightarrow} \Phi \quad \text{and} \quad \mu(A_{l_j})\tilde{\Phi}_{A_{l_j}} \overset{\mu_{A_{l_j}}}{\Longrightarrow} \tilde{\Phi} \quad \text{holds for} \quad j \to \infty. \tag{1.14}$$

*Proof.* For a fixed $d$ we can use the $d$-dimensional version of Helly's selection theorem to show that any sequence $m_i \to \infty$ of indices contains a subsequence $m_{i_j} \to \infty$ that converges, i.e. there exist random elements $\Phi_*^{[d]}$ and $\tilde{\Phi}_*^{[d]}$ of $[0,\infty]^{\mathbb{N}}$ such that

$$\mu(A_{l_j})\Phi_{A_{l_j}}^{[d]} \overset{\mu}{\Longrightarrow} \Phi_*^{[d]} \quad \text{and} \quad \mu(A_{l_j})\tilde{\Phi}_{A_{l_j}}^{[d]} \overset{\mu_{A_{l_j}}}{\Longrightarrow} \tilde{\Phi}_*^{[d]} \tag{1.15}$$

hold for $j \to \infty$. We can actually take the same subsequence if we take a converging one for the hitting time and then a further refinement such that also the return time converges. By diagonalisation, we can find a sequence such that (1.15) holds for any $d \geq 1$. But then we found a consistent family of probability measures, hence by Kolmogorov's extension theorem [11], there exists a random element $\Phi$ on $[0,\infty]^{\mathbb{N}}$ such that $\mu(A_{l_j})\Phi_{A_{l_j}} \overset{\mu}{\Longrightarrow} \Phi$ holds and $\Phi$ satisfies $\Phi^{[d]} = \Phi_*^{[d]}$ in law for every $d \geq 1$. The same arguments also hold for $\tilde{\Phi}_*^{[d]}$. $\square$

Now the core result for the later proof is the following observation from [22].

**Lemma 1.4.7.** *Let $S = (\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system and $(A_l)_{l\geq 1}$ a sequence of asymptotically rare events. If*

$$\mu(A_l)\Phi_{A_l} \overset{\mu}{\Longrightarrow} \Phi \quad \text{and} \quad \mu(A_l)\tilde{\Phi}_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \tilde{\Phi}$$

*holds for two random sequences in $[0,\infty]$, then the finite-dimensional distributions $F^{[d]}$ and $\tilde{F}^{[d]}$ satisfy for every $d \geq 0$ and $t_j \geq 0$ (1.13).*

*Proof.* (i) Note that the case $d = 0$ is given by theorem 1.2.2. So let us fix some $d \geq 1$. Since we assume convergence for the whole process, we get

$$\mu(A_l)\Phi_{A_l}^{[d]} \overset{\mu}{\Longrightarrow} \Phi^{[d]} \quad \text{and} \quad \mu(A_l)\tilde{\Phi}_{A_l}^{[d]} \overset{\mu_{A_l}}{\Longrightarrow} \tilde{\Phi}^{[d]} \tag{1.16}$$

automatically. We will show (1.13) on a dense subset $M \subseteq (0,\infty)^d$. Using the projection maps $\pi_j$ we define

$$D_j := \left\{t \in ]0,\infty[ \, : \mu(\pi_j\Phi = t) > 0\right\}$$
$$\tilde{D}_j := \left\{t \in ]0,\infty[ \, : \mu(\pi_{j-1}\tilde{\Phi} = t) > 0\right\}$$

with $\tilde{D}_0 := \emptyset$. Note that both sets are countable. Now set

$$M := \{(t_1,\dots,t_d) : t_j \notin D_j \cup \tilde{D}_j \text{ for } 0 \leq j \leq d\},$$

i.e. we remove countable many hyperplanes. Once we have (1.13) on $[0,\infty[\times M$, we can show that this set is dense in $[0,\infty[^{d+1}$ and that both sides of the equality define right-continuous functions on it. Consequently, the relation is true for all $(t_0,\dots,t_d) \in [0,\infty[^{d+1}$.

(ii) Now we try to show (1.13) on $M$. Note that by definition, we removed any discontinuous point with the definition of this set (the sets $D_j$ and $\tilde{D}_j$ are exactly the sets where mass is concentrated, i.e. a jump in the distribution occurs). In particular, every $(t_0,t_1,\dots,t_d) \in D_0^c \times M$ is a continuity point for $F^{[d+1]}$ and $\tilde{F}^{[d+1]}$.

Additionally $(t_1,\dots,t_d)$ is a continuity point of $\tilde{F}^{[d]}$ since $\tilde{\Phi}$ is stationary. Fix a point $(t_1,\dots,t_d)\in M$. We abbreviate for $t\in\ [0,\infty[$

$$
\begin{aligned}
F_l(t) &:= \mu\big(\mu(A_l)\Phi_{A_l}^{[d+1]} \le (t,t_1,\dots,t_d)\big), \quad F(t) := F_{[d+1]}(t,t_1,\dots,t_d),\\
\tilde{F}_l(t) &:= \mu\big(\mu(A_l)\tilde{\Phi}_{A_l}^{[d+1]} \le (t,t_1,\dots,t_d)\big), \quad \tilde{F}(t) := \tilde{F}_{[d+1]}(t,t_1,\dots,t_d).
\end{aligned}
$$

Note that since $\varphi_{A_l}$ is finite a.e. and $\tilde{\Phi}_{A_l}$ is stationary with respect to $\mu_{A_l}$, we get with dominated convergence

$$
\begin{aligned}
\tilde{F}_l(\infty) :&= \lim_{t\to\infty} \tilde{F}_l(t)\\
&= \lim_{t\to\infty} \mu_{A_l}\Big(\mu(A_l)\big(\varphi_{A_l}, \varphi_{A_l}\circ T_{A_l}^{-1},\dots,\varphi_{A_l}\circ T_{A_l}^{-(d+1)}\big) \le \big(t,t_1,\dots,t_d\big)\Big)\\
&= \mu_{A_l}\Big(\mu(A_l)\tilde{\Phi}_{A_l}^{[d]}\circ T_{A_l}^{-1} \le (t_1,\dots,t_d)\Big)\\
&= \mu_{A_l}\Big(\mu(A_l)\tilde{\Phi}_{A_l}^{[d]} \le (t_1,\dots,t_d)\Big).
\end{aligned}
$$

Since $(t_1,\dots,t_d)$ is a continuity point of $\tilde{F}^{[d]}$, the weak convergence in (1.16) ensures

$$
\lim_{t\to\infty} \tilde{F}_l(t) \xrightarrow{l\to\infty} \tilde{F}^{[d]}(t_1,\dots,t_d) =: \mathfrak{f} \in [0,1].
$$

Moreover if we take $t_0\in D_0^c$, $(t_0,\dots,t_d)$ is a continuity point of $F^{[d+1]}$ and $\tilde{F}^{[d+1]}$, so by assumption we have

$$
F_l(t_0) \longrightarrow F(t_0) \quad \text{and} \quad \tilde{F}_l(t_0) \longrightarrow \tilde{F}(t_0).
$$

Now we define the set

$$
B_l := A_l \cap \big\{\mu(A_l)\Phi_{A_l}^{[d]} \le (t_1,\dots,t_d)\big\} = \big\{\mu(A_l)\tilde{\Phi}_{A_l}^{[d]} \le (t_1,\dots,t_d)\big\}.
$$

With this definition, we can rewrite

$$
\begin{aligned}
&\big\{\mu(A_l)\Phi_{A_l}^{[d+1]} \le (t,t_1,\dots,t_d)\big\}\\
&\quad = \big\{\mu(A_l)\varphi_{A_l} \le t\big\} \cap \big\{\mu(A_l)(\varphi_{A_l}\circ T_{A_l}, \varphi_{A_l}\circ T_{A_l}^2,\dots,\varphi_{A_l}\circ T_{A_l}^{d+1}) \le (t_1,\dots,t_d)\big\}\\
&\quad = \big\{\mu(A_l)\varphi_{A_l} \le t\big\} \cap T_{A_l}^{-1}B_l.
\end{aligned}
$$

Using the identity $A\cap B = B\setminus(A^c\cap B)$, we end up with

$$
\begin{aligned}
\mu_{A_l}\big(\big\{\mu(A_l)\varphi_{A_l} > s\big\}\cap T_{A_l}^{-1}B_l\big) &= \mu_{A_l}\Big(T_{A_l}^{-1}B_l \setminus \big(\big\{\mu(A_l)\varphi_{A_l}\le t\big\}\cap T_{A_l}^{-1}B_l\big\}\big)\Big)\\
&= \mu_{A_l}\big(T_{A_l}^{-1}B_l\big) - \mu_{A_l}\big(\big\{\mu(A_l)\varphi_{A_l}\le t\big\}\cap T_{A_l}^{-1}B_l\big\}\big)\\
&= \mu_{A_l}(B_l) - \mu_{A_l}\big(\mu_{A_l}\tilde{\Phi}_{A_l}^{[d+1]} \le (s,t_1,\dots,t_d)\big)\\
&= \tilde{F}_l(\infty) - \tilde{F}_l(s).
\end{aligned}
$$

Now we use lemma 1.2.4 again to get

$$F_l(t) \leq \int_0^t \big(\tilde{F}_l(\infty) - \tilde{F}_l(s)\big) ds \leq F_l(t) + \mu(A_l)$$

for every $l \geq 1$ and $t > 0$. Since the measure of $A_l$ goes to zero, we get for a continuity point $t_0 \in D_0^c$

$$\int_0^{t_0} \big(\tilde{F}_l(\infty) - \tilde{F}_l(s)\big) ds \longrightarrow F(t_0)$$

as $l$ increases. But this applies for every $t_0$ in the set $D_0^c$ which is dense in $[0, \infty]$, so we can use the first part of proposition 1.2.5 to conclude that the limit distribution $\tilde{F}$ is uniquely defined by the relation

$$\int_0^t \big(\mathfrak{f} - \tilde{F}(s)\big) ds = F(t). \tag{1.17}$$

Hence we proved (1.13) for $[0, \infty) \times M$.
(iii) We still need to show uniqueness. By (1.13), $F^{[d+1]}$ is determined by $\tilde{F}^{d+1}$ and $\tilde{F}^d$. On the other hand (1.17) shows us that $\tilde{F}^{[d+1]}$ is given by $\tilde{F}^{[d]}$ and $F^{[d+1]}$ on a dense set. By right-continuity, this extends to all points. By theorem 1.2.2, $\tilde{F}^{[1]}$ and $F^{[1]}$ determine each other. The claim follows therefore by induction. □

With all this preparation, only a few steps are missing for the main result.

*Proof of theorem 1.4.5.* (i) In the first step, we show that (1.11) and (1.12) are equivalent. By the foregoing lemma, we know together they imply (1.13). So assume (1.12) and let $\tilde{F}^{[d]}$ be the distribution function of $\tilde{\Phi}^{[d]}$. Via (1.13), we find another variable $\Phi$ uniquely determined by the according family of finite distribution function $\{F^{[d]}\}$. By lemma 1.4.6, any sequence has a converging subsequence and by lemma 1.4.7, the limit coincides with $\Phi$. So assume (1.11) fails, i.e. there exists $d \geq 1$ and a continuity point $(t_0, \dots, t_{d-1})$ of $F^{[d]}$ such that

$$\mu\big(\mu(A_l)\Phi_{A_l}^{[d]} \leq (t_0, \dots, t_{d-1})\big) \nrightarrow F^{[d]}(t_0, \dots, t_{d-1}).$$

Therefore there exists $\varepsilon > 0$ and $l_j \to \infty$ such that the Lévy distance

$$\operatorname{dist}\big(F_{l_j}^{[d]}, F^{[d]}\big) \geq \varepsilon$$

for all $j \geq 1$. But as before, by lemma 1.4.6 there is a further subsequence where convergence takes place. This contradiction shows (1.11). The other direction works in the same way. □

### 1.4.2 Poisson asymptotics

Now we will show under which conditions we see the generalisation of theorem 1.3.2, i.e. convergence to some process

$$\Phi_{\mathrm{Exp}} = \big(\varphi_{\mathrm{Exp}}^{(0)}, \varphi_{\mathrm{Exp}}^{(1)}, \dots\big)$$

where the components are i.i.d. variables with exponential distribution. If we have convergence to such an object, we say we have **Poisson asymptotics** for the hitting and return time. Indeed, by theorem 1.21 in [10], it is shown that $\Phi_{\mathrm{Exp}}$ can be identified with the interarrival times of a standard Poisson point process and that the convergence of the return time process is in fact equivalent to the convergence of the corresponding point process.

We start with a generalised version of theorem 1.4.5.

**Proposition 1.4.8.** *Let $\Phi$ is a stationary random sequence in $[0,\infty)$. Then it's finite-dimensional marginals have distribution functions $F^{[d]}$ satisfying*

$$\int_0^{t_0} \big(F^{[d]}(t_1,\dots,t_d) - F^{[d+1]}(s,t_1,\dots,t_d)\big)ds = F^{[d+1]}(t_0,t_1,\dots,t_d), \tag{1.18}$$

*for all $d \geq 0$ and $t_j \geq 0$ if and only if $\Phi = \Phi_{\mathrm{Exp}}$.*

*Proof.* Assume (1.18) and rewrite $\Phi = (\varphi_0, \varphi_1, \dots)$. The case $d = 0$ gives a simple ordinary differential equation with solution with solution $F^{[1]}(t) = 1 - e^{-t}$. Hence we conclude all one-dimensional marginals $\varphi_i$ are exponentially distributed. In the independent case, those marginals would already define the joint distribution. So we try to show that $\varphi_0$ is independent of the other ones. By stationarity, we get that the whole sequence is independent. So fix $d \geq 1$ and $(t_1,\dots,t_d) \in [0,\infty)^d$. We show if $\mu\big((\varphi_1,\dots,\varphi_d) \leq (t_1,\dots,t_d)\big) > 0$ is true, then so is

$$\mu(\varphi_0 \leq t) = \mu\big(\varphi_0 \leq t | (\varphi_1,\dots,\varphi_d) \leq (t_1,\dots,t_d)\big)$$

for every $t \geq 0$. This is a simple consequence of (1.18) and stationarity, since

$$\begin{aligned}
\mu\big(\varphi_0 \leq t | (\varphi_1,\dots,\varphi_d) \leq (t_1,\dots,t_d)\big) &= \frac{F^{[d+1]}(t_0,\dots,t_d)}{F^{[d]}(t_1,\dots,t_d)} \\
&= \frac{\int_0^{t_0} F^{[d]}(t_1,\dots,t_d) - F^{[d+1]}(s,\dots,t_d)ds}{F^{[d]}(t_1,\dots,t_d)} \\
&= \int_0^{t_0} 1 - \mu(\varphi_0 \leq s | (\varphi_1,\dots,\varphi_d) \leq (t_1,\dots,t_d)\big)ds
\end{aligned}$$

holds. Hence the left hand side fulfils the same ODE as above, and therefore

$$\mu\big(\varphi_0 \leq t | (\varphi_1, \ldots, \varphi_d) \leq (t_1, \ldots, t_d)\big) = 1 - e^{-t} = \mu(\varphi_0 \leq t).$$

For the other direction, we need to show that $\Phi_{\mathrm{Exp}}$ fulfils (1.18). But since $\Phi_{\mathrm{Exp}}$ consists of independent random variables, this breaks down to the same equation we have seen in the one-dimensional case. $\square$

Now inspired by the foregoing proposition, we get the following sufficient condition for Poisson asymptotics.

**Theorem 1.4.9.** *Take a probability preserving system $S = (\mathbb{X}, \mathcal{A}, \mu, T)$ and a sequence of asymptotically rare events $(A_l)_{l \geq 1}$. Then we have equivalence of the following statements:*

*a) For the distribution functions $\tilde{F}_l^{[d]}$ and $F_l^{[d]}$ of $\Phi_{A_l}^{[d]}$ we have for each $d \geq 0$ on a dense subset of $[0, \infty[^d$*

$$\tilde{F}_l^{[d]} - F_l^{[d]} \longrightarrow 0 \text{ for } l \to \infty. \tag{1.19}$$

*b) We see Poisson asymptotics for the hitting-time process, that is*

$$\mu(A_l)\Phi_{A_l} \overset{\mu}{\Longrightarrow} \Phi_{\mathrm{Exp}}. \tag{1.20}$$

*c) We see Poisson asymptotics for the return-time process, that is*

$$\mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \Phi_{\mathrm{Exp}}. \tag{1.21}$$

*Proof.* By theorem 1.4.5 and the foregoing proposition the equivalence of (1.20) and (1.21) is clear. Also, it is obvious that the difference between the finite distribution functions tends to zero since. Now assume (1.19). By lemma 1.4.6, it is enough to show that every accumulation point of $\Psi$ coincides with $\Psi_{\mathrm{Exp}}$. So take a converging subsequence $(l_j)_{j \geq 1}$. Since $(A_{l_j})$ is also asymptotically rare, we can use theorem 1.4.5 to see that (1.13) holds for all $d \geq 0$ and $t_j \geq 0$. By definition of convergence of the process' distribution, we also have for every $d$

$$F_{l_j}^{[d]} \Rightarrow F^{[d]} \text{ and } \tilde{F}_{l_j}^{[d]} \Rightarrow \tilde{F}^{[d]}.$$

Since all the functions are right-continuous, (1.19) implies that $F^{[d]} = \tilde{F}^{[d]}$. Hence $\Psi$ and $\tilde{\Psi}$ have the same distribution and we already know that $\tilde{\Phi}$ is stationary. Applying proposition 1.4.8 gives $\Phi = \Phi_{\mathrm{Exp}}$. $\square$

For the next statement, we need the following definition:

**Definition 1.4.10.** Let $(R_l)_{l\geq 1}$ be a sequence of random variables on $[0,\infty]$. We call the sequence **asymptotically invariant in measure** if and only if

$$d_{\mathfrak{S}}(R_l \circ T, R_l) \xrightarrow{\mu} 0 \;\text{ as }\; l \to \infty$$

holds, where $\xrightarrow{\mu}$ denotes convergence in measure.

Following [20], we state the following generalisation of the principle of proposition 1.2.1, a uniform version of the idea of strong distributional convergence.

**Theorem 1.4.11.** *Let $S = (\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system. Assume $(R_l)_{l\geq 1}$ is a sequence of Borel measurable maps from $\mathbb{X}$ into some separable metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ which is asymptotically invariant in measure. Take a collection $\mathcal{U} \subseteq \mathcal{D}(\mu)$ which is relatively compact in $\mathcal{L}_1(\mu)$. If there exists $\nu \ll \mu$ and a random element $R$ of $\mathfrak{M}$ such that*

$$R_l \overset{\nu}{\Longrightarrow} R \text{ as } l \to \infty \tag{1.22}$$

*holds, then we have for every sequence $(u_l) \in \mathcal{U}$*

$$R_l \overset{u_l \odot \mu}{\Longrightarrow} R \text{ as } l \to \infty. \tag{1.23}$$

To prove this theorem, we use the following well-known characterisation of ergodicity. Both the statement and the proof can be found in [18].

**Theorem 1.4.12.** *A measure preserving map $T$ on a probability space $(\mathbb{X}, \mathcal{A}, \mu)$ is ergodic if and only if*

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} \hat{T}^k (u - u^*) \right\|_{\mathcal{L}_1(\mu)} \longrightarrow 0 \quad \forall\, u, u^* \in \mathcal{D}(\mu).$$

Now the only missing thing is some technical lemma.

**Lemma 1.4.13.** *Let $(\mathfrak{S}, d_{\mathfrak{S}})$ be a metric space, $w \in \mathfrak{S}$ constant and $m_n : \mathfrak{S} \to \mathfrak{S}$ a sequence of operators such that $m_n(u) \to w$ for all $u \in \mathfrak{S}$. Assume $(m_n)_{n\geq 1}$ is equicontinuous, i.e.*

$$\forall u \in \mathfrak{S}\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall n \in \mathbb{N}\ \forall u' \in \mathfrak{S} :\ d_{\mathfrak{S}}(u, u') \leq \delta \ \Rightarrow\ d_{\mathfrak{S}}\big(m_n(u), m_n(u')\big) \leq \varepsilon.$$

*Then $m_n \to w$ uniformly on the closure of any $\mathcal{U}$ which is relatively compact in $\mathfrak{S}$.*

*Proof.* Fix $\varepsilon > 0$ and take $\delta > 0$ in such a way that equicontinuity is fulfilled for $\varepsilon/2$. Take $\mathcal{U}$ arbitrary such that it's closure is compact and let $(U_i)_{i=1}^N$ be a finite $\delta$-cover. Denote the midpoint of $U_i$ with $u_i$ and take $n_i \in \mathbb{N}$ such that

$$d_{\mathfrak{S}}\big(m_n(u_i), w\big) \leq \frac{\varepsilon}{2} \quad \forall n \geq n_i.$$

If we define $n' = \max\{n_1, \ldots, n_N\}$, we can take an arbitrary $u \in U_i$ and obtain

$$d_{\mathfrak{S}}\big(m_n(u), w\big) \leq d_{\mathfrak{S}}\big(m_n(u), m_n(u_i)\big) + d_{\mathfrak{S}}\big(m_n(u_i), w\big) \leq \varepsilon.$$

Since the choice of $n'$ is independent of $u$, we are done. □

*Proof of theorem 1.4.11.* (i) W.l.o.g. we assume that $v := d\nu/d\mu$ is in $\mathcal{U}$. Otherwise we add $v$ to $\mathcal{U}$ and the set remains relatively compact. Fix a sequence $(u_n)_{n\geq 1}$ in $\mathcal{U}$. We know by (1.22) that

$$\int \psi \circ R_n \cdot v d\mu \longrightarrow \mathrm{E}[\psi(R)]$$

holds for every bounded Lipschitz function $\psi : \mathfrak{S} \to \mathbb{R}$. Therefore we get the claim as soon as we can show

$$\int \psi \circ R_n \cdot v d\mu - \int \psi \circ R_n \cdot u_n d\mu \xrightarrow{n\to\infty} 0$$

for all those $\psi$.
(ii) Note that since $\mu$ is the invariant probability, by theorem 1.4.12 and ergodicity, the averaged densities converge to 1, i.e.

$$\left\| \frac{1}{M} \sum_{m=0}^{M-1} \hat{T}^m u - 1 \right\|_{\mathcal{L}_1(\mu)} \longrightarrow 0.$$

The operators $M^{-1} \sum_{m=0}^{M-1} : \mathcal{L}_1(\mu) \to \mathcal{L}_1(\mu)$ all have norm 1 and are therefore equicontinuous. By the foregoing lemma, this implies uniform convergence, i.e.

$$\left\| \frac{1}{M} \sum_{m=0}^{M-1} \hat{T}^m u - \frac{1}{M} \sum_{m=0}^{M-1} \hat{T}^m u^* \right\|_{\mathcal{L}_1(\mu)} \xrightarrow{M\to\infty} 0 \quad \text{uniformly in } u, u^* \in \mathcal{U}.$$

Hence we can find some $M \geq 1$ such that for every $n \geq 1$

$$\left\| \frac{1}{M} \sum_{m=0}^{M-1} \hat{T}^m v - \frac{1}{M} \sum_{m=0}^{M-1} \hat{T}^m u_n \right\|_{\mathcal{L}_1(\mu)} < \frac{\varepsilon}{4||\psi||_\infty}$$

and consequently

$$\Big|\int \psi\circ R_n\cdot\Big(\frac{1}{M}\sum_{m=0}^{M-1}\hat{T}^m v\Big)d\mu - \int \psi\circ R_n\cdot\Big(\frac{1}{M}\sum_{m=0}^{M-1}\hat{T}^m u_n\Big)d\mu\Big| \tag{1.24}$$

$$= \Big|\,||\psi||_\infty \int \Big(\frac{1}{M}\sum_{m=0}^{M-1}\hat{T}^m v\Big) - \Big(\frac{1}{M}\sum_{m=0}^{M-1}\hat{T}^m u_n\Big)d\mu\Big| < \frac{\varepsilon}{4}. \tag{1.25}$$

(iii) Since $\psi$ is Lipschitz and $(R_n)_{n\geq 1}$ asymptotically rare in measure, we have

$$\Big|\psi\circ R_n\circ T, \psi\circ R_n\Big| \leq L_\psi d_{\mathfrak{M}}(R_n\circ T, R_n) \xrightarrow{\mu} 0.$$

So not only $(\psi\circ R_n)_{n\geq 1}$ is also asymptotically invariant in measure, the same holds for each sequence $(\Psi_{i,n})_{n\geq 1}$ with $\Psi_{i,n} = \psi\circ R_n\circ T^i$ since $\mu$ is $T$-invariant. Since $\mathcal{U}$ is precompact in $\mathcal{L}_1(\mu)$, it is especially uniformly integrable. Therefore we can find some $\delta > 0$ such that

$$\int_A u\ d\mu < \frac{\varepsilon}{16M||\psi||_\infty}$$

holds for all $u\in\mathcal{U}$ and $A\in\mathcal{A}$ with $\mu(A)<\delta$. Denote $\eta = (16m)^{-1}\varepsilon > 0$. Hence asymptotical invariance we find some $n_i\geq 1$ such that

$$\mu\Big(|\Psi_{i,n} - \Psi_{i,n\circ T}| > \eta\Big) < \delta$$

holds for $n\geq n_i$. Define $n' = \max\{n_0,\dots,n_{M-1}\}$. Then, using the definition $A := \{|\Psi_{i,n} - \Psi_{i,n}\circ T| > \eta\}$, we see

$$\begin{aligned}
&\int |\Psi_{i,n} - \Psi_{i,n}\circ T|\cdot u_n d\mu \\
&= \int_A |\Psi_{i,n} - \Psi_{i,n}\circ T|\cdot u_n d\mu + \int_{A^c} |\Psi_{i,n} - \Psi_{i,n}\circ T|\cdot u_n d\mu \\
&\leq 2||\psi||_\infty \int_A u_n d\mu + \eta\int_{A^c} u_n d\mu \\
&\leq 2||\psi||_\infty \frac{\varepsilon}{16M||\psi||_\infty} + \eta = \frac{3\varepsilon}{16M} \leq \frac{\varepsilon}{4M}
\end{aligned}$$

(iv) We can use this inequality to gain

$$\begin{aligned}
&\left|\int \psi \circ R_n \cdot u_n d\mu - \int \psi \circ R_n \cdot \left(\frac{1}{M}\sum_{m=0}^{M-1} \hat{T}^m u_n\right) d\mu\right| \\
&\leq \frac{1}{M}\sum_{k=0}^{M-1}\left|\int \psi \circ R_n \cdot u_n - \psi \circ R_n \cdot \hat{T}^k u_n d\mu\right| \\
&= \frac{1}{M}\sum_{k=0}^{M-1}\left|\int \left(\psi \circ R_n - \psi \circ R_n \circ T^k\right) \cdot u_n d\mu\right| \\
&= \frac{1}{M}\sum_{k=0}^{M-1}\left|\int \sum_{i=0}^{k} \left(\psi \circ R_n \circ T^{i-1} - \psi \circ R_n \circ T^i\right) \cdot u_n d\mu\right| \\
&\leq \frac{1}{M}\sum_{k=0}^{M-1}\sum_{i=0}^{k}\int \left|\psi \circ R_n \circ T^{i-1} - \psi \circ R_n \circ T^i\right| \cdot u_n d\mu \\
&= \frac{1}{M}\sum_{k=0}^{M-1}\sum_{i=0}^{k-1}\int \left|\Psi_{i,n} - \Psi_{i,n} \circ T\right| \cdot u_n d\mu \\
&\leq \frac{1}{M}\sum_{k=0}^{M-1}\sum_{i=0}^{k-1}\frac{\varepsilon}{4M} \leq \frac{\varepsilon}{4}
\end{aligned}$$

for all $n \geq n'$. Since we can repeat the whole procedure for $v$, we can obtain

$$\left|\int \psi \circ R_n \cdot v d\mu - \int \psi \circ R_n \cdot \left(\frac{1}{M}\sum_{m=0}^{M-1} \hat{T}^m v\right) d\mu\right| \leq \frac{\varepsilon}{4}$$

for all $n$ greater some other number $n''$. But now we can combine all we have shown to

$$\begin{aligned}
&\left| \int \psi \circ R_n \cdot v \; d\mu - \int \psi \circ R_n \cdot u_n \; d\mu \right| \\
&\leq \left| \int \psi \circ R_n \cdot u_n \; d\mu - \int \psi \circ R_n \cdot \left( \frac{1}{M} \sum_{m=0}^{M-1} \hat{T}^m u_n \right) \; d\mu \right| \\
&+ \left| \int \psi \circ R_n \cdot \left( \sum_{m=0}^{M-1} \hat{T}^m v \right) \; d\mu - \int \psi \circ R_n \cdot \left( \frac{1}{M} \sum_{m=0}^{M-1} \hat{T}^m u_n \right) \; d\mu \right| \\
&+ \left| \int \psi \circ R_n \cdot v \; d\mu - \int \psi \circ R_n \cdot \left( \frac{1}{M} \sum_{m=0}^{M-1} \hat{T}^m v \right) \; d\mu \right| \\
&\leq \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon
\end{aligned}$$

for all $n \geq \max\{n', n''\}$ and $M$ big enough such that (1.24) is fulfilled. □

Now as nice as theorem 1.4.9 is, it does not provide an easy checkable criterion to show Poisson asymptotics. In the last theorem of this chapter we try to find useful sufficient conditions which do not depend on anything but what happens to $A_l$ until some time which is $o(1/\mu(A_l)$. So we close this chapter with the statement from [20].

**Theorem 1.4.14.** *Let $(\mathbb{X}, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system and $(A_l)_{l \geq 1}$ a sequence of asymptotically rare events. Suppose there is a sequence of non-negative integers $(m_l)_{l \geq 0}$ and a collection $\mathcal{U} \subset \mathcal{D}(\mu)$ of probability densities which is relatively compact in $\mathcal{L}_1(\mu)$ such that*

$$m_l \mu(A_l) \longrightarrow 0 \textit{ for } l \to \infty \tag{1.26}$$

*and*

$$\mu_{A_l}(\varphi_{A_l} \leq m_l) \longrightarrow 0 \textit{ for } l \to \infty \tag{1.27}$$

*and finally*

$$\hat{T}^{m_l}\big(\mu(A_l)^{-1} \mathbb{1}_{A_l}\big) \in \mathcal{U} \textit{ for } l \geq 1. \tag{1.28}$$

*If all three assumptions are true, we see an exponential law for the first hitting time (and thus for the first return time), i.e.*

$$\mu(A_l)\Phi_{A_l} \overset{\mathcal{L}(\mu)}{\Longrightarrow} \Phi_{\mathrm{Exp}} \textit{ for } l \to \infty. \tag{1.29}$$

*Proof.* (i) Let $F_l$ be the distribution function of $\mu(A_l)\Phi_{A_l}$ and $F_{\mathrm{Exp}}$ the one of $\Phi_{\mathrm{Exp}}$. We prove the theorem by contradiction. So assume (1.29) fails, i.e. there is some metric $\mathfrak{d}$ on $\mathfrak{S}$ (e.g. the Prokhorov metric) which characterises weak convergence and some $\varepsilon > 0$ such that some subsequence $(F_{l_j})_{j\geq 1}$ fails to converge, i.e.

$$\mathfrak{d}(F_{l_j}, F_{\mathrm{Exp}}) > \varepsilon \quad \forall\, j \geq 1.$$

But due to lemma 1.4.6 there exists a further subsequence with indices $k_j := l_{i_j}$ and random elements $\Phi$ and $\tilde{\Phi}$ such that

$$\mu(A_{k_j})\Phi_{A_{k_j}} \overset{\mu}{\Longrightarrow} \Phi \quad \text{ and } \quad \mu(A_{k_j})\tilde{\Phi}_{A_{k_j}} \overset{\mu_{A_l}}{\Longrightarrow} \tilde{\Phi}$$

holds. In the second step, we will show that for the finite-dimensional distribution function $F_{k_j}^{[d]}$ and $\tilde{F}_{k_j}^{[d]}$ of $\Phi$ and $\tilde{\Phi}$ the assumptions (1.26)-(1.28) already imply for every $d \geq 1$

$$\tilde{F}_{k_j}^{[d]} - F_{k_j}^{[d]} \longrightarrow 0$$

on a dense set. Hence we can use theorem 1.4.9 for the subsequence to show that we have in fact convergence to $\Phi_{\mathrm{Exp}}$, which is a contradiction.
(ii) Fix $d \geq 1$. Note that on $\{\varphi_{A_l} > m_l\}$ we can rewrite for every $l \geq 1$

$$\Phi_{A_l}^{[d]} = \Phi_{A_l}^{[d]} \circ T^{m_l} + (m_l, 0, \ldots, 0).$$

Now multiplying this with $\mu(A_l)$ and using (1.26) we get

$$\mu(A_l)\Phi_{A_l}^{[d]} - \mu(A_l)\Phi_{A_l}^{[d]} \circ T^{m_l} = \big(\mu(A_l)m_l, 0\ldots, 0\big) \to \big(0, \ldots, 0\big)$$

where the convergence is even in probability, with respect to $\mu_{A_l}$. Consequently we know that

$$\mu(A_{k_j})\Phi_{A_{k_j}} \overset{\mu_{A_{k_j}}}{\Longrightarrow} \tilde{\Phi} \quad \text{ implies } \quad \mu(A_{k_j})\Phi_{A_{k_j}} \circ T^{m_{k_j}} \overset{\mu_{A_{k_j}}}{\Longrightarrow} \tilde{\Phi}.$$

If we define

$$u_l = \frac{d(\mu_{A_l} \circ T^{-m_l})}{d\mu} = \hat{T}^{m_l}\big(\mu(A_l)^{-1}\mathbb{1}_{A_l}\big)$$

we can rewrite this as

$$\mu(A_{k_j})\Phi_{A_{k_j}} \overset{u_{k_j}\odot\mu}{\Longrightarrow} \tilde{\Phi}.$$

On the other hand, $\mu(A_{k_j})\Phi_{A_{k_j}} \overset{\mu}{\Longrightarrow} \Phi$ implies

$$\mu(A_{k_j})\Phi_{A_{k_j}} \overset{u_{k_j}\odot\mu}{\Longrightarrow} \Phi$$

by theorem 1.4.11 because $u_l \in \mathcal{U}$ for all $l$ by assumption (1.28). Consequently, $\tilde{\Phi}$ and $\Phi$ have the same distribution and therefore their finite dimensional distribution functions fulfil

$$\tilde{F}^{[d]}_{k_j} - F^{[d]}_{k_j} \longrightarrow 0$$

on a dense set (the continuity points of $F$). By theorem 1.4.9 this contradicts the assumption that we do not have convergence to an $\Phi_{\text{Exp}}$. $\square$

Note that this theorem immediately implies the one-dimensional version, theorem 1.3.2. In the next chapter we will show an alternative way to find exponential laws for rare events. In contrast to the theorem above, we will use conditions which depend on the mixing properties of the system.

# Chapter 2

# An approach via extremal behaviour

This chapter follows the idea of some papers of Jorge Milhazes Freitas, namely [3,4,7,5,8]. We strengthen our assumption on $\mu$ in such a way that we now want it to be equivalent to the reference measure $\lambda$. We start with showing what it means that the maximum of some sequence has an extreme value law and specify the right conditions that such a law exists in the presence and absence of periodicity. It is quite simple to connect hitting times and extreme values by identifying the small sets from chapter 1 with topological $\varepsilon$-balls around some point $x^*$. If we consider a suitable function of the distance to $x^*$ in each time step, not visiting the ball equals the maximum of the sequence being smaller than some bound depending on $\varepsilon$. In section 2.2 we will examine that the occurrence of hitting time statistics and so called extreme value laws is equivalent. Finally, we define a point process that counts how often a sequence exceeds some level $u_n$. Also in this situation we will find a certain limit behaviour.

As already mentioned in the introduction, we look for an observable $f : \mathbb{X} \to \bar{\mathbb{R}}$ such that $X_n := f \circ T^n$ corresponds to the (negative) distance of some point $x^*$, i.e. it possesses a global maximum at $x^*$. We choose

$$X_n(x) = (f \circ T^n)(x) := -\mu\big(B_{|T^n x - x^*|}(x^*)\big)$$

since it fulfils a number of worthy properties, for example

EB1 If $u$ is close to the maximum, the event $U_u = \{X_0 > u\}$ describes a ball around $x^*$ and the radius of the ball depends on $u$ continuously.

EB2 Let $x^*$ be a repelling periodic point of minimal period $p$. Then for big enough

$u$

$$\{X_0 > u\} \cap \{X_j > u\} \begin{cases} = \emptyset \text{ for } 1 \leq j < p \\ \neq \emptyset \text{ for } j = p \end{cases}$$

holds. Additionally, since $x^*$ is repelling, the pre-image of the ball is again a ball of smaller radius. We want to specify the quantity of the growth with some parameter $\theta \in \,]0,1[$ such that

$$\mu\left(\bigcap_{j=0}^{i} T^{-jp}(X_0 > u)\right) = \mu\left(\bigcap_{j=0}^{i}(X_{jp} > u)\right) \sim (1-\theta)^i \mu(X_0 > u) \tag{2.1}$$

holds for large $u$.

## 2.1 Extreme value laws

We start by proving the existence of a limit distribution for the random variable measuring the maximum of a stationary sequence. We follow 3.1-3.4 of [3]. As in [4] we define

$$M_{i,j} = \max\left\{X_i, X_{i+1}, \ldots, X_{i+j-1}\right\}$$

and $M_n := M_{0,n-1}$. Similar as in the first chapter, we look for normalising sequences $a_n$ and $b_n$ such that

$$\mu\big(a_n(M_n - b_n) \leq y\big) \xrightarrow{n\to\infty} H(y) \tag{2.2}$$

holds for some non-trivial distribution function $H$. We say in this situation, the sequence $(X_n)_{n\geq 1}$ has the **extreme value law** (abbreviated EVL) $H$. Note that since $\mu$ is a probability measure, our choice of $f$ implies that the range of $X_n$ (and therefore of $M_n$) is $[-1,0]$. For our choice of $(X_n)_{n\geq 0}$, it turns out that the right normalisation is given by $a_n = n$ and $b_n = 0$, which is motivated from the i.i.d. case. The next proposition makes this a little bit clearer. To shorten notation we will write $u_n = y/n$.

**Proposition 2.1.1.** *For the sequence $u_n = y/n$ we get for all $y \leq 0$*

$$n\mu(X_0 > u_n) = -y. \tag{2.3}$$

*For $y \geq 0$, the left hand side expression is 0.*

*Proof.* Defining $l(\gamma) = \inf\{\eta > 0 : \mu(B_\eta(x^*)) = \gamma\}$, we get

$$\begin{aligned} n\mu(X_0 > u_n) &= n\mu\Big( - \mu(B_{|x-x^*|}(x^*)) > u_n\Big) = n\mu\big(\mu(B_{|x-x^*|}(x^*)) < -u_n\big) \\ &= n\mu\big(\mu(B_{|x-x^*|}(x^*)) < \mu(B_{l(-u_n)}(x^*))\big) \\ &= n\mu\big(|x-x^*| < l(-u_n)\big) = n\mu\Big(B_{l(-u_n)}(x^*)\Big) = -nu_n = -y. \end{aligned}$$

For $y \geq 0$, the measure is zero since the range of $X_0$ is given by $[-1, 0]$. □

Now (2.3) would be very useful if we were in the classical i.i.d. case, as shown in [3]. If $F$ is the distribution function of $X_n$, then we would get by independence

$$\mu(M_n \leq u_n) = \mu\left(\bigcap_{j=0}^{n-1} X_j \leq u_n\right) = F(u_n)^n = (1 - \mu(X_0 > u_n))^n = \left(1 - \frac{-y}{n}\right)^n \to e^y$$

for all $y \leq 0$. Clearly the i.i.d. setting does not apply in an general dynamical system. To get limit law nevertheless, we need some extra conditions. We follow the notation of [4] and [15].

**Definition 2.1.2.** A sequence $(X_n)_{n\geq 0}$ fulfils condition $D_2(u_n)$ if for all $l$, $t$ and $n$ we have

$$|\mu(\{X_0 > u_n\} \cap \{M_l \circ T^t \leq u_n\}) - \mu(X_0 > u_n)\mu(M_l \leq u_n)| \leq \gamma(n, t),$$

where $\gamma(n,t)$ is decreasing in $t$ for each $n$ and

$$n\gamma(n, t_n) \to 0 \text{ for } n \to \infty$$

holds for some sequence $(t_n)_{n\geq 0} \in o(n)$.

Intuitively $D_2(u_n)$ ensures that the further we look into the future, the more we gain independence from our first observation $X_0$. A second condition, stated in [3], is given by

**Definition 2.1.3.** Assume a sequence $(X_n)_{n\geq 0}$ fulfils condition $D_2(u_n)$. Then it fulfils in addition $D'(u_n)$ if there exists a diverging sequence $(k_n)_{n\geq 1}$ of integers such that

$$\lim_{n\to\infty} n \sum_{j=1}^{\lfloor n/k_n \rfloor} \mu(\{X_0 > u_n\} \cap \{X_j > u_n\}) = 0$$

is true and

$$(k_n t_n)_{n\geq 0} \in o(n)$$

holds for the sequence $(t_n)_{n\geq 1}$ from condition $D_2(u_n)$.

Regardless of its technical flavour condition $D'(u_n)$ can be seen as an easy way to impede multiple exceedances in a 'short' period of time. This behaviour, called clustering, occurs for periodic points and will be reconsidered in section 2.1.2. For the moment, as $u_n$ increases, we want isolated exceedances. To prove the existence of an EVL we first need two technical but helpful lemmas. To shorten the notation, we abbreviate $U_u := \{X_0 > u\}$ and $\mathcal{M}_i^j := \{M_{i,j} \leq u\}$.

**Lemma 2.1.4.** *For $s, t \in \mathbb{N}_0$ and $m \in \mathbb{N}$ we have*

$$|\mu(\mathcal{M}_0^{s+t+m}) - \mu(\mathcal{M}_0^s \cap \mathcal{M}_{s+t}^m)| \leq t\mu(U_u).$$

*Proof.* First notice that

$$\begin{aligned}\mathcal{M}_0^s \cap \mathcal{M}_{s+t}^m &= (\mathcal{M}_0^s \cap \mathcal{M}_s^t \cap \mathcal{M}_{s+t}^m) \uplus \big(\mathcal{M}_0^s \cap \mathcal{M}_{s+t}^m \cap (\mathcal{M}_s^t)^c\big) \\ &= \mathcal{M}_0^{s+t+m} \uplus \big(\mathcal{M}_0^s \cap \mathcal{M}_{s+t}^m \cap (\mathcal{M}_s^t)^c\big)\end{aligned}$$

holds and by stationarity of $(X_n)_{n\geq 0}$ we also get $\mu(X_n \leq u) = \mu(X_0 \leq u)$. So a short calculation gives

$$\begin{aligned}&\mu(\mathcal{M}_0^s \cap \mathcal{M}_{s+t}^m) - \mu(\mathcal{M}_0^{s+t+m}) = \mu\big(\mathcal{M}_0^s \cap \mathcal{M}_{s+t}^m \cap (\mathcal{M}_s^t)^c\big) \leq \mu\big((\mathcal{M}_s^t)^c\big) \\ &= \mu\big((\mathcal{M}_0^t)^c\big) = \mu\left(\bigcup_{j=0}^{t-1}\{X_j > u\}\right) \leq \sum_{j=0}^{t-1}\mu(\{X_j > u\}) = t\mu(U_u),\end{aligned}$$

where we used stationarity of $M_n$ in the second equality. □

We will use this result later also for $m = 0$. In this case, we define $\mathcal{M}_j^0 = \mathbb{X}$ for every $j$. The proof then starts with

$$\begin{aligned}&\mu(\mathcal{M}_0^s) - \mu(\mathcal{M}_0^{s+t}) = \mu(\mathcal{M}_0^s \setminus \mathcal{M}_0^{s+t}) = \mu\big(\mathcal{M}_0^s \cap (\mathcal{M}_0^{s+t})^c\big) \\ &= \mu(\{M_s \leq u\} \cap \{M_{s+t} > u\}) = \mu(\{M_s \leq u\} \cap \{M_{s,t} > u\}) \\ &= \mu\big(\mathcal{M}_0^s \cap (\mathcal{M}_s^t)^c\big)\end{aligned}$$

and continues analogously.

**Lemma 2.1.5.** *For any non-negative integers $s$, $t$ and $m$ we have*

$$\begin{aligned}&|\mu(\mathcal{M}_0^s \cap \mathcal{M}_{s+t}^m) - \mu(\mathcal{M}_0^m)\big(1 - s\mu(U_u)\big)| \\ &\leq \left|s\mu(U_u)\mu(\mathcal{M}_0^m) - \sum_{j=0}^{s-1}\mu(U_u \cap \mathcal{M}_{s+t-j}^m)\right| + 2s\sum_{j=1}^{s-1}\mu(U_u \cap T^{-j}U_u).\end{aligned}$$

*Proof.* First we use the triangle inequality to show

$$\begin{aligned}
&|\mu(\mathcal{M}_0^s \cap \mathcal{M}_{s+t}^m) - \mu(\mathcal{M}_0^m)\big(1 - s\mu(U_u)\big)| \\
&\leq \left| s\mu(U_u)\mu(\mathcal{M}_0^m) - \sum_{j=0}^{s-1} \mu(U_u \cap \mathcal{M}_{s+t-j}^m) \right| \\
&+ \left| \mu(\mathcal{M}_0^s \cap \mathcal{M}_{s+t}^m) - \mu(\mathcal{M}_0^m) + \sum_{j=0}^{s-1} \mu(U_u \cap \mathcal{M}_{s+t-j}^m) \right|
\end{aligned}$$

For the second term, note that

$$\mu(\mathcal{M}_0^m) = \mu(\mathcal{M}_{s+t}^m) = \mu(\mathcal{M}_0^s \cap \mathcal{M}_{s+t}^m) + \mu\big((\mathcal{M}_0^s)^c \cap \mathcal{M}_{s+t}^m\big)$$

holds by stationarity. Therefore we obtain

$$\begin{aligned}
&\left| \mu(\mathcal{M}_0^s \cap \mathcal{M}_{s+t}^m) - \mu(\mathcal{M}_0^m) + \sum_{j=0}^{s-1} \mu(U_u \cap \mathcal{M}_{s+t-j}^m) \right| \\
&= \left| \sum_{j=0}^{s-1} \mu(U_u \cap \mathcal{M}_{s+t-j}^m) - \mu\big((\mathcal{M}_0^s)^c \cap \mathcal{M}_{s+t}^m\big) \right| \\
&= \left| \sum_{j=0}^{s-1} \mu(U_u \cap \mathcal{M}_{s+t-j}^m) - \mu\left( \bigcup_{i=0}^{s-1} T^{-i}U_u \cap \mathcal{M}_{s+t}^m \right) \right| \\
&= \left| \underbrace{\sum_{j=0}^{s-1} \mu(T^{-j}U_u \cap \mathcal{M}_{s+t}^m) - \mu\left( \bigcup_{i=0}^{s-1} T^{-i}U_u \cap \mathcal{M}_{s+t}^m \right)}_{\geq 0} \right| \\
&\leq \sum_{j=0}^{s-1} \sum_{i=0}^{j-1} \mu\big(T^{-j}U_u \cap T^{-i}U_u\big) \leq 2s \sum_{j=1}^{s-1} \mu\big(U_u \cap T^{-j}U_u\big).
\end{aligned}$$

□

The next proposition is the actual core result in this section. The existence of a EVL will be a direct consequence. It is inspired by [3] but differs slightly.

**Proposition 2.1.6.** *Let $n, t \in \mathbb{N}$ be fixed and choose $l, k \in \mathbb{N}$ such that $l = \lfloor n/k \rfloor$ and $l\mu(U_u) < 2$. Then we have*

$$\begin{aligned}\left|\mu(\mathcal{M}_0^n) - \big(1 - l\mu(U_u)\big)^k\right| &\leq 2kt\mu(U_u) + 2n\sum_{j=1}^{l-1}\mu(U_u \cap T^{-j}U_u)\\ &+ k\max_{1\leq i\leq k}\left|l\mu(U_u)\mu\Big(\mathcal{M}_0^{(i-1)(l+t)}\Big) - \sum_{j=0}^{l-1}\mu\Big(U_u \cap \mathcal{M}_{l+t-j}^{(i-1)(l+t)}\Big)\right|.\end{aligned}$$

*Proof.* In the remark after lemma 2.1.4 we established

$$\left|\mu(\mathcal{M}_0^n) - \mu\Big(\mathcal{M}_0^{k(l+t)}\Big)\right| \leq kt\mu\big(U(u_n)\big).$$

Now our two lemmas from before show

$$\begin{aligned}&\left|\mu\Big(\mathcal{M}_0^{k(l+t)}\Big) - \Big(1 - l\mu(U_u)\Big)\mu\Big(\mathcal{M}_0^{(i-1)(l+t)}\Big)\right|\\ &\leq \left|\mu\Big(\mathcal{M}_0^{i(l+t)}\Big) - \mu\Big(\mathcal{M}_0^l \cap \mathcal{M}_{l+t}^{(i-1)(l+t)}\Big)\right|\\ &\quad + \left|\mu\Big(\mathcal{M}_0^l \cap \mathcal{M}_{l+t}^{(i-1)(l+t)}\Big) - \Big(1 - l\mu(U_u)\Big)\mu\Big(\mathcal{M}_0^{(i-1)(l+t)}\Big)\right|\\ &\leq t\mu(U_u) + \left|l\mu(U_u)\mu\Big(\mathcal{M}_0^{(i-1)(l+t)}\Big) - \sum_{j=0}^{l-1}\mu\Big(U_u \cap \mathcal{M}_{l+t-j}^{(i-1)(l+t)}\Big)\right|\\ &\quad + 2l\sum_{j=1}^{l-1}\mu(U_u \cap T^{-j}U_n)\\ &\leq t\mu(U_u) + \max_{1\leq i\leq k}\left|l\mu(U_u)\mu\Big(\mathcal{M}_0^{(i-1)(l+t)}\Big) - \sum_{j=0}^{l-1}\mu\Big(U_u \cap \mathcal{M}_{l+t-j}^{(i-1)(l+t)}\Big)\right|\\ &\quad + 2l\sum_{j=1}^{l-1}\mu(U_u \cap T^{-j}U_n) =: \Upsilon.\end{aligned}$$

Note that in the inequality above we additionally get

$$\left|\mu\big(\mathcal{M}_0^{l+t}\big) - \big(1 - l\mu(U_u)\big)\right| \leq \Upsilon$$

if we choose $i = 1$. Since $l\mu(U_n)$ is smaller 2 by assumption, clearly $|1 - l\mu(U_n)| < 1$

holds true. Now we can recursively estimate

$$\begin{aligned}
&\left|\mu\Big(\mathcal{M}_0^{k(l+t)}\Big) - \big(1 - l\mu(U_u)\big)^k\right| \\
&= \left|\sum_{i=0}^{k-1}\left(\big(1 - l\mu(U_u)\big)^{k-1-i}\mu\Big(\mathcal{M}_0^{i(l+t)}\Big) - \big(1 - l\mu(U_u)\big)^{k-i}\mu\Big(\mathcal{M}_0^{(i-1)(l+t)}\Big)\right)\right| \\
&\le \sum_{i=0}^{k-1}\left|\big(1 - l\mu(U_u)\big)^{k-1-i}\mu\Big(\mathcal{M}_0^{i(l+t)}\Big) - \big(1 - l\mu(U_u)\big)^{k-i}\mu\Big(\mathcal{M}_0^{(i-1)(l+t)}\Big)\right| \\
&= \sum_{i=0}^{k-1}\Big|1 - l\mu(U_u)\Big|^{k-1-i}\left|\mu\Big(\mathcal{M}_0^{i(l+t)}\Big) - \big(1 - l\mu(U_u)\big)\mu\Big(\mathcal{M}_0^{(i-1)(l+t)}\Big)\right| \\
&\le \sum_{i=0}^{k-1}\Big|1 - l\mu(U_u)\Big|^{k-1-i}\Upsilon \le k\Upsilon.
\end{aligned}$$

Finally we can conclude

$$\begin{aligned}
\big|\mu(\mathcal{M}_0^n) &- \big(1 - l\mu(U_u)\big)^k\big| \\
&\le \big|\mu(\mathcal{M}_0^n) - \mu\big(\mathcal{M}_0^{k(l+t)}\big)\big| + \big|\mu\big(\mathcal{M}_0^{k(l+t)}\big) - \big(1 - l\mu(U_u)\big)^k\big| \\
&\le kt\mu(U_u) + k\Upsilon,
\end{aligned}$$

which is exactly what we wanted to show. □

### 2.1.1 EVL for the non-clustering case

With the proposition above, it is rather easy to show the existence of an extreme value law.

**Theorem 2.1.7.** *If $(X_n)_{n\ge1}$ fulfils the conditions $D'(u_n)$ and $D_2(u_n)$, then we see an extreme value law with standard exponential distribution function, i.e.*

$$\lim_{n\to\infty}\mu(M_n \le u_n) = \lim_{n\to\infty}\mu(nM_n \le y) = \begin{cases} e^y \text{ for } y \le 0 \\ 1 \text{ for } y > 0. \end{cases}$$

As in [3] we use the foregoing proposition for the proof.

*Proof.* Use 2.1.6 with $l = \lfloor n/k_n \rfloor$, $k = k_n$ and $t = t_n$ as in the conditions $D'(u_n)$ and

$D_2(u_n)$. Then we get

$$\begin{aligned}
&\left|\mu(M_n \leq u_n) - \left(1 - \left\lfloor \frac{n}{k_n} \right\rfloor \mu(X_0 > u_n)\right)^{k_n}\right| \\
&\leq 2k_n t_n \mu(X_0 > u_n) + 2n \sum_{j=1}^{\lfloor n/k_n \rfloor - 1} \mu(X_0 > u_n \cap X_j > u_n) \\
&+ k_n \max_{1 \leq i \leq k_n} \left| l\mu(X_0 > u_n)\mu(M_{(i-1)(l+t_n)} \leq u_n) - \sum_{j=0}^{l-1} \mu\Big(X_0 > u_n \cap \mathcal{M}_{l+t_n-j}^{(i-1)(l+t_n)}\Big)\right|.
\end{aligned}$$

By $D'(u_n)$ the first two terms vanish. To prove that this holds also for the last term, we calculate

$$\begin{aligned}
&k_n \max_{1 \leq i \leq k_n} \left| l\mu(X_0 > u_n)\mu(M_{(i-1)(l+t_n)} \leq u_n) - \sum_{j=0}^{l-1} \mu\Big(X_0 > u_n \cap \mathcal{M}_{l+t_n-j}^{(i-1)(l+t_n)}\Big)\right| \\
&= k_n \max_{1 \leq i \leq k_n} \left| \sum_{j=0}^{l-1} \left(\mu(X_0 > u_n)\mu(M_{(i-1)(l+t_n)} \leq u_n) - \mu\Big(X_0 > u_n \cap \mathcal{M}_{l+t_n-j}^{(i-1)(l+t_n)}\Big)\right)\right| \\
&\leq k_n \max_{1 \leq i \leq k_n} \sum_{j=0}^{l-1} \left| \mu(X_0 > u_n)\mu(M_{(i-1)(l+t_n)} \leq u_n) - \mu\Big(X_0 > u_n \cap \mathcal{M}_{l+t_n-j}^{(i-1)(l+t_n)}\Big)\right| \\
&\leq k_n \max_{1 \leq i \leq k_n} l\gamma(n, t_n) = k_n \left\lfloor \frac{n}{k_n} \right\rfloor \gamma(n, t_n) \xrightarrow{n \to \infty} 0.
\end{aligned}$$

Now on the left side we use (2.3) to get

$$\left(1 - \left\lfloor \frac{n}{k_n} \right\rfloor \mu(X_0 > u_n)\right)^{k_n} \sim \left(1 - \frac{-y}{k_n}\right)^{k_n} = e^y.$$

□

**Exponential law for the Folklore maps - Part 2**

As in the first chapter, we now try to apply the established theory on our class of examples, the Folklore maps. Independent of our chosen size for the balls, we can show condition $D_2(u_n)$ with theorem 1.1.3. Since indicator functions are in $\mathcal{L}_1(\mu)$ and have bounded variation, we get

$$\begin{aligned}
&\big|\mu(\{X_0 > u_n\} \cap \{M_l \circ T^t \leq u_n\}) - \mu(\{X_0 > u_n\})\mu(\{M_l \leq u_n\})\big| \\
&\qquad \leq K\mu(X_0 > u_n)\big|\big|\mathbb{1}_{\{M_l \leq u_n\}}\big|\big|_{BV} \cdot q^t \leq 3K \cdot q^t
\end{aligned}$$

for some $q \in ]0,1[$. Hence we can for example choose $t_n = \sqrt{n}$ and we get $D_2(u_n)$.

Now to show $D'(u_n)$, take $k_n = \sqrt[3]{n}$ such that $(k_n t_n)_{n\geq 1} = o(n)$. Note that by decay of correlations, we get

$$\begin{aligned}\mu(\{X_0 > u_n\} \cap \{X_j > u_n\}) &\leq \mu(X_0 > u)\mu(X_j > u) \\ &\quad + K\big|\big|\mathbb{1}_{\{X_0>u\}}\big|\big|_{BV}\big|\big|\mathbb{1}_{\{X_j>u\}}\big|\big|_{\mathcal{L}_1(\mu)} \cdot q^n \\ &\leq \mu(X_0 > u_n)^2 + 3K\mu(X_0 > u_n) \cdot q^n.\end{aligned}$$

Now we prove that if the radius of the ball gets smaller and smaller, the minimal return time of the set tends to infinity. We show this for some subsequence $(A_{n_k})_{k\geq 1} := (\{X_0 > u_{n_k}\})_{k\geq 1}$ and use the fact that $\varphi_{A_n} \leq \varphi_{A_{n+1}}$ holds. We want to have

$$\mu_{A_{n_k}}\big(\varphi_{A_{n_k}} \leq k\big) = 0 \tag{2.4}$$

so we fix and define $\varepsilon_k = \min_{1\leq j\leq k} |T^j x^* - x^*| > 0$. Let $\gamma$ be the radius of the largest ball around $x^*$ such that $T$ and it's first $k$ iterations are continuous on that ball. Now choose $\delta_k < \gamma$ so small that

$$|x^* - x| < \delta_k \quad \Rightarrow \quad |T^j x - T^j x^*| < \frac{\varepsilon_k}{2}$$

holds. Then we have for all such close $x$ and for $1 \leq j \leq k$

$$|x^* - T^j x| \geq |x^* - T^j x^*| - |T^j x^* - T^j x| > \frac{\varepsilon_k}{2} > \delta_k. \tag{2.5}$$

Hence if we choose $n_k$ in such a way that the radius of the associated ball is smaller than $\delta_k$, (2.4) is fulfilled. This ensures the existence of some $R_n \nearrow \infty$ such that

$$\mu(A_n \cap T_{-j}A_n) = 0 \text{ for all } j < R_n.$$

But then we can conclude using proposition 2.1.1

$$\begin{aligned}\lim_{n\to\infty} n \sum_{j=1}^{\lfloor n/k_n \rfloor} \mu(A_n \cap T^{-j}A_n) &= \lim_{n\to\infty} n \sum_{j=R_n}^{\lfloor n/k_n \rfloor} \mu(A_n \cap T^{-j}A_n) \\ &\leq \lim_{n\to\infty} n \left\lfloor \frac{n}{k_n} \right\rfloor \mu(A_n)^2 + 2n\mu(A_n)K \sum_{j=R_n}^{\lfloor n/k_n \rfloor} q^j \\ &\leq \lim_{n\to\infty} \frac{(n\mu(A_n))^2}{k_n} + 2n\mu(A_n)K \sum_{j=R_n}^{\infty} q^j \\ &= \lim_{n\to\infty} \frac{y^2}{k_n} - 2yK \sum_{j=R_n}^{\infty} q^j = 0.\end{aligned}$$

### 2.1.2 EVL for the clustering case

If the point $x^*$ is periodic and repelling, then we get the same sort of periodicity for $X_n$ and exceedances may accumulate in a short time period. Unsurprisingly we have to use some different concepts to prove existence of an EVL. The main problem is that since (2.1) holds there exists a parameter $\theta \in ]0,1[$ with

$$\mu(X_0 > u \cap X_p > u) \sim (1-\theta)\mu(X_0 > u).$$

Hence condition $D'(u_n)$ cannot be true any longer. As in [5], we get by (2.3) and EB2

$$n \sum_{j=1}^{\lfloor n/k_n \rfloor} \mu(X_0 > u_n \cap X_j > u_n) \geq n\mu(X_0 > u_n \cap X_p > u_n) \xrightarrow{n\to\infty} (1-\theta)(-y).$$

Since this condition was crucial for proving the existence of an EVL, we have to replace it with some alternative. First we define the annulus of $u_n$ to be

$$Q_p(u) := \{X_0 > u\} \cap \{X_p \leq u\}.$$

This set represents the points that fail to return into the ball after $p$ steps. Besides we define

$$\mathcal{Q}_{p,s,l}(u) := \bigcap_{i=s}^{s+l-1} T^{-i}\big(Q_p(u)\big)^c,$$

which corresponds to no entrances to the annulus from time $s$ to $s+l-1$. As in [5], before getting to the actual EVL, we need the following proposition.

**Proposition 2.1.8.** *If $x^*$ is a repelling periodic point of minimal period $p$ such that EB1 and EB2 are fulfilled, then we have*

$$\lim_{n\to\infty} \mu(M_n \leq u_n) = \lim_{n\to\infty} \mu(\mathcal{Q}_{p,0,n}).$$

*Proof.* First we observe

$$\{M_n \leq u_n\} = \bigcap_{i=0}^{n-1} T^{-i}(\{X_0 \leq u_n\}) \subseteq \bigcap_{i=0}^{n-1} T^{-i}(\{X_0 \leq u_n\} \cup \{X_p > u_n\}) = \mathcal{Q}_{p,0,n}.$$

Now consider $\mathcal{Q}_{p,0,n} \setminus \{M_n \leq u_n\}$. Those are all sets which do not enter any the annulus, but nevertheless have at least one exceedance up to time $n$. Denote the first of those times with $i$. Then $p$ time steps later, we have to have an exceedance again, since the only way to leave $\{X_0 > u_n\}$ in $p$ steps is by ending up in the annulus $Q_p(u_n)$. But we already know this does not happen. The same argument holds for every other multiple of $p$, hence $\{X_{i+jp} > u_n\}$ occurs for all $i + jp < n$. Thereby we showed

$$\mathcal{Q}_{p,0,n} \setminus \{M_n \leq u_n\} \subseteq \bigcup_{i=0}^{n-1} \bigcap_{j=0}^{\lfloor (n-1-i)/p \rfloor} \{X_{i+jp} > u_n\}.$$

Now it is not hard to see

$$\begin{aligned}
\mu(\mathcal{Q}_{p,0,n}) - \mu(M_n \leq u_n) &= \mu\left( \bigcup_{i=0}^{n-1} \bigcap_{j=0}^{\lfloor (n-1-i)/p \rfloor} \{X_{i+jp} > u_n\} \right) \\
&= \mu\left( \bigcup_{i=0}^{n-1} T^{-i} \Big( \bigcap_{j=0}^{\lfloor (n-1-i)/p \rfloor} T^{-jp}(\{X_0 > u_n\}) \Big) \right) \\
&\leq \sum_{i=0}^{n-1} \mu\left( T^{-i} \Big( \bigcap_{j=0}^{\lfloor (n-1-i)/p \rfloor} T^{-jp}(\{X_0 > u_n\}) \Big) \right) \\
&= \sum_{i=0}^{n-1} \mu\left( \bigcap_{j=0}^{\lfloor (n-1-i)/p \rfloor} T^{-jp}(\{X_0 > u_n\}) \right) \\
&\overset{EV2}{\sim} \sum_{i=0}^{n-1} (1-\theta)^{\lfloor (n-1-i)/p \rfloor} \mu(X > u_n) \\
&\leq \underbrace{\mu(X > u_n)}_{\to 0} \frac{p}{1-(1-\theta)} \to 0.
\end{aligned}$$

□

The ideas to proof existence of an EVL in the clustering case a similar to the ones before, but now annuli play the part of the sets $T^{-j}\{X_0 \geq u\}$. Following [5], we introduce new conditions the sequence has to satisfy. The first one captures the ideas from EV1 and EV2.

**Definition 2.1.9.** A sequence $(X_n)_{n \geq 0}$ fulfils condition $SP_{p,\theta}$ for $p \in \mathbb{N}$ and $\theta \in [0,1]$ if

$$\lim_{n\to\infty} \sup_{0<j<p} \mu(X_j > u_n | X_0 > u_n) = 0 \quad \text{and} \quad \lim_{n\to\infty} \mu(X_p > u_n | X_0 > u_n) \to (1-\theta)$$

and furthermore

$$\lim_{n\to\infty}\sum_{i=0}^{\lfloor (n-1)/p\rfloor}\mu\Big(\bigcap_{j=0}^{i}\{X_{jp}>u_n\}\Big)=0$$

for any sequence $(u_n)_{n\geq 0}) = (y/n)_{n\geq 0}$ with $y \leq 0$.

Note that this condition together with (2.3) implies

$$n\mu(X_0 > u_n \cap X_p \leq u_n) = n\mu(X_0 > u_n)\mu(X_p \leq u_n | X_0 > u_n) \xrightarrow{n\to\infty} -y\theta. \tag{2.6}$$

We need two more conditions from [5]:

**Definition 2.1.10.**

- A sequence fulfils condition $D^p(u_n)$ if for any integers $l$, $t$ and $n$

  $$\big|\mu\big(Q_p(u_n)\cap \mathcal{Q}_{p,t,l}(u_n)\big)-\mu\big(Q_p(u_n)\big)\mu\big(\cap\,\mathcal{Q}_{p,0,l}(u_n)\big)\big|\leq \gamma(n,t)$$

  holds where $\gamma(n,t)$ is non-increasing in $t$ and

  $$n\gamma(n,t_n)\to 0 \text{ for } n\to\infty$$

  holds for some sequence $(t_n)_{n\geq 0}\in o(n)$.

- A sequence $(X_n)_{n\geq 0}$ fulfils condition $D'_p(u_n)$ if it fulfils $D^p(u_n)$ and additionally we have

  $$\lim_{n\to\infty} n\sum_{j=1}^{\lfloor n/k_n\rfloor}\mu\big(Q_p\cap T^{-j}(Q_p)\big)=0,$$

  where $k_n$ goes to infinity such that $(k_n t_n)_{n\geq 0}\in o(n)$ for $(t_n)_{n\geq 0}$ from condition $D^p(u_n)$.

Condition $D^p(u_n)$ resembles $D_2(u_n)$ the same way $D'_p(u_n)$ resembles $D'(u_n)$ up to the fact that we replaced $\{X_0 > u_n\}$ with the annulus $Q_p(u_n)$. So analogously to theorem 2.1.7 we can now state

**Theorem 2.1.11.** *If for the sequence $(X_n)_{n\geq 0}$ the conditions $D^p(u_n)$, $D'_p(u_n)$ and $SP_{p,\theta}(u_n)$ hold for some $p\in\mathbb{N}$ and $\theta\in\,]0,1[$, we get the limit law*

$$\lim_{n\to\infty}\mu(M_n\leq u_n)=\lim_{n\to\infty}\mu\big(\mathcal{Q}_{p,0,n}(u_n)\big)=e^{\theta y}$$

*for all $y\leq 0$ and $1$ for $y>0$. We say in this case the sequence has **extremal index** $0<\theta<1$.*

We follow the proof from [5], therefore we need two lemmas.

**Lemma 2.1.12.** *For any integers $p$ and $l \in \mathbb{N}$ and $s \in \mathbb{N}_0$ and real number $u$ we have*

$$\sum_{j=s}^{s+l-1} \mu\big(T^{-j}Q_p(u)\big) \geq \mu\big(\mathcal{Q}_{p,s,l}(u)^c\big)$$

$$\geq \sum_{j=s}^{s+l-1} \mu\big(T^{-j}Q_p(u)\big) - \sum_{j=s}^{s+l-1} \sum_{i=s, i\neq j}^{s+l-1} \mu\big(T^{-j}Q_p(u) \cap T^{-i}Q_p(u)\big).$$

*Proof.* The first inequality comes in fact with the subadditivity of the measure by

$$\mathcal{Q}_{p,s,l}(u)^c = \left(\bigcap_{j=s}^{s+l-1} T^{-j}\big(Q_p(u)\big)^c\right)^c = \bigcup_{j=s}^{s+l-1} T^{-j}\big(Q_p(u)\big).$$

The second inequality is a rough estimate for the error we make by the first one. Denote $T^{-j}Q_p(u) = A_j$. By summing up the measures of the single $A_j$'s, we count subsets which are in the intersection of $k$ single sets $k$ times. We fix this by subtracting it $2k$ times with the double sum. □

**Lemma 2.1.13.** *Let $t$, $m$, $l$ and $s$ be non-negative integers and $u \in \mathbb{R}$. Then on one hand we have*

$$\mu\big(\mathcal{Q}_{p,s,l}(u)\big) - \mu\big(\mathcal{Q}_{p,s,l+t}(u)\big) \leq t\mu\big(Q_p(u)\big) \tag{2.7}$$

*and on the other hand*

$$\left|\mu\big(\mathcal{Q}_{p,0,s+t+m}(u)\big) - \mu\big(\mathcal{Q}_{p,0,m}(u)\big) + \sum_{j=0}^{s-1} \mu\big(Q_p(u) \cap \mathcal{Q}_{p,s+t-j,m}(u)\big)\right|$$

$$\leq t\mu\big(Q_p(u)\big) + 2s\sum_{j=1}^{s-1} \mu\big(Q_p(u) \cap T^{-j}Q_p(u)\big). \tag{2.8}$$

We adapt the proof of a similar result for the non-clustering case from [4].

*Proof.* Proofing the inequality (2.7) works completely analogous to lemma 2.1.4, in fact both statements are the same but this time we consider annuli. So to show (2.8) we observe that we can decompose

$$\mathcal{Q}_{p,0,s+t+m} = \mathcal{Q}_{p,0,s} \cap \mathcal{Q}_{p,s,t} \cap \mathcal{Q}_{p,s+t,m}.$$

Thus

$$\big(\mathcal{Q}_{p,0,s} \cap \mathcal{Q}_{p,s+t,m}\big) \setminus \mathcal{Q}_{p,0,s+t+m} \subseteq \mathcal{Q}^c_{p,s,t}.$$

By $\sigma$-subadditivity we obtain

$$\begin{aligned}\mu(\mathcal{Q}_{p,0,s} \cap \mathcal{Q}_{p,s+t,m}) - \mu(\mathcal{Q}_{p,0,s+t+m}) \leq & \mu(\mathcal{Q}^c_{p,s,t}) = \mu\left(\bigcup_{i=s}^{s+t-1} T^{-i} Q_p(u)\right) \\ & \leq \sum_{i=s}^{s+t-1} \mu\big(T^{-i} Q_p(u)\big) = t\mu\big(Q_p(u)\big). \end{aligned} \tag{2.9}$$

By stationarity we get furthermore

$$\begin{aligned}\mu(\mathcal{Q}_{p,0,s} \cap \mathcal{Q}_{p,s+t,m}) &= \mu(\mathcal{Q}_{p,s+t,m}) - \mu(\mathcal{Q}^c_{p,0,s} \cap \mathcal{Q}_{p,s+t,m}) \\ &= \mu(\mathcal{Q}_{p,0,m}) - \mu(\mathcal{Q}^c_{p,0,s} \cap \mathcal{Q}_{p,s+t,m}) \\ &\geq \mu(\mathcal{Q}_{p,0,m}) - \sum_{j=0}^{s-1} \mu\big(T^{-j} Q_p(u) \cap \mathcal{Q}_{p,s+t,m}\big).\end{aligned}$$

The non-trivial inequality of lemma 2.1.12 and stationarity lead us to

$$\begin{aligned}\mu(\mathcal{Q}_{p,0,s} \cap \mathcal{Q}_{p,s+t,m}) &= \mu(\mathcal{Q}_{p,s+t,m}) - \mu(\mathcal{Q}^c_{p,0,s} \cap \mathcal{Q}_{p,s+t,m}) \\ &= \mu(\mathcal{Q}_{p,0,m}) - \mu(\mathcal{Q}^c_{p,0,s} \cap \mathcal{Q}_{p,s+t,m}) \\ &\leq \mu(\mathcal{Q}_{p,0,m}) - \sum_{j=0}^{s-1} \mu\big(T^{-j} Q_p(u) \cap \mathcal{Q}_{p,s+t,m}\big) \\ &+ \sum_{j=0}^{s-1} \sum_{l=0, j\neq l}^{s-1} \mu\big(T^{-j} Q_p(u) \cap T^{-l} Q_p(u) \cap \mathcal{Q}_{p,s+t,m}\big).\end{aligned}$$

The two foregoing inequalities together give

$$
\begin{aligned}
&\left|\mu(\mathcal{Q}_{p,0,s} \cap \mathcal{Q}_{p,s+t,m}) - \mu(\mathcal{Q}_{p,0,m}) + \sum_{j=0}^{s-1} \mu\big(T^{-j} Q_p(u) \cap \mathcal{Q}_{p,s+t,m}\big)\right| \\
&\leq \sum_{j=0}^{s-1} \sum_{l=0, j\neq l}^{s-1} \mu\big(T^{-j} Q_p(u) \cap T^{-l} Q_p(u) \cap \mathcal{Q}_{p,s+t,m}\big) \\
&\leq \sum_{j=0}^{s-1} \sum_{l=0, j\neq l}^{s-1} \mu\big(T^{-j} Q_p(u) \cap T^{-l} Q_p(u)\big) \\
&= \sum_{j=0}^{s-1} \sum_{l=0, j\neq l}^{s-1} \mu\big(Q_p(u) \cap T^{-|l-j|} Q_p(u)\big) \\
&\leq 2s \sum_{j=1}^{s-1} \mu\big(Q(u) \cap T^{-j} Q(u)\big),
\end{aligned}
$$

where we reordered the double sum using stationarity. Finally, we can combine this with 2.9 and conclude

$$
\begin{aligned}
&\left|\mu\big(\mathcal{Q}_{p,0,s+t+m}(u)\big) - \mu\big(\mathcal{Q}_{p,0,m}(u)\big) + \sum_{j=0}^{s-1} \mu\big(Q_p(u) \cap \mathcal{Q}_{p,s+t-j,m}(u)\big)\right| \\
&\leq \left|\mu(\mathcal{Q}_{p,0,s} \cap \mathcal{Q}_{p,s+t,m}) - \mu(\mathcal{Q}_{p,0,s+t+m})\right| \\
&+ \left|\mu(\mathcal{Q}_{p,0,s} \cap \mathcal{Q}_{p,s+t,m}) - \mu(\mathcal{Q}_{p,0,m}) + \sum_{j=0}^{s-1} \mu\big(T^{-j} Q_p(u) \cap \mathcal{Q}_{p,s+t,m}\big)\right| \\
&\leq t\mu\big(Q_p(u)\big) + 2s \sum_{j=1}^{s-1} \mu\big(Q_p(u) \cap T^{-j} Q_p(u)\big).
\end{aligned}
$$

□

With all those technical preparations we can now finally proof the theorem.

*Proof of theorem 2.1.11.* We have already shown the first equality in proposition 2.1.8. For the other one, define for the use of $D_p'(u_n)$ $k = k_n$ and $l = l_n := \lfloor \frac{n}{k_n} \rfloor$. Using (2.7) and $Q_p(u_n) \subseteq \{X_0 > u_n\}$ we get

$$
\left|\mu\big(\mathcal{Q}_{p,0,n}(u_n) - \mu(\mathcal{Q}_{p,0,k(l+t)}(u_n)\big)\right| \leq kt\mu\big(Q_p(u_n)\big) \leq kt\mu(X_0 > u_n). \tag{2.10}
$$

Now we use (2.8) and stationarity to show for $1 \le i \le k$

$$
\begin{aligned}
&\left|\mu(\mathcal{Q}_{p,0,i(l+t)}(u_n)) - \big(1 - l\mu(Q_p(u_n))\big)\mu\big(\mathcal{Q}_{p,0,(i-1)(l+t)}(u_n)\big)\right| \\
&\le \left| l\mu\big(Q_p(u_n)\big)\mu\big(\mathcal{Q}_{p,0,(i-1)(l+t)}(u_n)\big) - \sum_{j=0}^{l-1} \mu\big(T^j Q_p(u_n) \cap \mathcal{Q}_{p,l+t,(i-1)(l+t)}(u_n)\big) \right| \\
&+ \left| \sum_{j=0}^{l-1} \mu\big(T^j Q_p(u_n) \cap \mathcal{Q}_{p,l+t,(i-1)(l+t)}(u_n)\big) + \mu(\mathcal{Q}_{p,0,i(l+t)}(u_n)) - \mu(\mathcal{Q}_{p,0,(i-1)(l+t)}(u_n)) \right| \\
&\le \left| l\mu\big(Q_p(u_n)\big)\mu\big(\mathcal{Q}_{p,0,(i-1)(l+t)}(u_n)\big) - \sum_{j=0}^{l-1} \mu\big(T^{-j} Q_p(u_n) \cap \mathcal{Q}_{p,l+t,(i-1)(l+t)}(u_n)\big) \right| \\
&+ t\mu(X_0 > u_n) + 2l \sum_{j=1}^{l-1} \mu\big(Q_p(u_n) \cap T^{-j} Q_p(u_n)\big) := \Gamma_{n,i}
\end{aligned}
$$

Note that for the case $i = 0$ we identify $\mathcal{Q}_{p,0,0}$ with the whole space $\mathbb{X}$. The given estimate is not good enough, we want to find a bound which is actually independent of $i$. Hence we use $D^p(u_n)$ and the fact that $\gamma(n,t)$ is non-increasing in $t$. So

$$
\begin{aligned}
\Gamma_{n,i} &= \left| \sum_{j=0}^{l-1} \Big( \mu\big(Q_p(u_n)\big)\mu\big(\mathcal{Q}_{p,0,(i-1)(l+t)}(u_n)\big) - \mu\big(T^{-j} Q_p(u_n) \cap \mathcal{Q}_{p,l+t,(i-1)(l+t)}(u_n)\big) \Big) \right| \\
&\quad + t\mu(X_0 > u_n) + 2l \sum_{j=1}^{l-1} \mu\big(Q_p(u_n) \cap T^{-j} Q_p(u_n)\big) \\
&\sum_{j=0}^{l-1} \left| \mu\big(Q_p(u_n)\big)\mu\big(\mathcal{Q}_{p,0,(i-1)(l+t)}(u_n)\big) - \mu\big(Q_p(u_n) \cap \mathcal{Q}_{p,l+t-j,(i-1)(l+t)}(u_n)\big) \right| \\
&\quad + t\mu(X_0 > u_n) + 2l \sum_{j=1}^{l-1} \mu\big(Q_p(u_n) \cap T^{-j} Q_p(u_n)\big) \\
&\le l\gamma(n,t) + t\mu(X_0 > u_n) + 2l \sum_{j=1}^{l-1} \mu\big(Q_p(u_n) \cap T^{-j} Q_p(u_n)\big) := \Upsilon_n.
\end{aligned}
$$

In (2.6) we showed $n\mu(Q_p(u_n)) \to -\theta y$. Hence there exist $n$ and $k_n$ large enough such that

$$
l\mu(Q_p(u_n)) \le \left(\frac{n}{k_n} + 1\right)\mu\big(Q_p(u_n)\big) < 2
$$

holds. With the exact same induction arguments as in the proof of proposition 2.1.6, we obtain

$$\left|\mu\big(\mathcal{Q}_{p,0,k(l+t)}(u_n)\big) - \big(1 - l\mu(Q_p(u_n))\big)^k\right| \leq k\Upsilon_n.$$

Using this together with (2.10) now gives

$$\begin{aligned}\left|\mu\big(\mathcal{Q}_{p,0,n}(u_n)\big) - \big(1 - l\mu(Q_p(u_n))\big)^k\right| &\leq \left|\mu\big(\mathcal{Q}_{p,0,n}(u_n) - \mu\big(\mathcal{Q}_{p,0,k(l+t)}(u_n)\big)\right| \\ &\quad + \left|\mu\big(\mathcal{Q}_{p,0,k(l+t)}(u_n)\big)\big(1 - l\mu(Q_p(u_n))\big)^k\right| \\ &\leq kt\mu\big(Q_p(u_n)\big) + k\Upsilon_n.\end{aligned}$$

Note that by (2.6) the subtrahend converges

$$\begin{aligned}\lim_{n\to\infty}\left(1 - \left\lfloor\frac{n}{k}\right\rfloor\mu\big(Q_p(u_n)\big)\right)^k &\sim \lim_{n\to\infty}\left(1 - \frac{n\mu\big(Q_p(u_n)\big)}{k}\right)^k \\ &\sim \lim_{n\to\infty}\left(1 - \frac{-\theta y}{k}\right)^k = e^{\theta y}.\end{aligned}$$

So it suffices to show that $kt\mu\big(Q_p(u_n)\big) + k\Upsilon_n$ vanishes in the limit, i.e.

$$\lim_{n\to\infty} kt\mu\big(Q_p(u_n)\big) + kt\mu(X_0 > u_n) + kl\gamma(n,t) + 2kl\sum_{j=1}^{l-1}\mu\big(Q_p(u_n) \cap T^{-j}Q_p(u_n)\big) = 0.$$

Since we took $k = k_n$ and $t = t_n$ as in $D'_p(u_n)$, we know that the first part vanishes since $kt = k_n t_n \in o(n)$ and by (2.3) $n\mu\big(Q_p(u_n)\big) \leq n\mu(X_0 > u_n) = -y$ for $u_n = y/n$ holds. The same argument applies for the second term. The product $n\gamma(n,t)$ goes to zero by $D^p(u_n)$ and finally the sum vanishes by $D'_p(u_n)$ again. □

### 2.1.3 Classic distributions for independent sequences

In [3] and [7], there are three families of distribution functions presented which occur in the study of maxima of independent random sequences. The **Gumbel** extreme value distribution is given by

$$F(y) = e^{-e^{-y}} \text{ for } y \in \mathbb{R},$$

the family of **Fréchet** extreme value distributions for $\alpha > 0$ by

$$F(y) = \begin{cases} e^{\frac{1}{y^{\alpha}}} & \text{for } y > 0 \\ 0 & \text{else}\end{cases}$$

and finally the **Weibull** family of distributions for $\alpha > 0$ by

$$F(y) = \begin{cases} 1 & \text{for } y > 0 \\ e^{-(-y)^\alpha} & \text{for } y \leq 0. \end{cases}$$

As a matter of fact, with the right normalisation we can find those distribution in our system, too. Consider

$$F(y) = \mu\big(\gamma_n \psi(M_n) - \vartheta_n \leq y\big).$$

- The choice $\psi(x) = -\ln(-x)$, $\gamma_n = 1$ and $\vartheta_n = \ln n$ gives

$$\begin{aligned} \mu\big(-\ln(-M_n) - \ln n \leq y\big) &= \mu\big(-\ln(-M_n) \leq y + \ln n\big) \\ &= \mu\big(\ln(-M_n) \geq -(y + \ln n)\big) \\ &= \mu\Big(-M_n \geq e^{-(y+\ln n)}\Big) \\ &= \mu\Big(nM_n \leq -ne^{-(y+\ln n)}\Big) \\ &= \mu\Big(nM_n \leq -e^{-y}\Big) \to e^{-e^{-y}}, \end{aligned}$$

  so the Gumbel distribution.

- To get Weibull in the limit, we choose $\psi(x) = (-x)^{\frac{1}{\alpha}}$, $\gamma_n = -n^{\frac{1}{\alpha}}$ and $\vartheta_n = 0$. Then for $y \leq 0$ we obtain

$$\begin{aligned} \mu\Big(-n^{\frac{1}{\alpha}}(-M_n)^{\frac{1}{\alpha}} \leq y\Big) &= \mu\left((-M_n)^{\frac{1}{\alpha}} \geq \frac{-y}{n^{\frac{1}{\alpha}}}\right) \\ &= \mu\left(-M_n \geq \frac{(-y)^\alpha}{n}\right) = \mu\big(nM_n \leq -(-y)^\alpha\big) \to e^{-(-y)^\alpha} \end{aligned}$$

  and for $y > 0$

$$\mu\Big(-n^{\frac{1}{\alpha}}(-M_n)^{\frac{1}{\alpha}} \leq y\Big) = \mu\left(\underbrace{(-M_n)^{\frac{1}{\alpha}}}_{\geq 0} \geq \underbrace{\frac{-y}{n^{\frac{1}{\alpha}}}}_{\leq 0}\right) = 1.$$

## 2.2 Connection between extreme value laws and hitting-time statistics

Now there is a certain duality between hitting times and extreme values, in fact, they are two sides of the same coin. Notice for a ball with radius $r$ around some point $x^*$ we have the following useful

**Lemma 2.2.1.** *Let $M_n$ and $\varphi$ be defined as above. Then we have*

$$\big\{\varphi_{B_r(x^*)} > n\big\} = \big\{M_n \circ T \le -\mu\big(B_r(x^*)\big)\big\}.$$

*Proof.* A short calculation gives

$$\begin{aligned}
\big\{\varphi_{B_r(x^*)} > n\big\} &= \big\{\forall i \in \{1,\dots,n\} : |T^i x - x^*| \ge r\big\} \\
&= \big\{\forall i \in \{1,\dots,n\} : B_{|T^i x - x^*|}(x^*) \supseteq B_r(x^*)\big\} \\
&= \big\{\forall i \in \{1,\dots,n\} : \mu\big(B_{|T^i x - x^*|}(x^*)\big) \ge \mu\big(B_r(x^*)\big)\big\} \\
&= \big\{\forall i \in \{1,\dots,n\} : -\mu\big(B_{|T^i x - x^*|}(x^*)\big) \le -\mu\big(B_r(x^*)\big)\big\} \\
&= \big\{\forall i \in \{1,\dots,n\} : X_i \le -\mu\big(B_r(x^*)\big)\big\} \\
&= \big\{M_n \circ T \le -\mu\big(B_r(x^*)\big)\big\}.
\end{aligned}$$

□

Note that in the third equality we used the equivalence of $\mu$ and the reference measure. The conclusion that two balls with equal measure have equal radius is wrong, as it depends on the null sets of the measure. Following [6] and [3], we can finally combine HTS and EVL.

**Theorem 2.2.2.** *Let $(X, A, \mu, T)$ be a dynamical system and $\mu$ an equivalent probability measure with respect to Lebesgue. Let $M_n$ be defined as above. Then we have hitting time statistics for small balls around $x^*$ if and only if we have an extreme value law for $M_n$. In particular we get*

$$nM_n \xRightarrow{\mu} R \quad \text{if and only if} \quad \mu\big(B_l(x^*)\big)\varphi_{B_l(x^*)} \xRightarrow{\mu} -R.$$

The proof follows the ideas of [6]. However, by our choice of $X_n$ it gets simplified since we have now limits in (2.3).

*Proof.* (i) Assume we have hitting time statistics for balls. By a short calculation we obtain

$$\begin{aligned}
\{M_n \le u_n\} = \bigcap_{j=0}^{n-1} \{X_j \le u_n\} &= \bigcap_{j=0}^{n-1} \big\{x : -\mu(B_{|T^j x - x^*|}(x^*)) \le u_n\big\} \\
&= \bigcap_{j=0}^{n-1} \big\{x : \mu\big(B_{|T^j x - x^*|}(x^*)\big) \ge -u_n\big\}.
\end{aligned}$$

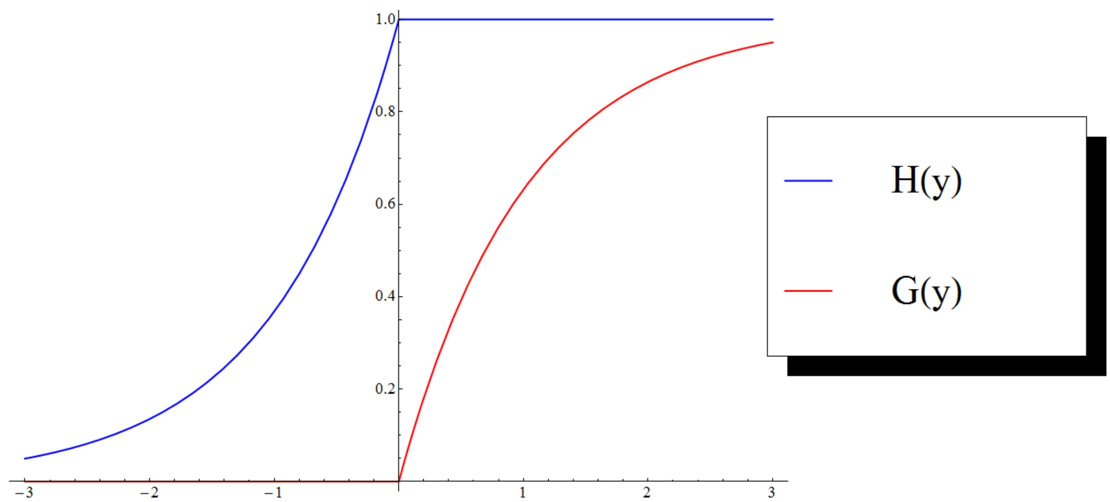


Figure 2.1: The two distribution functions for the exponential case

Recalling $l(\gamma) = \inf\{\eta > 0 : \mu(B_\eta(x^*)) = \gamma\}$ we can continue

$$\begin{aligned}
\mu(\{M_n \leq u_n\}) &= \mu\left(\bigcap_{j=0}^{n-1} \left\{x : \mu\big(B_{|T^j x - x^*|}(x^*)\big) \geq \mu\big(B_{l(-u_n)}(x^*)\big)\right\}\right) \\
&= \mu\left(\bigcap_{j=0}^{n-1} \left\{x : |T^j x - x^*| \geq l(-u_n)\right\}\right) \\
&= \mu\left(\bigcap_{j=1}^{n} \left\{x : |T^j x - x^*| \geq l(-u_n)\right\} \circ T^{-1}\right) \\
&= \mu\Big(\Big\{x : \varphi_{B_{l(-u_n)}(x^*)}(x) > n\Big\}\Big) \qquad (2.11)
\end{aligned}$$

Note that we have

$$n = \frac{y}{u_n} = \frac{-y}{\mu(B_{l(-u_n)}(x^*))}. \qquad (2.12)$$

Now we can conclude

$$H(y) = \lim_{n\to\infty} \mu(M_n \leq u_n) = \lim_{n\to\infty} \mu\left(\varphi_{B_{l(-u_n)}(x^*)}(x) > \frac{-y}{\mu(B_{l(-u_n)}(x^*))}\right) = 1 - G(-y).$$

(ii) To show the other direction, we use almost the same strategy. Let $(\delta_n)_{n\geq 0}$ be a sequence of decreasing radii, in particular $\delta_n \to 0$. For the hitting time distribution, we need positive arguments while $y < 0$. Hence we define

$$\kappa_n := \left\lceil \frac{-y}{\mu\big(B_{\delta_n}(x^*)\big)} \right\rceil$$

and note that $\kappa_n \to \infty$ holds for $n \to \infty$. Since $-u_n n = y$ holds, we especially get with (2.3)

$$-u_{\kappa_n} = \frac{y}{\kappa_n} \sim \mu\big(B_{\delta_n}(x^*)\big). \tag{2.13}$$

Now as in (2.11) we get

$$\mu(M_{\kappa_n} \leq u_{\kappa_n}) = \mu\Big(\varphi_{B_{l(-u_{\kappa_n})}(x^*)} \geq \kappa_n\Big).$$

Our goal is to show

$$1 - G(-y) = \lim_{n\to\infty} \mu\left(\varphi_{B_{\delta_n}(x^*)} \geq \frac{-y}{\mu\big(B_{\delta_n}(x^*)\big)}\right) = \lim_{n\to\infty} \mu\big(M_{\kappa_n} \leq u_{\kappa_n}\big) = H(y).$$

So we are still missing the first equality. To show it, we use the triangle inequality for

$$\begin{aligned}\mu\left(\varphi_{B_{\delta_n}(x^*)} \geq \frac{t}{\mu\big(B_{\delta_n}(x^*)\big)}\right) - \mu(M_{\kappa_n} \leq u_{\kappa_n}) &= \Big(\mu\big(\varphi_{B_{\delta_n}(x^*)} \geq \kappa_n\big) - \mu(M_{\kappa_n} \leq u_{\kappa_n})\Big)\\ &+\left(\mu\left(\varphi_{B_{\delta_n}(x^*)} \geq \frac{t}{\mu\big(B_{\delta_n}(x^*)\big)}\right) - \mu\big(\varphi_{B_{\delta_n}(x^*)} \geq \kappa_n\big)\right).\end{aligned}$$

For the first term, observe that if $A \subseteq B$ or $B \subseteq A$ holds, we get

$$|\mu(A) - \mu(B)| = \mu(A \triangle B).$$

With that idea, we obtain

$$\begin{aligned}
&\Big|\mu\big(\varphi_{B_{\delta_n}(x^*)} \geq \kappa_n\big) - \mu(M_{\kappa_n} \leq u_{\kappa_n})\Big| \\
&= \Big|\mu\big(\varphi_{B_{\delta_n}(x^*)} \geq \kappa_n\big) - \mu\big(\varphi_{B_{l(-u_{\kappa_n})}(x^*)} \geq \kappa_n\big) \\
&= \mu\left(\bigcup_{j=1}^{\kappa_n} T^{-j}\big(B_{\delta_n}(x^*) \triangle B_{l(-u_{\kappa_n})}(x^*)\big)\right) \\
&\leq \sum_{j=1}^{\kappa_n} \mu\Big(T^{-j}\big(B_{\delta_n}(x^*) \triangle B_{l(-u_{\kappa_n})}(x^*)\big)\Big) \\
&= \kappa_n \mu\Big(\big(B_{\delta_n}(x^*) \triangle B_{l(-u_{\kappa_n})}(x^*)\big)\Big) \\
&\sim \frac{-y}{\mu\big(B_{\delta_n}(x^*)\big)}\big|\mu\big(B_{\delta_n}(x^*)\big) - \mu\big(B_{l(-u_{\kappa_n})}(x^*)\big)\big| \\
&= -y\left|1 - \frac{\mu\big(B_{l(-u_{\kappa_n})}(x^*)\big)}{\mu\big(B_{\delta_n}(x^*)\big)}\right| = -y\left|1 - \frac{-u_{\kappa_n}}{\mu\big(B_{\delta_n}(x^*)\big)}\right|,
\end{aligned}$$

which tends to 0 by (2.13). For the second term, observe that $\varphi_{B_{\delta_n}(x^*)}$ is integer-valued. Hence

$$\left\{\varphi_{B_{\delta_n}(x^*)} \geq \frac{t}{\mu\big(B_{\delta_n}(x^*)\big)}\right\} = \left\{\varphi_{B_{\delta_n}(x^*)} \geq \left\lceil \frac{t}{\mu\big(B_{\delta_n}(x^*)\big)} \right\rceil\right\} = \big\{\varphi_{B_{\delta_n}(x^*)} \geq \kappa_n\big\},$$

and the difference is 0 in the first place. □

# Bibliography


[1] Aytaç H, Freitas JM. Laws of rare events for deterministic and random dynamical systems. arXic:1207.5188v2

[2] Billingsley P, Convergence of probability measures. 2nd edition, John Wiley & Sons, Inc 1999

[3] Freitas JM. Extremal behaviour of chaotic dynamics. Dynamical Systems, 2013 Vol. 28, No. 3, p. 302-332

[4] Freitas ACM, Freitas JM. On the link between dependence and independence in extreme value theory for dynamical systems. Stat Probab Lett. 2008;78(9):1088-1093

[5] Freitas ACM, Freitas JM, Todd M. Extremal index, hitting time statistics and periodicity. arXiv:1008.1350

[6] Freitas ACM, Freitas JM, Todd M. Extreme value laws in dynamical systems for non-smooth observations. arXiv:1006.3276v1

[7] Freitas ACM, Freitas JM, Todd M. Hitting time statistics and extreme value theory. arXiv:0804.2887

[8] Freitas ACM, Freitas JM, Todd M. The compound Poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics. arXiv:1204.2304

[9] Haydn N, Lacroix Y, Vaienti S. Hitting and return times in ergodic dynamical systems. Ann. Probab. 33 (2005), 2043-2050

[10] Karr A. Point processes and their statistical inference. 2nd edition, Marcel Dekker, Inc 1991

[11] Klenke A, Wahrscheinlichkeitstheorie. 3. Auflage Springer-Verlag Berlin Heidelberg 2013

[12] Kupsa M, Lacroix Y. Asymptotics for hitting times. Ann. Probab. 33 (2005) 610-619.

[13] Lacroix Y. Possible limit laws for entrance times of an ergodic aperiodic dynamical system. Israel J. Math. 132 (2002), 253-263.

[14] Leadbetter MR, Lindgren G, Rootzén H. Extremes and related properties of random sequences and processes. Springer Serien in Statistics, New York: Springer Verlag (1983)

[15] Leadbetter MR. On extreme values in stationary sequences. Z Wahrscheinlichkeitstheorie und verwandte Gebiete. 1973/74;28:289-303

[16] Rychlik M. Bounded variation and invariant measures. Studia Math. 76 (1983) 69–80

[17] Zweimüller R. Asymptotically rare events, hitting-time and return-time statistics. Unpublished lecture notes.

[18] Zweimüller R. Lecture notes "Ergodentheorie 1", winter term 2015, University of Vienna

[19] Zweimüller R. Lecture notes "Ergodentheorie 2", sommer term 2016, University of Vienna

[20] Zweimüller R. Mean ergodic theory and Poisson limits. Unpublished preprint, version from 10.5.2016

[21] Zweimüller R. Mixing limit theorems for ergodic transformations. Journal of Theoretical Probability 20 (2007), 1059-1071

[22] Zweimüller R. The general asymptotic return-time process. Isr. J. Math. (2016) 212: 1.

# Abstract

Für seltene Ereignisse werden hinreichende Bedingungen formuliert, durch welche im Grenzübergang eine Exponentialverteilung erzeugt wird. Dies wird durch zwei verschiedene Herangehensweisen verwirklicht, einerseits durch Annahme bestimmter Mischungseigenschaften des Systems entsprechend einiger Arbeiten von Jorge Milhazes Freitas, andererseits durch aktuelle Forschungsergebnisse von Roland Zweimüller, wobei hierbei auf Mischungseigenschaften verzichtet wird. Die Resultate werden auf eine Klasse von Intervallabbildungen angewendet.

Wien, 30. August 2016