Annales Henri Lebesgue
5 (2022) 1361-1415

ANNALES
HENRI LEBESGUE

ROLAND ZWEIMÜLLER

HITTING TIMES AND POSITIONS IN RARE EVENTS

TEMPS D'ATTEINTE ET POSITIONS POUR LES ÉVÉNEMENTS RARES

Abstract. — We establish abstract limit theorems which provide sufficient conditions for a sequence (A_l) of rare events in an ergodic probability preserving dynamical system to exhibit Poisson asymptotics, and for the consecutive positions inside the A_l to be asymptotically iid (spatiotemporal Poisson limits). The limit theorems only use information on what happens to A_l before some time τ_l which is of order $o(1/\mu(A_l))$. In particular, no assumptions on the asymptotic behavior of the system akin to classical mixing conditions are used. We also discuss some general questions about the asymptotic behaviour of spatial and spatiotemporal processes, and illustrate our results in a setup of simple prototypical systems.

Résumé. — Nous établissons des théorèmes limites abstraits qui fournissent des conditions suffisantes pour qu'une suite (A_l) d'événements rares dans un système préservant une mesure de probabilité ergodique satisfasse des asymptotiques de Poisson, et pour que les positions dans A_l soient asymptotiquement iid (limites de Poisson spatio-temporelles). Les théorèmes limites n'utilisent que des informations sur ce qui arrive à A_l avant un certain temps τ_l d'ordre $o(1/\mu(A_l))$. En particulier, nous n'utilisons aucune hypothèse sur le comportement asymptotique du système du type conditions de mélange classiques. Nous discutons également quelques questions générales sur le comportement asymptotique des processus spatiaux et spatio-temporels, et illustrons nos résultats avec des systèmes prototypiques simples.

Keywords: invariant measure, limit distribution, rare events, Poisson process.
2020 *Mathematics Subject Classification:* 28D05, 37A25, 37A50, 37C30, 60F05, 11K50.
DOI: `https://doi.org/10.5802/ahl.154`

1. Introduction

Consider an ergodic measure-preserving map T on the probability space $(X, \mathcal{A}, \mu)$, and a sequence $(A_l)_{l \geqslant 1}$ of sets for which $0 < \mu(A_l) \to 0$. Let φ_{A_l} denote the first hitting time function of A_l. The asymptotic behavior of the distributions of the rescaled hitting times $\mu(A_l)\varphi_{A_l}$ as $l \to \infty$ is a well-studied circle of questions. In many interesting situations, mixing properties have been used to show that these *hitting time distributions* converge to an exponential law,

$$\mu\left(\mu(A_l)\varphi_{A_l} \leqslant t\right) \longrightarrow 1 - e^{-t} \quad \text{as} \quad l \to \infty \quad \text{for} \quad t > 0, \tag{1.1}$$

and so do the corresponding *return distributions*,

$$\mu_{A_l}\left(\mu(A_l)\varphi_{A_l} \leqslant t\right) \longrightarrow 1 - e^{-t} \quad \text{as} \quad l \to \infty \quad \text{for} \quad t > 0, \tag{1.2}$$

where μ_{A_l} denotes the normalized restriction of μ to A_l. In fact, (1.1) and (1.2) are equivalent: as a consequence of [HSV99, Theorem 2.1] one has

Proposition 1.1 (Characterizing convergence to the exponential law). — *Let T be an ergodic measure-preserving map on the probability space $(X, \mathcal{A}, \mu)$, and $(A_l)_{l \geqslant 1}$ a sequence in $\mathcal{A}$ such that $0 < \mu(A_l) \to 0$. Then each of (1.1) and (1.2) is equivalent to*

$$\mu\left(\mu(A_l)\varphi_{A_l} \leqslant t\right) - \mu_{A_l}\left(\mu(A_l)\varphi_{A_l} \leqslant t\right) \longrightarrow 0 \quad \text{as} \quad l \to \infty \quad \text{for} \quad t > 0. \tag{1.3}$$

This fact leads to one standard approach (of many) to proving (1.1) and (1.2): Checking condition (1.3) means to show that, asymptotically, changing μ_{A_l} to μ does not affect the resulting distribution. This is not trivial, because the μ_{A_l} become increasingly singular, being concentrated on smaller and smaller sets.

For many classes of concrete systems, strong results on the decay of correlations (or mixing properties) provide information on how the system forgets the difference between two initial probabilities over time. A basic form of this might state that $d_{\mathfrak{N}}(\nu \circ T^{-n}, \mu) \leqslant c_n$ for $\nu \in \mathfrak{N}$ and $n \geqslant 0$, where $c_n \to 0$, and $\mathfrak{N}$ is a family of normalized measures, typically rather small, and equipped with some metric $d_{\mathfrak{N}}$.

We can then hope to establish (1.3) once we check it is possible to replace the measures μ_{A_l} there by push-forwards $\mu_{A_l} \circ T^{-\tau_l}$ (with integers τ_l) which are nicer in that they belong to $\mathfrak{N}$, and hence allow comparison to μ via the control of $d_{\mathfrak{N}}(\nu \circ T^{-n}, \mu)$ on $\mathfrak{N}$. Taking the push-forward means to skip the first τ_l time steps, and we need those to be negligible compared to the variable $\mu(A_l)\varphi_{A_l}$ itself, meaning that $\mu(A_l)\tau_l \to 0$. Also, one has to check that in skipping these steps, we do not miss (with positive asymptotic probability) the awaited visit to A_l, which amounts to requiring that $\mu_{A_l}(\varphi_{A_l} \leqslant \tau_l) \longrightarrow 0$.

Note that the existence of τ_l meeting the last two conditions is in fact necessary for (1.2) (otherwise, the limit law contains an atom at the origin), and that this is a property ("no short returns") of the specific sequence (A_l) of sets, which always needs to be checked (since every system contains sequences for which it fails). The correlation decay, on the other hand, is a feature of the whole system, and the two are tied together by the requirement that $\mu_{A_l} \circ T^{-\tau_l} \in \mathfrak{N}$ for all $l \geqslant 1$.

The first purpose of the present paper is to point out that the same strategy can be used, even for functional versions for the processes of consecutive hitting-times, without assuming any information on the decay of correlations (and without the system being mixing). We only ask for a sequence of (not necessarily constant) delay times τ_l satisfying the necessary conditions $\mu(A_l)\tau_l \to 0$ and $\mu_{A_l}(\varphi_{A_l} \leqslant \tau_l) \longrightarrow 0$, and such that the time-τ_l-measures $\mu_{A_l} \circ T^{-\tau_l}$ belong to some set $\mathfrak{K}$ of probabilities which is compact in total variation norm. This latter condition can be seen as a short-time (since $\tau_l = o(1/\mu(A_l))$) decorrelation property of (A_l).

Second, we show that the same approach can be used to analyse distributional limits of the sequences of consecutive positions, inside the A_l, of orbits upon their visits to these small sets, and of joint time-position processes. Such results on spatiotemporal Poisson limits have recently been introduced in [PS20]. We also include a general discussion of some aspects of spatial and spatio-temporal process limits in the abstract setup, and illustrate our results in the context of some simple prototypical systems.

There is a large body of literature devoted to the asymptotics of return- and hitting-time distributions. While some sources directly relevant for the present work are mentioned in the bibliography, we do not attempt to provide a complete overview. A closer look at earlier articles which derive similar limit theorems on the basis of mixing conditions or on correlation decay reveals that, effectively, these assumptions are often only used to obtain control on the time scale mentioned above, see for example [Aba04, AS11, CC13, FFT12, HP14, HY16, PS16, PS20]. The present paper makes this more explicit, and emphasizes that in the presence of mere ergodicity a comparatively weak form of short-time control suffices.

Acknowledgements

I am grateful to the refereree for comments and questions which helped to clarify some points, and to Francoise Pène, Benoit Saussol and Max Auer for discussions on related matters.

2. General setup and preparations

2.1. Hitting- and return-times. Inducing

Throughout, $(X, \mathcal{A}, \mu)$ is a probability space, and $T : X \to X$ is an ergodic μ-preserving map. Also, A and A_l will always denote measurable sets of (strictly) *positive* measure. By ergodicity and the Poincaré recurrence theorem, the measurable *(first) hitting time* function of A, $\varphi_A : X \to \overline{\mathbb{N}} := \{1, 2, \ldots, \infty\}$ with $\varphi_A(x) := \inf\{n \geqslant 1 : T^n x \in A\}$, is finite a.e. on X. When restricted to A it is called the *(first) return time* function of the set. Define $T_A x := T^{\varphi_A(x)} x$ for a.e. $x \in X$, which gives the *first entrance map* $T_A : X \to A$. It is a standard fact that its restriction to A, the *first return map* $T_A : A \to A$ is an ergodic measure preserving map on the probability space $(A, \mathcal{A} \cap A, \mu_A)$, where $\mu_A(B) := \mu(A \cap B)/\mu(A)$, $B \in \mathcal{A}$. By Kac' formula, $\int_A \varphi_A \, d\mu_A = 1/\mu(A)$. That is, when regarded as a random variable on

$(A, \mathcal{A} \cap A, \mu_A)$, the return time has expectation $1/\mu(A)$, and we will often normalize these functions accordingly, thus considering $\mu(A)\,\varphi_A$.

The focus of this work is on asymptotic distributions of such normalized hitting (or return) times, and of the positions inside the target set at which an orbit hits. We shall study processes of consecutive hitting times and hitting places in the limit of very small sets. Call $(A_l)_{l \geqslant 1}$ a *sequence of asymptotically rare events* (or an *asymptotically rare sequence*) provided that $A_l \in \mathcal{A}$ and $0 < \mu(A_l) \to 0$.

It will be natural to view various observables defined on (parts of) X through different probability measures ν. In the present paper we shall focus on the family $\mathfrak{P} := \{\nu : \text{probability measure on } (X, \mathcal{A}),\ \nu \ll \mu\}$, equipped with the *total variation distance* $d_{\mathfrak{P}}(\nu, \nu') := 2 \sup_{A \in \mathcal{A}} \mid \nu(A) - \nu'(A) \mid$. This induces a topology so weak that mere ergodicity ensures convergence of averages (via Theorem 6.2), but strong enough for the latter to have substantial consequences. (The results below rely on compactness conditions which clearly would not be sufficient if we used a w*-topology instead.)

The push-forward of a measure ν by T will be denoted $T_*\nu := \nu \circ T^{-1}$, and likewise for measurable maps other than T. Indeed, we shall use suitable *times*, that is, measurable functions $\tau : B \to \mathbb{N}_0 := \{0, 1, \ldots\}$ with $B \in \mathcal{A}$, to define auxiliary *induced maps* $T^\tau : B \to X$ via $T^\tau x := T^{\tau(x)} x$. Given $\nu \in \mathfrak{P}$, the push-forward $T_*^\tau \nu := \nu \circ (T^\tau)^{-1}$ then is the distribution, at the (possibly random) time τ, of the process $(T^n)_{n \geqslant 0}$, all defined on the probability space $(X, \mathcal{A}, \nu)$.

2.2. Distributional convergence

Let $(\mathfrak{E}, d_{\mathfrak{E}})$ be a compact metric space with Borel σ-algebra $\mathcal{B}_{\mathfrak{E}}$. As usual, a sequence $(Q_l)_{l \geqslant 1}$ of probability measures on $(\mathfrak{E}, \mathcal{B}_{\mathfrak{E}})$ is said to *converge weakly* to the probability measure Q on $(\mathfrak{E}, \mathcal{B}_{\mathfrak{E}})$, written $Q_l \Longrightarrow Q$, if the integrals of all continuous real functions χ on $\mathfrak{E}$ converge, $\int \chi \, dQ_l \longrightarrow \int \chi \, dQ$ as $l \to \infty$ for $\chi \in \mathcal{C}(\mathfrak{E})$. This is w*-convergence in $\mathfrak{M}(\mathfrak{E})$, the set of Borel probabilities on $\mathfrak{E}$, regarded as a subset of the space of all finite signed Borel measures on $\mathfrak{E}$ (equipped with the total variation norm), which by the Riesz representation theorem constitute the topological dual space of $\mathcal{C}(\mathfrak{E})$.

If R_l, $l \geqslant 1$, are measurable maps of $(X, \mathcal{A})$ into $(\mathfrak{E}, \mathcal{B}_{\mathfrak{E}})$, ν_l are probability measures on $(X, \mathcal{A})$, and R is another *random element* of $\mathfrak{E}$ (defined on some $(\Omega, \mathcal{F}, \Pr)$), then we write

$$R_l \overset{\nu_l}{\Longrightarrow} R \quad \text{as} \quad l \to \infty \tag{2.1}$$

to indicate that $\text{law}_{\nu_l}(R_l) := \nu_l \circ R_l^{-1} \Longrightarrow \text{law}(R) = \Pr \circ R^{-1}$. This is *distributional convergence* to R of the R_l when the latter functions are regarded as random variables on the probability spaces $(X, \mathcal{A}, \nu_l)$, respectively. This includes the case of a single measure ν, where $R_l \overset{\nu}{\Longrightarrow} R$ means that the distributions $\text{law}_\nu(R_l) = \nu \circ R_l^{-1}$ of the R_l under ν converge weakly to the law of R.

A sequence $R = (R^{(0)}, R^{(1)}, \ldots)$ of measurable functions $R^{(j)} : X \to \mathfrak{E}$ can be regarded as a single function into the (compact) sequence space $\mathfrak{E}^{\mathbb{N}_0} = \{(r^{(j)})_{j \geqslant 0} : r^{(j)} \in \mathfrak{E}\}$, equipped with the *product metric* $d_{\mathfrak{E}^{\mathbb{N}_0}}(q, r) := \sum_{j \geqslant 0} 2^{-(j+1)} d_{\mathfrak{E}}(q^{(j)}, r^{(j)})$.

Recall that weak convergence $Q_l \Longrightarrow Q$ in $\mathfrak{M}(\mathfrak{E}^{\mathbb{N}_0})$ of Borel probabilities on $\mathfrak{E}^{\mathbb{N}_0}$ is equivalent to convergence of all finite-dimensional marginals, $\pi^d_* Q_l \Longrightarrow \pi^d_* Q$ in $\mathfrak{M}(\mathfrak{E}^d)$ for $d \geqslant 1$, where $\pi^d : \mathfrak{E}^{\mathbb{N}_0} \to \mathfrak{E}^d$ denotes the canonical projection onto the first d factors.

3. When do orbits hit small sets?

3.1. Hitting-time and return-time processes

To accommodate normalized hitting times and their possible limits, we will use the (compact) target space $(\mathfrak{E}, d_{\mathfrak{E}}) = ([0,\infty], d_{[0,\infty]})$ with $d_{[0,\infty]}(s,t) := | e^{-s} - e^{-t} |$. The sequence space $\mathfrak{E}^{\mathbb{N}_0} = [0,\infty]^{\mathbb{N}_0}$ will be equipped with the corresponding product metric $d_{[0,\infty]^{\mathbb{N}_0}}$ as above.

We first study, for sets A as above, the random sequences of consecutive return- and hitting-times, that is, we are going to consider the sequences $\Phi_A : X \to [0,\infty]^{\mathbb{N}_0}$ of functions given by

$$\Phi_A := \left(\varphi_A, \varphi_A \circ T_A, \varphi_A \circ T_A^2, \ldots\right) \quad \text{on } X. \tag{3.1}$$

When regarded as a random sequence defined on $(X, \mathcal{A}, \nu)$, we shall call Φ_A the *hitting-time process of A under ν*. If no measure is mentioned, this means that $\nu = \mu$. In case we restrict Φ_A to A and view it through μ_A, we call it the *return-time process* of A. From the properties of T_A on $(A, \mathcal{A} \cap A, \mu_A)$ it is immediate that

(3.2) any return-time process Φ_A is stationary and ergodic (under μ_A),

and by relating return-time processes to hitting-time processes with different initial measures, stationarity often carries over to limits of the latter.

3.2. Asymptotic hitting-time and return-time processes for rare events

Assume now that $(A_l)_{l \geqslant 1}$ is a sequence of asymptotically rare events. It is immediate from (3.2) and Kac' formula that for any random sequence $\widetilde{\Phi} = (\widetilde{\varphi}^{(0)}, \widetilde{\varphi}^{(1)}, \ldots)$ in $[0,\infty]$,

$$\text{if} \quad \mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \widetilde{\Phi} \text{ as } l \to \infty, \quad \text{then } \widetilde{\Phi} \text{ is stationary with } \mathbb{E}\left[\widetilde{\varphi}^{(0)}\right] \leqslant 1. \tag{3.3}$$

Beyond that, little can be said about the general *asymptotic return-time process* $\widetilde{\Phi}$. In fact, it has been shown in [Zwe16] that every stationary sequence $\widetilde{\Phi}$ with $\mathbb{E}[\widetilde{\varphi}^{(0)}] \leqslant 1$ does appear as the limit for a suitable asymptotically rare sequence (A_l) if only T acts on a nonatomic space $(X, \mathcal{A}, \mu)$.

Turning to *asymptotic hitting-time processes* Φ, that is, distributional limits of hitting-time processes under one fixed probability $\nu \in \mathfrak{P}$, we first recall that these do not depend on the particular choice of ν. (The following is [Zwe07b, Corollary 6].)

Proposition 3.1 (Strong distributional convergence of $\mu(A_l)\Phi_{A_l}$). — *Suppose that $(X, \mathcal{A}, \mu, T)$ is an ergodic probability preserving system, and (A_l) an asymptotically rare sequence in $\mathcal{A}$. Let Φ be any random sequence in $[0,\infty]$. Then*

$$\mu(A_l)\Phi_{A_l} \overset{\nu}{\Longrightarrow} \Phi \text{ for some } \nu \in \mathfrak{P} \quad \text{iff} \quad \mu(A_l)\Phi_{A_l} \overset{\nu}{\Longrightarrow} \Phi \text{ for all } \nu \in \mathfrak{P}. \tag{3.4}$$

Despite this, even if both exist, the asymptotic hitting-time process Φ for a given sequence (A_l) need not coincide with the asymptotic return-time process $\widetilde{\Phi}$ for that very sequence. Indeed, the relation between the two types of limit processes will be of central importance in what follows.

3.3. Relating limit processes under μ_{A_l} to limit processes under μ.

It is well known that for any asymptotically rare sequence (A_l), limit laws for the normalized first return-times, $\mu(A_l)\varphi_{A_l}$ under μ_{A_l}, are intimately related to limit laws of the normalized first hitting-times, $\mu(A_l)\varphi_{A_l}$ under μ (see [AS11, HLV05]), and that this leads to an efficient way of proving convergence (of both) to an exponential law. In [Zwe16] we have extended the crucial duality to processes $\mu(A_l)\Phi_{A_l}$, see Section 7 below for more details.

A key ingredient of our present approach is the following generalization of Proposition 3.1 which provides conditions under which the processes $\mu(A_l)\Phi_{A_l}$, when started with suitable measures ν_l, exhibit the same asymptotic distributional behaviour as hitting time processes started with μ. The assumptions on the delay times τ_l below are those already mentioned in the introduction, but we now allow non-constant τ_l. Parallel to (2.1) we write, for measurable functions $R_l, R : X \to \mathfrak{E}$,

$$R_l \overset{\nu_l}{\longrightarrow} R \quad \text{as} \quad l \to \infty \tag{3.5}$$

provided that $\nu_l(d_{\mathfrak{E}}(R_l, R) > \varepsilon) \to 0$ as $l \to \infty$ whenever $\varepsilon > 0$. This includes the case of a single measure, $\nu_l = \nu$, in which case $R_l \overset{\nu}{\longrightarrow} R$ is the usual convergence in measure, $\nu(d_{\mathfrak{E}}(R_l, R) > \varepsilon) \to 0$ for $\varepsilon > 0$. In the results to follow, compact subsets $\mathfrak{K}$ of $(\mathfrak{P}, d_{\mathfrak{P}})$ play the role of families of measures which only differ from μ in a controllable way, as one can always assume w.l.o.g. that $\mu \in \mathfrak{K}$.

Theorem 3.2 (Asymptotic hitting-time process - ν_l versus μ). — *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system, $(A_l)_{l\geqslant 1}$ a sequence of asymptotically rare events, and (ν_l) a sequence in $\mathfrak{P}$. Assume that there are measurable functions $\tau_l : X \to \mathbb{N}_0$ such that*

$$\mu(A_l)\,\tau_l \overset{\nu_l}{\longrightarrow} 0 \quad \text{as} \quad l \to \infty, \tag{3.6}$$

and

$$\nu_l\,(\tau_l < \varphi_{A_l}) \longrightarrow 1 \quad \text{as} \quad l \to \infty, \tag{3.7}$$

while there is some compact subset $\mathfrak{K}$ of $(\mathfrak{P}, d_{\mathfrak{P}})$ such that

$$T_*^{\tau_l}\nu_l \in \mathfrak{K} \quad \text{for} \quad l \geqslant 1, \tag{3.8}$$

Then, for any random element Φ of $[0,\infty]^{\mathbb{N}_0}$,

$$\mu(A_l)\Phi_{A_l} \overset{\nu_l}{\Longrightarrow} \Phi \quad \text{iff} \quad \mu(A_l)\Phi_{A_l} \overset{\mu}{\Longrightarrow} \Phi \quad \text{as} \quad l \to \infty. \tag{3.9}$$

Remark 3.3. — For a constant sequence $(\nu_l) = (\nu)$ we can take $\tau_l := 0$ for all l, and obtain Proposition 3.1. Given any sequence (ν_l) in $\mathfrak{P}$, the $\tau_l = 0$ case of the theorem shows that (3.9) holds whenever all the ν_l belong to some compact subset $\mathfrak{K}$ of $\mathfrak{P}$.

3.4. Convergence to iid exponential limit processes

In the most prominent case the limit process is an iid sequence of normalized exponentially distributed random variables, henceforth denoted by Φ_{Exp}. This is the process of interarrival times of an elementary standard Poisson (counting) process, and we shall say that (A_l) *exhibits Poisson asymptotics* if

$$\mu(A_l)\Phi_{A_l} \overset{\mu}{\Longrightarrow} \Phi_{\mathrm{Exp}} \quad \text{and} \quad \mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \Phi_{\mathrm{Exp}}. \tag{3.10}$$

This terminology is further justified by

Remark 3.4 (Convergence of associated counting processes). — Any set $A \in \mathcal{A}$ of positive measure comes with an associated *normalized counting process* $\mathrm{N}_A : X \to \mathcal{D}[0,\infty)$ (the collection of cadlag paths $\mathrm{x} = (\mathrm{x}_t)_{t\geqslant 0} : [0,\infty) \to \mathbb{R}$) given by $\mathrm{N}_A = (\mathrm{N}_{A,t})_{t\geqslant 0}$ with

$$\mathrm{N}_{A,t} := \sum_{k=1}^{\lfloor t/\mu(A)\rfloor} 1_A \circ T^k = \sum_{j\geqslant 1} 1_{\left[\mu(A)\sum_{i=0}^{j-1}\varphi_A\circ T_A^i,\infty\right)}(t). \tag{3.11}$$

Observe then that for any probability measures ν_l on $(X,\mathcal{A})$,

$$\mu(A_l)\Phi_{A_l} \overset{\nu_l}{\Longrightarrow} \Phi_{\mathrm{Exp}} \quad \text{implies} \quad \mathrm{N}_{A_l} \overset{\nu_l}{\Longrightarrow} \mathrm{N} \text{ in } (\mathcal{D}[0,\infty),\mathcal{J}_1)\,, \tag{3.12}$$

where $\mathrm{N} = (\mathrm{N}_t)_{t\geqslant 0}$ denotes a standard Poisson counting process. To see this, define $\mathcal{W} := \{(\varphi^{(i)}) \in [0,\infty)^{\mathbb{N}_0} : \sum_{i=0}^{j-1}\varphi^{(i)} \nearrow \infty \text{ as } j\to\infty\}$, and $\Upsilon : [0,\infty]^{\mathbb{N}_0} \to \mathcal{D}[0,\infty)$ by

$$\Upsilon\left(\left(\varphi^{(i)}\right)_{i\geqslant 0}\right) := \sum_{j\geqslant 1} 1_{\left[\sum_{i=0}^{j-1}\varphi^{(i)},\infty\right)}$$

for $(\varphi^{(i)})_{i\geqslant 0} \in \mathcal{W}$, and $\Upsilon((\varphi^{(i)})_{i\geqslant 0}) := 0$ otherwise. An elementary argument shows that Υ is continuous on $\mathcal{W}_+ := \mathcal{W} \cap (0,\infty)^{\mathbb{N}_0}$ for the Skorokhod $\mathcal{J}_1$-topology of $\mathcal{D}[0,\infty)$ (see [Bil99, Whi02]). Since $\Upsilon(\Phi_{\mathrm{Exp}})$ has the same law as N, and $\Phi_{\mathrm{Exp}} \in \mathcal{W}_+$ almost surely, (3.12) thus follows by the " extended continuous mapping theorem" [Bil99, Theorem 2.7].

Remark 3.5 (Convergence of associated point processes). — Alternatively, consider the associated *normalized point process of visiting times*, $\mathcal{N}_A : X \to \mathcal{M}_p[0,\infty)$ (the collection of Radon point measures $\mathrm{n} : \mathcal{B}_{[0,\infty)} \to \{0,1,2,\ldots,\infty\}$, equipped with the topology of vague convergence, see [Res08, Chapter 3]) given by

$$\mathcal{N}_A := \sum_{j\geqslant 1} \delta_{\left\{\mu(A)\sum_{i=0}^{j-1}\varphi_A\circ T_A^i\right\}}, \tag{3.13}$$

where $\delta_{\{s\}}$ denotes the unit point mass at s. That is, $\mathcal{N}_A(x)$ has distribution function $t \mapsto \mathrm{N}_{A,t}(x)$. Then, for arbitrary probability measures ν_l on $(X, \mathcal{A})$,

$$\mu(A_l)\Phi_{A_l} \overset{\nu_l}{\Longrightarrow} \Phi_{\mathrm{Exp}} \quad \text{implies} \quad \mathcal{N}_{A_l} \overset{\nu_l}{\Longrightarrow} \mathrm{PRM}\left(\lambda^1_{[0,\infty)}\right) \text{ in } \mathcal{M}_p[0,\infty), \tag{3.14}$$

with $\mathrm{PRM}(\lambda^1_{[0,\infty)})$ denoting the Poisson random measure of intensity λ^1 on $[0,\infty)$. This, too, follows via the " extended continuous mapping theorem" since it is easily seen that the map

$$\Theta : [0,\infty]^{\mathbb{N}_0} \to \mathcal{M}_p[0,\infty) \quad \text{with} \quad \Theta\left(\left(\varphi^{(i)}\right)_{i\geqslant 0}\right) := \sum_{j\geqslant 1} \delta_{\left\{\sum_{i=0}^{j-1} \varphi^{(i)}\right\}}$$

for $(\varphi^{(i)})_{i\geqslant 0} \in \mathcal{W}$, and $\Theta((\varphi^{(i)})_{i\geqslant 0}) := 0$ otherwise, is continuous on $\mathcal{W}$, while $\Theta(\Phi_{\mathrm{Exp}})$ has the same law as $\mathrm{PRM}(\lambda^1_{[0,\infty)})$.

We will show that this happens in the $\nu_l = \mu_{A_l}$ case of the previous theorem:

Theorem 3.6 (Convergence to an iid exponential sequence). — *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system, and $(A_l)_{l\geqslant 1}$ a sequence of asymptotically rare events. Let $\nu_l := \mu_{A_l}$ and assume that there are measurable functions $\tau_l : X \to \mathbb{N}_0$ such that*

$$\mu(A_l)\,\tau_l \overset{\nu_l}{\longrightarrow} 0 \quad \text{as} \quad l \to \infty, \tag{3.15}$$

and

$$\nu_l\left(\tau_l < \varphi_{A_l}\right) \longrightarrow 1 \quad \text{as} \quad l \to \infty, \tag{3.16}$$

while there is some compact subset $\mathfrak{K}$ of $(\mathfrak{P}, d_{\mathfrak{P}})$ such that

$$T_*^{\tau_l}\nu_l \in \mathfrak{K} \quad \text{for} \quad l \geqslant 1, \tag{3.17}$$

Then (A_l) exhibits Poisson asymptotics,

$$\mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \Phi_{\mathrm{Exp}} \quad \text{and} \quad \mu(A_l)\Phi_{A_l} \overset{\mu}{\Longrightarrow} \Phi_{\mathrm{Exp}} \quad \text{as} \quad l \to \infty. \tag{3.18}$$

More generally, (3.18) *follows provided that for every $\varepsilon > 0$ there are a sequence $(\nu_l)_{l\geqslant 1}$ in $\mathfrak{P}$ with $d_{\mathfrak{P}}(\nu_l, \mu_{A_l}) < \varepsilon$ for all l, a compact set $\mathfrak{K} \subseteq \mathfrak{P}$, and measurable functions $\tau_l : X \to \mathbb{N}_0$ such that* (3.15)-(3.17) *hold.*

We shall see that this opens up a very easy way of proving Poisson asymptotics in several interesting situations.

Remark 3.7. —

(a) An obvious necessary condition for the first component $\mu(A_l)\,\varphi_{A_l}$ to have an exponential limit law is the absence of a point mass at zero in the limit. That is, for any sequence (τ_l) satisfying condition (3.15) we need to have (3.16).
(b) Note that condition (3.17) only uses information on what happens to A_l before the time τ_l which is of order $o(1/\mu(A_l))$ as $l \to \infty$. We do not assume mixing (let alone any quantitative mixing conditions), but only use what little asymptotic information follows from ergodicity alone.
(c) The assumptions (3.15)-(3.17) make precise the condition that the system should forget, sufficiently fast, whether or not it started in A_l. We can regard (3.15)-(3.17) as a short-time decorrelation (or mixing) condition.

(d) Of course, mixing properties of specific systems can still be very useful for validating conditions (3.15)-(3.17). However, in our discussion of examples in Section 10, we make a point of not using any asymptotic mixing properties for this purpose.
(e) Allowing measures ν_l more general than μ_{A_l} in the final statement of the theorem sometimes enables us to replace the density $\mu(A_l)^{-1}1_{A_l}$ of μ_{A_l} by an approximating density of higher regularity for which (3.15)-(3.17) are easier to verify (e.g. if they belong to a space on which the transfer operator is well understood).
(f) Another way of using this flexibility is to replace μ_{A_l} by $\mu_{A'_l}$ for nicer sets $A'_l \in \mathcal{A}$. This works if, for every $\widetilde{\varepsilon} > 0$, one can pick a sequence (A'_l) such that $\mu_{A_l}(A_l \triangle A'_l) < \widetilde{\varepsilon}$ for all $l \geqslant 1$ while the $\nu_l := \mu_{A'_l}$ admit $\mathfrak{K}$ and τ_l satisfying (3.15)-(3.17).

3.5. Allowing immediate returns

While for many important classes of concrete dynamical systems one typically observes Poisson asymptotics for natural families of rare events (cylinders or general ε-balls shrinking to a typical point x^*), there are often distinguished exceptional points x^*, like the periodic points of the system, to which a positive proportion $1-\theta$ with $\theta \in (0,1)$ of a neighbourhood can return after a fixed number of steps. This will result in a point mass at zero in the limit of return time distributions. If the situation is nice otherwise, the part which did escape in the first step may return after a rescaled exponential time, so that

$$\mu(A_l)\varphi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \widetilde{\varphi},$$

where the limit variable $\widetilde{\varphi}$ is characterized by the distribution function

$$\widetilde{F}_{(\mathrm{Exp},\theta)}(t) := (1-\theta) + \theta\left(1 - e^{-\theta t}\right), \quad t \geqslant 0. \tag{3.19}$$

Turning to processes, for any $\theta \in (0,1)$ we let $\widetilde{\Phi}_{(\mathrm{Exp},\theta)}$ denote an iid sequence of random variables, each distributed according to $\widetilde{F}_{(\mathrm{Exp},\theta)}$.

The following complement to Theorem 3.6 covers such situations. Here and later, when given a sequence (ν_l) of probabilities on $(X,\mathcal{A})$ and some $B \in \mathcal{A}$, we shall simply write $\nu_{l,B}$ for the normalized restriction given by $\nu_{l,B}(A) := \nu_l(B)^{-1}\nu_l(B \cap A)$, $A \in \mathcal{A}$.

Theorem 3.8 (Convergence to an iid $\widetilde{F}_{(\mathrm{Exp},\theta)}$ sequence). — *Let $(X,\mathcal{A},\mu,T)$ be an ergodic probability preserving system, $(A_l)_{l\geqslant 1}$ a sequence of asymptotically rare events. Let $\nu_l := \mu_{A_l}$ and suppose that $A_l = A_l^\bullet \cup A_l^\circ$ (disjoint) with*

$$\nu_l(A_l^\circ) \longrightarrow \theta \in (0,1) \quad \textit{as} \quad l \to \infty. \tag{3.20}$$

Set $\nu_l^\circ := \nu_{l,A_l^\circ}$ and $\nu_l^\bullet := \nu_{l,A_l^\bullet}$ (the normalized restrictions of ν_l to A_l° and $A_l^\bullet$, respectively). Assume further that there are measurable functions $\tau_l : X \to \mathbb{N}_0$, $l \geqslant 1$, such that

$$\mu(A_l)\,\tau_l \overset{\nu_l}{\longrightarrow} 0 \quad \textit{as} \quad l \to \infty, \tag{3.21}$$

and

$$\nu_l^\circ\,(\tau_l < \varphi_{A_l}) \longrightarrow 1 \quad \text{as} \quad l \to \infty, \tag{3.22}$$

while there is some compact subset $\mathfrak{K}$ *of* $(\mathfrak{P}, d_{\mathfrak{P}})$ *such that*

$$T_*^{\tau_l}\nu_l^\circ \in \mathfrak{K} \quad \text{for} \quad l \geqslant 1, \tag{3.23}$$

whereas

$$\nu_l^\bullet\,(\tau_l \geqslant \varphi_{A_l}) \longrightarrow 1 \quad \text{as} \quad l \to \infty, \tag{3.24}$$

and

$$d_{\mathfrak{P}}\left((T_{A_l})_*\,\nu_l^\bullet, \nu_l\right) \longrightarrow 0 \quad \text{as} \quad l \to \infty. \tag{3.25}$$

Then,

$$\mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \widetilde{\Phi}_{(\mathrm{Exp},\theta)} \quad \text{as} \quad l \to \infty. \tag{3.26}$$

More generally, (3.26) *follows provided that for every* $\varepsilon > 0$ *there are a sequence* $(\nu_l)_{l\geqslant 1}$ *in* $\mathfrak{P}$ *with* $d_{\mathfrak{P}}(\nu_l, \mu_{A_l}) < \varepsilon$ *for all* l*, a compact set* $\mathfrak{K} \subseteq \mathfrak{P}$*, and measurable functions* $\tau_l : X \to \mathbb{N}_0$ *such that* (3.21)-(3.25) *hold.*

In Section 10 we illustrate how this can be used very easily in some standard situations.

4. Where do orbits hit small sets?

4.1. Local observables and local processes

Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system. We introduce a large class of random processes associated to the visits of orbits to a given small set $A \in \mathcal{A}$. The idea is to focus on what exactly happens upon each visit, and record the position inside A by means of some function ψ_A on this set. As we are interested in small sets and, ultimately, in limits as the size of the sets tends to zero, it is natural to consider functions encoding the *relative* position inside A, thus effectively rescaling the set.

For instance, if the relevant sets A are subintervals of some larger interval X, the *normalizing interval charts* $\psi_A : A \to [0,1]$ with $\psi_A(x) := (x-a)/(b-a)$ for $A = [a,b]$, are a natural choice. This is the prototypical example to keep in mind, but for the general theory we simply allow measurable maps $\psi_A : A \to \mathfrak{Z}$, not necessarily invertible, into some space $\mathfrak{Z}$ which does not depend on A.

In the following we fix some compact metric space $(\mathfrak{Z}, d_{\mathfrak{Z}})$ with Borel σ-algebra $\mathcal{B}_{\mathfrak{Z}}$ to represent the relative positions (or some other relevant aspect) of points inside the distinguished small sets we wish to study. Any $\mathcal{A}$-$\mathcal{B}_{\mathfrak{Z}}$-measurable map $\psi_A : A \to \mathfrak{Z}$ will be called an ($\mathfrak{Z}$-valued) *local observable* on A, and we shall use the uppercase Ψ_A to denote the sequence of consecutive local observations of an orbit which starts anywhere in X,

$$\Psi_A : X \to \mathfrak{Z}^{\mathbb{N}}, \quad \Psi_A := \left(\psi_A \circ T_A, \psi_A \circ T_A^2, \ldots\right). \tag{4.1}$$

Equip $\mathfrak{Z}^{\mathbb{N}}$ with its compact Polish product topology, induced by the product metric $d_{\mathfrak{Z}^{\mathbb{N}}}$, then Ψ_A is $\mathcal{A}$-$\mathcal{B}_{\mathfrak{Z}^{\mathbb{N}}}$-measurable (and can thus be regarded as a single $\mathfrak{Z}^{\mathbb{N}}$-valued local observable on A). We can include the local observable at time zero provided that the orbit starts in A. To this end, define

$$\widetilde{\Psi}_A : A \to \mathfrak{Z}^{\mathbb{N}_0}, \quad \widetilde{\Psi}_A := \left(\psi_A, \psi_A \circ T_A, \psi_A \circ T_A^2, \ldots\right). \tag{4.2}$$

Given any probability measure ν on $(X, \mathcal{A})$, we can view Ψ_A as a random process on the probability space $(X, \mathcal{A}, \nu)$. If ν is concentrated on A, the same is true for $\widetilde{\Psi}_A$. We shall refer to either variant as a *local process under* ν. Again, the properties of T_A on $(A, \mathcal{A} \cap A, \mu_A)$ entail that

$$\text{any local process } \widetilde{\Psi}_A \text{ is stationary and ergodic under } \mu_A, \tag{4.3}$$

and we shall exploit this by linking local processes with different initial measures to this particular version. By (4.3), $\mathrm{law}_{\mu_A}(\widetilde{\Psi}_A) = \mathrm{law}_{\mu_A}(\Psi_A)$, so that we need not distinguish between the two variants as far as the laws of individual processes under the corresponding measures μ_A are concerned.

Remark 4.1 (Normalized return times and processes as local observables and processes). — The normalized return times $\mu(A)\varphi_A : A \to [0, \infty]$ studied in the previous section also define local observables, somewhat special in that they are defined in terms of the dynamics. Moreover, the processes $\mu(A)\Phi_A : A \to [0, \infty]^{\mathbb{N}_0}$ are particular local processes: using the notation introduced above, we have $\mu(A)\Phi_A = \widetilde{\Psi}_A$ for $\psi_A := \mu(A)\varphi_A$.

Remark 4.2 (Warning regarding normalized hitting processes). — In contrast, the normalized hitting time process $\mu(A)\Phi_A : X \to [0, \infty]^{\mathbb{N}_0}$ with Φ_A as defined in (3.1) is *not* a local process Ψ_A as in (4.1). Indeed, Φ_A is not constant on $\{x, Tx, \ldots, T^{\varphi_A(x)-1}x\}$ for $x \in A^c$. This is why we do not need an analogue of condition (3.6) in Theorem 4.7 below.

4.2. Local processes for asymptotically rare events

Assume now that $(A_l)_{l \geqslant 1}$ is a sequence of asymptotically rare events, and that for each A_l we are given a local observable $\psi_{A_l} : A \to \mathfrak{Z}$. Our goal is to provide useful conditions under which the sequence of local processes $(\widetilde{\Psi}_{A_l})_{l \geqslant 1}$ or $(\Psi_{A_l})_{l \geqslant 1}$ converges in distribution as $l \to \infty$. Here, again, it makes sense to study these random variables either through one fixed initial probability ν (in case of $(\Psi_{A_l})_{l \geqslant 1}$), say $\nu = \mu$, or to view them through the sequence (μ_{A_l}) of normalized restrictions to these sets.

We first look at the $\widetilde{\Psi}_{A_l}$ under the measures μ_{A_l}. Due to (4.3) we see that for any random sequence $\widetilde{\Psi}$ in $\mathfrak{Z}$,

$$\text{if} \quad \widetilde{\Psi}_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \widetilde{\Psi} \quad \text{as} \quad l \to \infty, \quad \text{then } \widetilde{\Psi} \text{ is stationary.} \tag{4.4}$$

Beyond that, little can be said about the general *asymptotic local process* $\widetilde{\Psi}$. In fact, we are going to show that unless the system acts on a discrete space (and hence is essentially a cyclic permutation), every $\mathfrak{Z}$-valued stationary sequence arises as the limit of local processes for any given sequence (A_l) if only we use suitable local observables ψ_{A_l}. (This is parallel to [Zwe16, Theorem 2.1].) In particular, $\widetilde{\Psi}$ need not be independent, and doesn't even have to be ergodic.

Theorem 4.3 (Prescribing the asymptotic internal state process). — *Let T be an ergodic measure preserving map on the nonatomic probability space $(X, \mathcal{A}, \mu)$, let (A_l) be an asymptotically rare sequence in $\mathcal{A}$, and let $\widetilde{\Psi}$ be any $\mathfrak{Z}$-valued stationary sequence. Then there is a sequence (ψ_{A_l}) of local observables for the A_l such that*

$$\widetilde{\Psi}_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \widetilde{\Psi} \quad \text{as} \quad l \to \infty. \tag{4.5}$$

Observe that the distributions $\mathrm{law}_{\mu_{A_l}}(\psi_{A_l})$ of the first components of the $\widetilde{\Psi}_{A_l}$ may not involve any dynamics, but their convergence is of course necessary for convergence of the processes as in (4.5). For the abstract theory we will therefore take the assumption

$$\psi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \psi \quad \text{as} \quad l \to \infty, \tag{4.6}$$

that they converge to the law of some particular random element ψ of $\mathfrak{Z}$, as our starting point. (Note that this is not particularly restrictive. By compactness of $\mathfrak{M}(\mathfrak{Z})$, every sequence contains a subsequence along which (4.6) is satisfied.) The question will then be under what conditions (4.6) entails convergence of the processes Ψ_{A_l} to some (or some particular) random sequence Ψ.

In some natural situations, condition (4.6) relates the local observables ψ_{A_l} to the local regularity of μ on the A_l:

Example 4.4 (Normalizing interval charts). — Assume that X is an interval, and μ is absolutely continuous w.r.t. Lebesgue measure λ. Suppose that the A_l are subintervals which shrink to a distinguished point $x^* \in X$ at which (a suitable version of) the density $d\mu/d\lambda$ is continuous and strictly positive. Define $\psi_{A_l} : A_l \to [0, 1]$ to be the normalizing interval chart from the previous subsection. Then (4.6) holds, with the limit variable ψ uniformly distributed on $\mathfrak{Z} = [0, 1]$.

Considering a sequence of local processes under one probability measure ν which doesn't depend on l, we usually lose stationarity, but gain the possibility of freely switching measures.

Proposition 4.5 (Strong distributional convergence of Ψ_{A_l}). — *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system, (A_l) an asymptotically rare sequence in $\mathcal{A}$, with $(\Psi_{A_l})_{l \geqslant 1}$ a sequence of $\mathfrak{Z}$-valued local processes for the A_l. Then*

$$\Psi_{A_l} \overset{\nu}{\Longrightarrow} \Psi \ \textit{for}\ \text{some}\ \nu \in \mathfrak{P} \qquad \textit{iff} \qquad \Psi_{A_l} \overset{\nu}{\Longrightarrow} \Psi \ \textit{for}\ \text{all}\ \nu \in \mathfrak{P}. \tag{4.7}$$

This is an immediate consequence of [Zwe07b], see the start of Section 8 for details. Variations on this theme will be the key to the limit theorems below.

4.3. Relating limit processes under μ_{A_l} to limit processes under μ

As mentioned before, the intimate relation between return- and hitting times, that is, the relation between the laws of φ_A under μ_A and μ respectively, is often crucial for the analysis of these variables. For general local observables there is no such principle:

(4.8) In general, $\psi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \psi$ does not imply convergence of $\psi_{A_l} \circ T_{A_l}$ under μ.

Example 4.6. — Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system, (A'_l) an asymptotically rare sequence in $\mathcal{A}$ such that $A'_l \cap T^{-1}A'_l = \varnothing$ for $l \geqslant 1$. Set

$$A_l := T^{-1}A'_l \cup A'_l, \quad \text{and} \quad \psi_{A_l} : A_l \to \{0,1\}$$

with

$$\psi_{A_{2l-1}} := 1_{A'_l} \quad \text{while} \quad \psi_{A_{2l}} := 1_{T^{-1}A'_l}.$$

Then $\psi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \psi$, with ψ denoting a fair coin, while $\psi_{A_{2l-1}} \circ T_{A_l} \overset{\mu}{\Longrightarrow} 0$ and $\psi_{A_{2l}} \circ T_{A_l} \overset{\mu}{\Longrightarrow} 1$ since $T_{A_l} = T_{T^{-1}A'_l}$ outside $T^{-1}A'_l$.

Nonetheless, we can provide a very useful condition which ensures that possibly localized measures ν_l can be replaced by any fixed probability $\nu \in \mathfrak{P}$.

Theorem 4.7 (Asymptotic local process - ν_l versus μ). — *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system, (A_l) an asymptotically rare sequence in $\mathcal{A}$, and $(\psi_{A_l})_{l \geqslant 1}$ a sequence of $\mathfrak{Z}$-valued local observables for the A_l, with corresponding local processes Ψ_{A_l}. Let (ν_l) be a sequence in $\mathfrak{P}$.*

Assume that there are measurable $\tau_l : X \to \mathbb{N}_0$ such that

$$\nu_l\left(\tau_l < \varphi_{A_l}\right) \longrightarrow 1 \quad \text{as} \quad l \to \infty, \tag{4.9}$$

while there is some compact subset $\mathfrak{K}$ of $(\mathfrak{P}, d_{\mathfrak{P}})$ such that

$$T_*^{\tau_l}\nu_l \in \mathfrak{K} \quad \text{for} \quad l \geqslant 1, \tag{4.10}$$

Then, for any random sequence Ψ in $\mathfrak{Z}$,

$$\Psi_{A_l} \overset{\nu_l}{\Longrightarrow} \Psi \quad \text{iff} \quad \Psi_{A_l} \overset{\mu}{\Longrightarrow} \Psi \quad \text{as} \quad l \to \infty. \tag{4.11}$$

Of course, the most important case will be that of $\nu_l = \mu_{A_l}$.

4.4. Convergence to iid limit processes

A variant of the above assumption in which we now take ν_l to be μ_{A_l} conditioned on suitable subsets of A_l actually allows us to prove (under the necessary assumption (4.6) discussed above) convergence of the local processes to an *independent* stationary sequence.

Theorem 4.8 (Convergence to an iid local process). — *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system, (A_l) an asymptotically rare sequence in $\mathcal{A}$, and $(\psi_{A_l})_{l \geqslant 1}$ a sequence of $\mathfrak{Z}$-valued local observables for the A_l such that*

$$\psi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \psi \quad \text{as} \quad l \to \infty, \tag{4.12}$$

for some random element ψ of $\mathfrak{Z}$. Let Ψ_{A_l} be the corresponding local processes. Suppose that $\mathcal{B}_{\mathfrak{Z}}^{\pi} \subseteq \mathcal{B}_{\mathfrak{Z}}$ is a π-system generating $\mathcal{B}_{\mathfrak{Z}}$ with $\mathfrak{Z} \in \mathcal{B}_{\mathfrak{Z}}^{\pi}$, while $\Pr[\psi \in \partial F] = 0$ *for all $F \in \mathcal{B}_{\mathfrak{Z}}^{\pi}$.*

Set $\nu_l := \mu_{A_l}$, *and assume further that for every* $F \in \mathcal{B}^\pi_\mathfrak{Z}$ *with* $\Pr[\psi \in F] > 0$ *there are measurable* $\tau_{l,F} : X \to \mathbb{N}_0$, $l \geqslant 1$, *such that, letting* $\nu_{l,F}$ *denote the normalized restriction* $\nu_{l,\{\psi_{A_l} \in F\}}$ *of* ν_l *to* $\{\psi_{A_l} \in F\}$, *we have*

$$\nu_{l,F}\,(\tau_l < \varphi_{A_l}) \longrightarrow 1 \quad \text{as} \quad l \to \infty, \tag{4.13}$$

while there is some compact subset $\mathfrak{K}_F$ *of* $(\mathfrak{P}, d_\mathfrak{P})$ *such that*

$$T_*^{\tau_{l,F}}(\nu_{l,F}) \in \mathfrak{K}_F \quad \text{for} \quad l \geqslant 1, \tag{4.14}$$

Then

$$\widetilde{\Psi}_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \Psi^* \quad \text{and} \quad \Psi_{A_l} \overset{\mu}{\Longrightarrow} \Psi^* \quad \text{as} \quad l \to \infty, \tag{4.15}$$

where $\Psi^* = (\psi^{*(j)})_{j \geqslant 0}$ *is an iid sequence in* $\mathfrak{Z}$ *with* $\text{law}(\psi^{*(0)}) = \text{law}(\psi)$.

More generally, (4.15) *follows if every* $\varepsilon > 0$ *there are a sequence* $(\nu_l)_{l \geqslant 1}$ *in* $\mathfrak{P}$ *with* $d_\mathfrak{P}(\nu_l, \mu_{A_l}) < \varepsilon$ *for all* l, *and, for every* $F \in \mathcal{B}^\pi_\mathfrak{Z}$ *with* $\Pr[\psi \in F] > 0$, *a compact set* $\mathfrak{K}_F \subseteq \mathfrak{P}$, *measures* $\nu_{l,F} \in \mathfrak{P}$ *with* $d_\mathfrak{P}(\nu_{l,F}, \nu_{l,\{\psi_{A_l} \in F\}}) \to 0$, *and* $\tau_{l,F}$ *such that* (4.13) *and* (4.14) *hold.*

Remark 4.9 (Using rich generating families of conditioning events). — Note that there are natural situations in which the condition $\Pr[\psi \in \partial F] = 0$ for all $F \in \mathcal{B}^\pi_\mathfrak{Z}$ need not be checked explicitly. For instance, if $\mathfrak{Z} = [a, b]$, then $\mathcal{G} := \{(s, b] : s \in [a, b)\}$ is a collection of sets with pairwise disjoint boundaries. Thus, the requirement $\Pr[\psi \in \partial F] = 0$ only rules out countably many $F \in \mathcal{G}$, and the remaining family $\mathcal{B}^\pi_\mathfrak{Z} := \{F \in \mathcal{G} : \Pr[\psi \in \partial F] = 0\}$ still is a generating π-system.

Therefore, if we can check the other conditions (4.13) and (4.14) for all $F \in \mathcal{G}$, then there is automatically a suitable collection $\mathcal{B}^\pi_\mathfrak{Z}$.

4.5. Robustness of the asymptotic behaviour

It will also be useful to know that asymptotic local processes do not change if the sets A_l are replaced by sets A'_l which are asymptotically equivalent (mod μ) in that $\mu(A_l \vartriangle A'_l) = o(\mu(A_l))$, and if the local observables ψ_{A_l} are replaced by $\psi'_{A'_l}$ which are close in measure.

Theorem 4.10 (Robustness of asymptotic local processes). — *Let* $(X, \mathcal{A}, \mu, T)$ *be an ergodic probability preserving system,* (A_l) *and* (A'_l) *two asymptotically rare sequences in* $\mathcal{A}$, *and* (ψ_{A_l}), $(\psi'_{A'_l})$ *two sequences of* $\mathfrak{Z}$*-valued local observables for the sets* A_l *and* A'_l, *respectively. Assume that*

$$\mu\,(A_l \vartriangle A'_l) = o\,(\mu(A_l)) \quad \text{as} \quad l \to \infty, \tag{4.16}$$

and

$$d_\mathfrak{Z}\left(\psi_{A_l}, \psi'_{A'_l}\right) \overset{\mu_{A_l \cap A'_l}}{\longrightarrow} 0 \quad \text{as} \quad l \to \infty. \tag{4.17}$$

Then, for any random sequence $\widetilde{\Psi}$ *in* $\mathfrak{Z}$,

$$\widetilde{\Psi}_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \widetilde{\Psi} \quad \text{iff} \quad \widetilde{\Psi}'_{A'_l} \overset{\mu_{A'_l}}{\Longrightarrow} \widetilde{\Psi}, \tag{4.18}$$

where $\widetilde{\Psi}_{A_l}$ *and* $\widetilde{\Psi}'_{A'_l}$ *are the local processes given by* ψ_{A_l} *and* $\psi'_{A'_l}$, *respectively.*

For basic situations in which this applies consider the following.

Example 4.11 (Normalizing interval charts). — Let X be an interval containing the sets $A_l = [a_l, b_l]$, $A'_l = [a'_l, b'_l]$, and let ψ_{A_l} and $\psi'_{A'_l}$ be the corresponding normalizing interval charts. It is immediate that $\sup_{A_l \cap A'_l} d_{[0,1]}(\psi_{A_l}, \psi'_{A'_l}) \to 0$ provided that $\lambda(A_l \triangle A'_l) = o(\lambda(A_l))$ as $l \to \infty$. In this case, assumptions (4.16) and (4.17) of the theorem are satisfied if $c\lambda \leqslant \mu \leqslant c^{-1}\lambda$ for some constant $c > 0$.

Example 4.12 (Normalized return times). — Let $(X, \mathcal{A}, \mu, T)$ be ergodic and probability preserving, (A_l), (A'_l) two asymptotically rare sequences, and consider the $[0,\infty]$-valued local observables $\psi_{A_l} := \mu(A_l)\varphi_{A_l} \mid_{A_l}$, $\psi'_{A'_l} := \mu(A'_l)\varphi_{A'_l} \mid_{A'_l}$. Then,

$$\mu\left(A_l \triangle A'_l\right) = o\left(\mu(A_l)\right) \quad \text{implies} \quad d_{[0,\infty]}\left(\psi_{A_l}, \psi'_{A'_l}\right) \overset{\mu_{A_l \cap A'_l}}{\longrightarrow} 0. \tag{4.19}$$

In particular, [Zwe16, Theorem 2.2] is a special case of Theorem 4.10 above.

(To check (4.19), assume w.l.o.g. that $A'_l \subseteq A_l$, take $\eta > 0$, and note that due to $d_{[0,\infty]}(s,t) \leqslant \mid s - t \mid$ it suffices to control the measure of $\{\mid \mu(A_l)\varphi_{A_l} - \mu(A'_l)\varphi_{A'_l} \mid \geqslant \eta\} \subseteq \{\mu(A'_l) \mid \varphi_{A_l} - \varphi_{A'_l} \mid \geqslant \eta/2\} \cup \{\mu(A_l \setminus A'_l)\, \varphi_{A_l} \geqslant \eta/2\}$ which is easy since $(A_l \cup A'_l) \cap \{\varphi_{A_l} \neq \varphi_{A'_l}\} = T^{-1}_{A_l \cup A'_l}(A_l \triangle A'_l)$ and $\int \mu(A'_l)\varphi_{A'_l} d\mu_{A'_l} = 1$ for all l.)

5. Joint limit processes

Again, the space for local observables will be a compact metric space $(\mathfrak{Z}, d_{\mathfrak{Z}})$. Given an asymptotically rare sequence (A_l) for $(X, \mathcal{A}, \mu, T)$ and local observables ψ_{A_l} we now consider the joint distribution of $\mu(A_l)\Phi_{A_l}$ and $\widetilde{\Psi}_{A_l}$ under μ_{A_l}, and that of $\mu(A_l)\Phi_{A_l}$ and Ψ_{A_l} under μ (or some other fixed probability $\nu \in \mathfrak{P}$). For the second variant we find, as expected:

Proposition 5.1 (Strong distributional convergence of $(\mu(A_l)\Phi_{A_l}, \Psi_{A_l})$). — *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system, (A_l) an asymptotically rare sequence in $\mathcal{A}$, with $(\Psi_{A_l})_{l \geqslant 1}$ a sequence of $\mathfrak{Z}$-valued local processes for the A_l. Let (Φ, Ψ) be any random sequence in $[0,\infty] \times \mathfrak{Z}$. Then $(\mu(A_l)\Phi_{A_l}, \Psi_{A_l}) \overset{\nu}{\Longrightarrow} (\Psi, \Phi)$ for* some *$\nu \in \mathfrak{P}$ iff $(\mu(A_l)\Phi_{A_l}\Psi_{A_l}) \overset{\nu}{\Longrightarrow} (\Psi, \Phi)$ for* all *$\nu \in \mathfrak{P}$.*

The main result of this section, Theorem 5.3 below, gives sufficient conditions for convergence to an independent pair of iid sequences. Before stating it, we record that, in certain situations, this takes place under the measure μ iff it takes place under the measures μ_{A_l}. Recall (4.8), which shows that the latter statement can only be correct under some extra condition. We will use the same assumption, (5.1) and (5.2) below, which already appeared in Theorem 4.7.

Theorem 5.2 (Independent joint limit processes - $\mu_{A_l}\backslash$ versus μ). — *Suppose $(X, \mathcal{A}, \mu, T)$ is an ergodic probability preserving system, (A_l) an asymptotically rare sequence in $\mathcal{A}$, and $(\psi_{A_l})_{l \geqslant 1}$ a sequence of $\mathfrak{Z}$-valued local observables for the A_l with corresponding local processes Ψ_{A_l}. Assume there are measurable $\tau_l : X \to \mathbb{N}_0$ s.t.*

$$\mu_{A_l}\left(\tau_l < \varphi_{A_l}\right) \longrightarrow 1 \quad as \quad l \to \infty, \tag{5.1}$$

while there is some compact subset $\mathfrak{K}$ of $(\mathfrak{P}, d_{\mathfrak{P}})$ such that

$$T_*^{\tau_l}\mu_{A_l} \in \mathfrak{K} \quad \textit{for} \quad l \geqslant 1, \tag{5.2}$$

Let $(\Phi_{\mathrm{Exp}}, \Psi^*)$ be an independent pair of iid sequences. Then

$$(\mu(A_l)\Phi_{A_l}, \Psi_{A_l}) \overset{\mu}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*) \quad \textit{as} \quad l \to \infty, \tag{5.3}$$

iff

$$\left(\mu(A_l)\Phi_{A_l}, \widetilde{\Psi}_{A_l}\right) \overset{\mu_{A_l}}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*) \quad \textit{as} \quad l \to \infty. \tag{5.4}$$

We can then formulate our abstract spatiotemporal Poisson limit theorem.

Theorem 5.3 (Joint iid limit processes). — *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system, (A_l) an asymptotically rare sequence in $\mathcal{A}$, and $(\psi_{A_l})_{l\geqslant 1}$ a sequence of $\mathfrak{Z}$-valued local observables for the A_l such that*

$$\psi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \psi \quad \textit{as} \quad l \to \infty, \tag{5.5}$$

for some random element ψ of $\mathfrak{Z}$. Let Ψ_{A_l} be the corresponding local processes. Moreover, assume that

(A) *for every $s \in [0, \infty)$ there are a compact subset $\mathfrak{K}_s$ of $(\mathfrak{P}, d_{\mathfrak{P}})$, measures $\nu_{l,s} \in \mathfrak{P}$ with $d_{\mathfrak{P}}(\nu_{l,s}, \mu_{A_l \cap \{\mu(A_l)\varphi_{A_l} > s\}}) \to 0$, and measurable $\tau_{l,s} : X \to \mathbb{N}_0$ s.t.*

$$\nu_{l,s}\left(\tau_{l,s} < \varphi_{A_l}\right) \longrightarrow 1 \quad \textit{as} \quad l \to \infty, \tag{5.6}$$

while $T_*^{\tau_{l,s}}\nu_{l,s} \in \mathfrak{K}_s$ for $l \geqslant 1$, and

(B) *there is some π-system $\mathcal{B}_{\mathfrak{Z}}^{\pi}$ generating $\mathcal{B}_{\mathfrak{Z}}$ with $\mathfrak{Z} \in \mathcal{B}_{\mathfrak{Z}}^{\pi}$, while $\Pr[\psi \in \partial F] = 0$ for all $F \in \mathcal{B}_{\mathfrak{Z}}^{\pi}$; in addition, for every $F \in \mathcal{B}_{\mathfrak{Z}}^{\pi}$ with $\Pr[\psi \in F] > 0$ there are a compact subset $\mathfrak{K}_F$ of $(\mathfrak{P}, d_{\mathfrak{P}})$, measures $\nu_{l,F} \in \mathfrak{P}$ with $d_{\mathfrak{P}}(\nu_{l,F}, \mu_{A_l \cap \{\psi_{A_l} \in F\}})$ $\to 0$, and measurable $\tau_{l,F} : X \to \mathbb{N}_0$ such that*

$$\mu(A_l)\,\tau_{l,F} \overset{\nu_{l,F}}{\longrightarrow} 0 \quad \textit{as} \quad l \to \infty, \tag{5.7}$$

and

$$\nu_{l,F}\left(\tau_{l,F} < \varphi_{A_l}\right) \longrightarrow 1 \quad \textit{as} \quad l \to \infty, \tag{5.8}$$

while $T_*^{\tau_{l,F}}\nu_{l,F} \in \mathfrak{K}_F$ for $l \geqslant 1$.

Then,

$$(\mu(A_l)\Phi_{A_l}, \Psi_{A_l}) \overset{\mu}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*) \quad \textit{as} \quad l \to \infty, \tag{5.9}$$

and

$$\left(\mu(A_l)\Phi_{A_l}, \widetilde{\Psi}_{A_l}\right) \overset{\mu_{A_l}}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*) \quad \textit{as} \quad l \to \infty, \tag{5.10}$$

where $(\Phi_{\mathrm{Exp}}, \Psi^*)$ is an independent pair of iid processes.

Remark 5.4 (Convergence of associated spatiotemporal point processes). — Generalizing Remark 3.5 we can interpret the above as a result about convergence of associated spatiotemporal point processes $\mathcal{N}_{A,\psi_A} : X \to \mathcal{M}_p([0,\infty)\times\mathfrak{Z})$ (the collection of Radon point measures $\mathrm{n} : \mathcal{B}_{[0,\infty)\times\mathfrak{Z}} \to \{0,1,2,\,\ldots,\,\infty\}$ with the topology of vague convergence, [Res08]) defined by

$$\mathcal{N}_{A,\psi_A} := \sum_{j\geqslant 1} \delta_{\left\{\left(\mu(A)\sum_{i=0}^{j-1}\varphi_A\circ T_A^i,\psi_A\circ T_A^j\right)\right\}}. \tag{5.11}$$

Parallel to our earlier remarks, for $(\Phi_{\mathrm{Exp}},\Psi^*)$ an independent pair of iid processes and arbitrary probabilities ν_l on $(X,\mathcal{A})$,

$$\begin{aligned} (\mu(A_l)\Phi_{A_l},\Psi_{A_l}) &\overset{\nu_l}{\Longrightarrow} (\Phi_{\mathrm{Exp}},\Psi^*) \quad \text{implies} \\ \mathcal{N}_{A_l,\psi_{A_l}} &\overset{\nu_l}{\Longrightarrow} \mathrm{PRM}(\lambda^1_{[0,\infty)}\otimes \mathrm{law}(\psi)) \quad \text{in} \quad \mathcal{M}_p([0,\infty)\times\mathfrak{Z}), \end{aligned} \tag{5.12}$$

and Ψ_{A_l} may be replaced by $\widetilde{\Psi}_{A_l}$ if for each $l\geqslant 1$, ν_l is supported on A_l. As before, this merely requires an application of the "extended continuous mapping theorem" ([Bil99, Theorem 2.7]), because the map $\Theta^\times : [0,\infty]^{\mathbb{N}_0}\times\mathfrak{Z}^{\mathbb{N}} \to \mathcal{M}_p([0,\infty)\times\mathfrak{Z})$ with

$$\Theta^\times\left(\left(\varphi^{(i)}\right)_{i\geqslant 0},\left(\psi^{(i)}\right)_{i\geqslant 1}\right) := \sum_{j\geqslant 1}\delta_{\left\{\left(\sum_{i=0}^{j-1}\varphi^{(i)},\psi^{(j)}\right)\right\}}$$

for $(\varphi^{(i)})_{i\geqslant 0}\in\mathcal{W}$, and $\Theta^\times((\varphi^{(i)})_{i\geqslant 0},(\psi^{(i)})_{i\geqslant 1}) := 0$ otherwise, is continuous on $\mathcal{W}\times\mathfrak{Z}^{\mathbb{N}}$.

Remark 5.5 (Robustness of joint limit processes). — The conclusion in (5.10) is a statement about the $(\mu(A_l)\Phi_{A_l},\widetilde{\Psi}_{A_l})$ which can be regarded as local processes taking values in $[0,\infty]\times\mathfrak{Z}$. Recalling Example 4.12 we can therefore use Theorem 4.10 to see that (5.10) is equivalent to

$$\left(\mu\left(A_l'\right)\Phi_{A_l'},\widetilde{\Psi'}_{A_l'}\right) \overset{\mu_{A_l'}}{\Longrightarrow} (\Phi_{\mathrm{Exp}},\Psi^*) \quad \text{as} \quad l\to\infty,$$

whenever $\mu(A_l\,\triangle\,A_l') = o(\mu(A_l))$ and $d_{\mathfrak{Z}}(\psi_{A_l},\psi'_{A_l'}) \overset{\mu_{A_l\cap A_l'}}{\longrightarrow} 0$. This sometimes allows us to replace the original sequence by one for which the conditions of the present theorem can be verified more easily.

6. Mean ergodic theory and distributions under varying measures

The present section discusses the abstract core of our approach. (It is based on ideas which arose in the study of probabilistic properties of infinite measure preserving systems, see [PSZ17, RZ20, Zwe07a].) Throughout, $(\mathfrak{E},d_{\mathfrak{E}})$ is a compact metric space with Borel σ-algebra $\mathcal{B}_{\mathfrak{E}}$. For a Lipschitz function $\varkappa:\mathfrak{E}\to\mathbb{R}$ we let $\mathrm{Lip}(\varkappa)$ denote its (optimal) Lipschitz constant.

6.1. More on distributional convergence

Recall that weak convergence of probabilities on $(\mathfrak{E},\mathcal{B}_{\mathfrak{E}})$ is metrisable: There are various metrics $D_{\mathfrak{E}}$ on $\mathfrak{M}(\mathfrak{E})$ such that $D_{\mathfrak{E}}(Q_l,Q)\to 0$ iff $Q_l\Longrightarrow Q$. For instance, it is well known that $\mathcal{C}(\mathfrak{E})$ is separable, and by a standard argument for w*-topologies, every dense sequence $(\vartheta_j)_{j\geqslant 1}$ in $\mathcal{C}(\mathfrak{E})$ allows us to define a suitable metric with $D_{\mathfrak{E}}\leqslant 1$ by setting

$$D_{\mathfrak{E}}(Q,Q'):=\sum_{j\geqslant 1}2^{-(j+1)}\left|\int\chi_j\,dQ-\int\chi_j\,dQ'\right|,\quad Q,Q'\in\mathfrak{M}(\mathfrak{E}),\tag{6.1}$$

where $\chi_j:=c_j^{-1}\vartheta_j$ for constants $c_j\geqslant\sup|\vartheta_j|$ (so that $|\chi_j|\leqslant 1$). By the Stone-Weierstrass theorem, Lipschitz functions are dense in $\mathcal{C}(\mathfrak{E})$. As a consequence, we can *define* $D_{\mathfrak{E}}$ *as in* (6.1) *using a particular dense sequence* (ϑ_j)*, henceforth fixed, of Lipschitz functions* $\vartheta_j:\mathfrak{E}\to\mathbb{R}$, and $c_j:=\max(\sup|\vartheta_j|,\mathrm{Lip}(\vartheta_j))$, which ensures that $\mathrm{Lip}(\varkappa_j)\leqslant 1$ for all $j\geqslant 1$.

Observe that for any convex combinations $Q=\theta Q_\triangle+(1-\theta)Q_\triangledown$ and $Q'=\theta Q'_\triangle+(1-\theta)Q'_\triangledown$ in $\mathfrak{M}(\mathfrak{E})$ with the same $\theta\in[0,1]$, we have

$$\begin{aligned}D_{\mathfrak{E}}\left(Q,Q'\right)&\leqslant\theta D_{\mathfrak{E}}\left(Q_\triangle,Q'_\triangle\right)+(1-\theta)\,D_{\mathfrak{E}}\left(Q_\triangledown,Q'_\triangledown\right)\\&\leqslant D_{\mathfrak{E}}\left(Q_\triangle,Q'_\triangle\right)+(1-\theta).\end{aligned}\tag{6.2}$$

It is straightforward that for $\nu,\widetilde{\nu}\in\mathfrak{P}$, and Borel measurable $R:X\to\mathfrak{E}$,

$$D_{\mathfrak{E}}\left(\mathrm{law}_\nu(R),\mathrm{law}_{\widetilde{\nu}}(R)\right)\leqslant d_{\mathfrak{P}}\left(\nu,\widetilde{\nu}\right),\tag{6.3}$$

so that for any sequences (ν_l) and $(\widetilde{\nu}_l)$ in $\mathfrak{P}$, and Borel measurable $R_l:X\to\mathfrak{E}$,

$$d_{\mathfrak{P}}(\nu_l,\widetilde{\nu}_l)\to 0\quad\text{implies}\quad D_{\mathfrak{E}}\left(\mathrm{law}_{\nu_l}(R_l),\mathrm{law}_{\widetilde{\nu}_l}(R_l)\right)\to 0.\tag{6.4}$$

Also, for further measurable $R'_l:X\to\mathfrak{E}$,

$$d_{\mathfrak{E}}\left(R_l,R'_l\right)\overset{\nu_l}{\longrightarrow}0\quad\text{implies}\quad D_{\mathfrak{E}}\left(\mathrm{law}_{\nu_l}\left(R_l\right),\mathrm{law}_{\nu_l}\left(R'_l\right)\right)\to 0,\tag{6.5}$$

because $D_{\mathfrak{E}}(\mathrm{law}_{\nu_l}(R_l),\mathrm{law}_{\nu_l}(R'_l))\leqslant\sum_{j=1}^J\mathrm{Lip}(\chi_j)\int d_{\mathfrak{E}}(R_l,R'_l)\,d\nu_l+2^{-J}$ for each $J\geqslant 1$, and $\int d_{\mathfrak{E}}(R_l,R'_l)\,d\nu_l\to 0$ as $\mathfrak{E}$ is bounded. Finally, note that

$$R=R'\quad\text{on}\quad A\in\mathcal{A}\quad\text{implies}\quad D_{\mathfrak{E}}\left(\mathrm{law}_\nu(R),\mathrm{law}_\nu\left(R'\right)\right)\leqslant\nu\left(A^c\right).\tag{6.6}$$

6.2. Strong distributional convergence

If $(X,\mathcal{A},\mu,T)$ is an ergodic probability preserving system, distributional limit theorems for dynamically defined quantities are often stated in terms of the distinguished measure $\nu:=\mu$. Since the latter is not the only potentially relevant initial distribution for the process, it is both interesting and useful to observe that in many cases such a limit theorem automatically carries over to all probability measures ν absolutely continuous with respect to μ.

For measurable maps R_l, $l\geqslant 1$, of a probability space $(X,\mathcal{A},\mu)$ into $(\mathfrak{E},\mathcal{B}_{\mathfrak{E}})$, *strong distributional convergence* to a random element R of $\mathfrak{E}$, written

$$R_l\overset{\mathcal{L}(\mu)}{\Longrightarrow}R\quad\text{as}\quad l\to\infty,\tag{6.7}$$

means that $R_l \overset{\nu}{\Longrightarrow} R$ for all probability measures $\nu \ll \mu$. (This is equivalent to $R_l \overset{\mu}{\Longrightarrow} R$ (*mixing*), meaning that $R_l \overset{\mu_E}{\Longrightarrow} R$ for every fixed $E \in \mathcal{A}$ with $\mu(E) > 0$. The latter concept dates back to [Rén58], see also [Eag76].)

A property often responsible for this sort of behaviour is that the sequence (R_l) be T-invariant in the long run. Let T be a measure-preserving map on the probability space $(X, \mathcal{A}, \mu)$. For Borel measurable maps $R_l : X \to \mathfrak{E}$, we call the sequence (R_l) *asymptotically T-invariant in (μ-)measure* in case

$$d_{\mathfrak{E}}\left(R_l \circ T, R_l\right) \overset{\mu}{\longrightarrow} 0 \quad \text{as} \quad l \to \infty \tag{6.8}$$

(with $\overset{\mu}{\longrightarrow}$ indicating convergence in measure, recall (3.5)). Note that this is equivalent to convergence in distribution, $d_{\mathfrak{E}}(R_l \circ T, R_l) \overset{\mu}{\Longrightarrow} 0$, because the limit is constant. The important role of this concept becomes clear through

Theorem 6.1 (Strong distributional convergence of asymptotically invariant sequences). — *Let T be an ergodic measure-preserving map on the probability space $(X, \mathcal{A}, \mu)$. Suppose that the sequence (R_l) of Borel measurable maps $R_l : X \to \mathfrak{E}$ into the compact metric space $(\mathfrak{E}, d_{\mathfrak{E}})$ is asymptotically T-invariant in measure. Then*

$$D_{\mathfrak{E}}\left(\mathrm{law}_\nu(R_l), \mathrm{law}_{\overline{\nu}}(R_l)\right) \longrightarrow 0 \quad \textit{as} \quad l \to \infty \quad \textit{for} \quad \nu, \overline{\nu} \in \mathfrak{P}. \tag{6.9}$$

Hence, for R a random element of $\mathfrak{E}$, and any $\nu, \overline{\nu} \in \mathfrak{P}$,

$$R_l \overset{\overline{\nu}}{\Longrightarrow} R \qquad \textit{implies} \qquad R_l \overset{\nu}{\Longrightarrow} R \quad \textit{as} \;\; l \to \infty. \tag{6.10}$$

Proof. — It is clear that (6.9) entails (6.10). The implication (6.10) is the content of [Zwe07b, Theorem 1]. A stronger version of assertion (6.9) is contained in Theorem 6.3 below, whose proof does not use the present theorem.

Alternatively, it is not hard to check directly that (6.10) implies (6.9): Suppose that (6.9) fails, meaning that there are $\nu, \overline{\nu} \in \mathfrak{P}$ with $\delta > 0$ and $l_j \nearrow \infty$ such that

$$D_{\mathfrak{E}}\left(\mathrm{law}_\nu\left(R_{l_j}\right), \mathrm{law}_{\overline{\nu}}\left(R_{l_j}\right)\right) \geqslant \delta \quad \text{for} \;\; j \geqslant 1. \tag{6.11}$$

By Alaoglu's theorem, the metric space $(\mathfrak{M}(\mathfrak{E}), D_{\mathfrak{E}})$ is compact, which allows us to select a further subsequence $l'_i = l_{j_i} \nearrow \infty$ of indices such that $\mathrm{law}_{\overline{\nu}}(R_{l'_i}) \Longrightarrow Q$ for some $Q \in \mathfrak{M}(\mathfrak{E})$. But then (6.10) shows that $\mathrm{law}_\nu(R_{l'_i}) \Longrightarrow Q$ as well, which contradicts (6.11). □

6.3. The transfer operator and mean ergodic theory

Since we shall improve on the above result, we review the main ingredient of its proof. Recall the *transfer operator* $\widehat{T} : L_1(\mu) \to L_1(\mu)$ of T on $(X, \mathcal{A}, \mu)$ which describes the evolution of probability densities under T. That is, if ν has density u w.r.t. μ, $u = d\nu/d\mu$, then $\widehat{T}u := d(\nu \circ T^{-1})/d\mu$. Equivalently, $\int (g \circ T) \cdot u \, d\mu = \int g \cdot \widehat{T}u \, d\mu$ for all $u \in L_1(\mu)$ and $g \in L_\infty(\mu)$, i.e. $\widehat{T}$ is dual to $g \longmapsto g \circ T$. We let $\mathcal{D}(\mu)$ denote the set of probability densities w.r.t. μ.

The following classical companion of the mean ergodic theorem is essentially due to Yosida [Yos38] (see also [Kre85, Theorem 2.1.3], or [Zwe07b, Theorem 2]). Statements (6.12) and (6.13) below are equivalent since we can identify $(\mathfrak{P}, d_{\mathfrak{P}})$ with $(\mathcal{D}(\mu), \|\bullet\|_{L_1(\mu)})$ via $\nu \mapsto d\nu/d\mu$, where $\|d\nu/d\mu - d\overline{\nu}/d\mu\|_{L_1(\mu)} = d_{\mathfrak{P}}(\nu, \overline{\nu})$ for $\nu, \overline{\nu} \in \mathfrak{P}$.

Theorem 6.2 (Characterization of ergodicity). — *Let T be a measure-preserving map on a probability space $(X,\mathcal{A},\mu)$. Then T is ergodic if and only if*

$$\left\| \frac{1}{n}\sum_{k=0}^{n-1} \widehat{T}^k(u-\overline{u}) \right\|_{L_1(\mu)} \longrightarrow 0 \quad \text{as} \quad l\to\infty \quad \text{for all } u,\overline{u}\in\mathcal{D}(\mu), \tag{6.12}$$

which is equivalent to

$$d_{\mathfrak{P}}\left(\frac{1}{n}\sum_{k=0}^{n-1} T_*^k\nu, \frac{1}{n}\sum_{k=0}^{n-1} T_*^k\overline{\nu} \right) \longrightarrow 0 \quad \text{as} \quad l\to\infty \quad \text{for all } \nu,\overline{\nu}\in\mathfrak{P}. \tag{6.13}$$

6.4. Uniform distributional convergence

At the heart of the present paper is a uniform version of the principle of strong distributional convergence for asymptotically invariant sequences of observables quoted above. We capture the key point in the following result.

Theorem 6.3 (Uniform distributional convergence of asymptotically invariant sequences). — *Let $(X,\mathcal{A},\mu,T)$ be an ergodic probability preserving system and $(R_l)_{l\geqslant 1}$ a sequence of Borel measurable maps $R_l : X\to\mathfrak{E}$, asymptotically T-invariant in measure, into a compact metric space $(\mathfrak{E},d_{\mathfrak{E}})$. Let $\mathfrak{K}$ be a compact set in $(\mathfrak{P},d_{\mathfrak{P}})$. Then,*

$$D_{\mathfrak{E}}\left(\mathrm{law}_\nu(R_l),\mathrm{law}_{\overline{\nu}}(R_l)\right) \longrightarrow 0 \quad \begin{array}{c} \text{as } l\to\infty, \\ \text{uniformly in } \nu,\overline{\nu}\in\mathfrak{K}. \end{array} \tag{6.14}$$

Hence, for R a random element of $\mathfrak{E}$, and any two sequences (ν_l), $(\overline{\nu}_l)$ in $\mathfrak{K}$,

$$R_l \overset{\overline{\nu}_l}{\Longrightarrow} R \qquad \text{implies} \qquad R_l \overset{\nu_l}{\Longrightarrow} R \quad \text{as} \quad l\to\infty. \tag{6.15}$$

The key to this refinement of Theorem 6.1 is the following basic principle.

Remark 6.4 (Uniform convergence by equicontinuity). — Let $(\mathfrak{P},d_{\mathfrak{P}})$ be any metric space and $\gamma_M:\mathfrak{P}\to\mathfrak{P}$, $M\geqslant 1$, a sequence of maps which converges pointwise to the continuous map $\gamma:\mathfrak{P}\to\mathfrak{P}$. If $(\gamma_M)_{M\geqslant 1}$ is equicontinuous, then

$$\gamma_M \longrightarrow \gamma \quad \text{uniformly on } \mathfrak{K} \text{ as } M\to\infty$$

whenever $\mathfrak{K}$ is a compact subset of $\mathfrak{P}$. (Indeed, for every $\varepsilon>0$ there is some $\delta>0$ such that $d_{\mathfrak{P}}(\nu,\widetilde{\nu})<\delta$ implies $d_{\mathfrak{P}}(\gamma_M(\nu),\gamma_M(\widetilde{\nu}))<\varepsilon$ for all $M\geqslant 1$. But the compact set $\mathfrak{K}$ contains a finite δ-dense subset, and on the latter $\gamma_M \longrightarrow \gamma$ uniformly as $M\to\infty$.)

Proof of Theorem 6.3. —
(i) The second assertion, implication (6.15), is immediate from (6.14), so we focus on proving the latter. In steps (6.4)-(6.4) below we are going to show that for every $\varepsilon>0$, there is some $l'=l'(\varepsilon)$ such that for any $\chi:\mathfrak{E}\to\mathbb{R}$ with $|\chi|\leqslant 1$ and $\mathrm{Lip}(\varkappa)\leqslant 1$,

$$\left| \int \chi\circ R_l\, d\overline{\nu} - \int \chi\circ R_l\, d\nu \right| < \varepsilon \quad \text{for } l\geqslant l' \quad \text{and} \quad \nu,\overline{\nu}\in\mathfrak{K}. \tag{6.16}$$

To see that this implies (6.14), recall the definition (6.1) of $D_{\mathfrak{E}}$ via our specific sequence (χ_j), and that the χ_j satisfy the above assumptions on χ. Therefore (6.14) follows, since

$$\begin{aligned} D_{\mathfrak{E}}\left(\mathrm{law}_{\nu}(R_l), \mathrm{law}_{\overline{\nu}}(R_l)\right) &= \sum_{j\geqslant 1} 2^{-(j+1)} \left| \int \chi_j \circ R_l \, d\overline{\nu} - \int \chi_j \circ R_l \, d\nu \right| \\ &< \sum_{j\geqslant 1} 2^{-(j+1)} \varepsilon = \varepsilon \quad \text{ for } \nu, \overline{\nu} \in \mathfrak{K} \quad \text{ if } l \geqslant l'. \end{aligned}$$

(ii) Consider the maps $\gamma_M : \mathfrak{P} \to \mathfrak{P}$ with $\gamma_M(\nu) := M^{-1} \sum_{m=0}^{M-1} T_*^m \nu$. By ergodicity and Theorem 6.2, (γ_M) converges pointwise to the constant map $\gamma(\nu) := \mu$, as $\gamma_M(\nu) \to \mu$ for every $\nu \in \mathfrak{P}$. But $(\gamma_M)_{M\geqslant 1}$ is equicontinuous. Indeed, due to the identification of the ν with their densities, it suffices to observe that all the operators $M^{-1} \sum_{m=0}^{M-1} \widehat{T}^m$ have norm equal to 1 on $L_1(\mu)$. Hence, compactness of $\mathfrak{K}$ entails uniform convergence (see Remark 6.4),

$$d_{\mathfrak{P}}\left(\frac{1}{M}\sum_{m=0}^{M-1} T_*^m \overline{\nu}, \frac{1}{M}\sum_{m=0}^{M-1} T_*^m \nu\right) \longrightarrow 0 \qquad \begin{array}{r} \text{as } M \to \infty, \\ \text{uniformly in } \nu, \overline{\nu} \in \mathfrak{K}. \end{array}$$

Therefore there is some $M_\varepsilon \geqslant 1$ (henceforth fixed) such that

$$d_{\mathfrak{P}}\left(\frac{1}{M_\varepsilon}\sum_{m=0}^{M_\varepsilon-1} T_*^m \overline{\nu}, \frac{1}{M_\varepsilon}\sum_{m=0}^{M_\varepsilon-1} T_*^m \nu\right) < \frac{\varepsilon}{4} \quad \text{for} \quad \nu, \overline{\nu} \in \mathfrak{K}.$$

Consequently (as $|\chi| \leqslant 1$), for every l and all $\nu, \overline{\nu} \in \mathfrak{K}$,

$$\left| \int \chi \circ R_l \, d\left(\frac{1}{M_\varepsilon}\sum_{m=0}^{M_\varepsilon-1} T_*^m \overline{\nu}\right) - \int \chi \circ R_l \, d\left(\frac{1}{M_\varepsilon}\sum_{m=0}^{M_\varepsilon-1} T_*^m \nu\right) \right| < \frac{\varepsilon}{4}. \tag{6.17}$$

(iii) As $\mathfrak{K}$ is compact in $(\mathfrak{P}, d_{\mathfrak{P}})$, the family $\{d\nu/d\mu : \nu \in \mathfrak{K}\}$ is compact in $L_1(\mu)$, and hence uniformly integrable. Thus, there is some $\delta = \delta(\varepsilon) > 0$ such that

$$\nu(A) = \int_A \frac{d\nu}{d\mu} \, d\mu < \frac{\varepsilon}{16 M_\varepsilon} \quad \text{for} \quad \nu \in \mathfrak{K} \text{ and } A \in \mathcal{A} \text{ with } \mu(A) < \delta. \tag{6.18}$$

Set $\eta := (8M_\varepsilon)^{-1}\varepsilon > 0$ and define sequences $(\Xi_{i,l})_{l\geqslant 1}$ via $\Xi_{i,l} := \chi \circ R_l \circ T^i$, $i \geqslant 0$. We claim that there is some $l' = l'(\varepsilon)$ s.t.

$$\mu\left(|\Xi_{i,l} - \Xi_{i,l} \circ T| > \eta\right) < \delta \quad \text{for} \quad i \geqslant 0 \text{ and } l \geqslant l'. \tag{6.19}$$

Indeed, using asymptotic T-invariance in measure of (R_l) we see that there is some l' such that

$$\mu\left(d_{\mathfrak{E}}(R_l, R_l \circ T) > \eta\right) < \delta \quad \text{for } l \geqslant l'.$$

Due to T-invariance of μ and $\mathrm{Lip}(\varkappa) \leqslant 1$ we then find that

$$\begin{aligned} \mu\left(|\Xi_{i,l} - \Xi_{i,l} \circ T| > \eta\right) &= \mu\left(|\chi \circ R_l - \chi \circ R_l \circ T| > \eta\right) \\ &\leqslant \mu\left(d_{\mathfrak{E}}(R_l, R_l \circ T) > \eta\right) < \delta \quad \text{ for } i \geqslant 0 \text{ and } l \geqslant l', \end{aligned}$$

as required. Using $|\Xi_{i,l}| \leqslant 1$ and (6.18) we then see that

$$\begin{aligned}\int |\Xi_{i,l} - \Xi_{i,l} \circ T| \, d\nu &\leqslant \eta + \int_{\{|\Xi_{i,l}-\Xi_{i,l}\circ T| > \eta\}} |\Xi_{i,l} - \Xi_{i,l} \circ T| \, d\nu \\ &< \frac{\varepsilon}{4M_\varepsilon} \quad \text{for } i \geqslant 0, l \geqslant l' \text{ and } \nu \in \mathfrak{K}.\end{aligned} \tag{6.20}$$

(iv) Note that by duality and (6.20),

$$\begin{aligned}&\left| \int \chi \circ R_l \, d\nu - \int \chi \circ R_l \, d\left(\frac{1}{M_\varepsilon} \sum_{m=0}^{M_\varepsilon - 1} T_*^m \nu \right) \right| \\ &\qquad \leqslant \frac{1}{M_\varepsilon} \sum_{m=0}^{M_\varepsilon-1} \left| \int (\chi \circ R_l - \chi \circ R_l \circ T^m) \, d\nu \right| \\ &\qquad \leqslant \frac{1}{M_\varepsilon} \sum_{m=0}^{M_\varepsilon-1} \sum_{i=0}^{m-1} \int |\Xi_{i,l} - \Xi_{i,l} \circ T| \, d\nu \\ &\qquad < \frac{1}{M_\varepsilon} \sum_{m=0}^{M_\varepsilon-1} \sum_{i=0}^{m-1} \frac{\varepsilon}{4M_\varepsilon} \leqslant \frac{\varepsilon}{4} \quad \text{for } l \geqslant l' \text{ and } \nu \in \mathfrak{K}.\end{aligned} \tag{6.21}$$

Now take any $\nu, \overline{\nu} \in \mathfrak{K}$, and combine (6.17) with an application of (6.21) to ν and another application of (6.21) to $\overline{\nu}$ to obtain (6.16). □

6.5. Waiting for good measure(s)

We shall say that the measurable functions $\tau_l : X \to \mathbb{N}_0$ form an *admissible delay sequence* $(\tau_l)_{l \geqslant 1}$ *for* (R_l) *and* (ν_l) if

$$D_{\mathfrak{E}} \left(\mathrm{law}_{\nu_l}(R_l), \mathrm{law}_{\nu_l}(R_l \circ T^{\tau_l}) \right) \longrightarrow 0 \quad \text{as} \quad l \to \infty. \tag{6.22}$$

An easy sufficient condition for this is that

$$d_{\mathfrak{E}}(R_l \circ T^{\tau_l}, R_l) \overset{\nu_l}{\longrightarrow} 0 \quad \text{as} \quad l \to \infty. \tag{6.23}$$

(By compactness of $\mathfrak{M}(\mathfrak{E})$ it suffices to show that for any subsequence $l_j \nearrow \infty$ of indices and any random element R of $\mathfrak{E}$, $R_{l_j} \overset{\nu_{l_j}}{\Longrightarrow} R$ implies $R_{l_j} \circ T^{\tau_l} \overset{\nu_{l_j}}{\Longrightarrow} R$, which follows from (6.23) by a standard (Slutsky) argument, see [Bil99, Theorem 3.1]).

Since $\mathrm{law}_{\nu_l}(R_l \circ T^{\tau_l}) = \mathrm{law}_{\overline{\nu}_l}(R_l)$ where $\overline{\nu}_l := T_*^{\tau_l} \nu_l$, we can efficiently use admissible delays in situations where the sequence $(\overline{\nu}_l)$ of these push-forwards allows for better control than (ν_l). The latter phrase will mean that $\overline{\nu}_l \in \mathfrak{K}$ for all l, where $\mathfrak{K} \subseteq \mathfrak{P}$ is compact, in which case we can use the following straightforward consequence of Theorem 6.3.

Proposition 6.5 (Asymptotically invariant sequences - ν_l versus μ). — *Let* $(X, \mathcal{A}, \mu, T)$ *be an ergodic probability preserving system,* $(R_l)_{l \geqslant 1}$ *a sequence of Borel measurable maps* $R_l : X \to \mathfrak{E}$*, asymptotically* T*-invariant in measure, into a compact metric space* $(\mathfrak{E}, d_{\mathfrak{E}})$*, and* $(\nu_l)_{l \geqslant 1}$ *a sequence in* $\mathfrak{P}$*. Suppose that* $(\tau_l)_{l \geqslant 1}$ *is an admissible delay sequence for* (R_l) *and* (ν_l) *such that there is some compact set* $\mathfrak{K}$ *in* $(\mathfrak{P}, d_{\mathfrak{P}})$ *for which*

$$T_*^{\tau_l} \nu_l \in \mathfrak{K} \quad \textit{for} \quad l \geqslant 1. \tag{6.24}$$

Then

$$D_{\mathfrak{E}}\left(\mathrm{law}_{\nu_l}(R_l), \mathrm{law}_{\mu}(R_l)\right) \longrightarrow 0 \quad \textit{as} \quad l \to \infty. \tag{6.25}$$

Proof. — Letting $\overline{\nu}_l := T_*^{\tau_l}\nu_l$ we have $D_{\mathfrak{E}}(\mathrm{law}_{\nu_l}(R_l), \mathrm{law}_{\overline{\nu}_l}(R_l)) \to 0$ due to admissibility of (τ_l), but also $D_{\mathfrak{E}}(\mathrm{law}_{\overline{\nu}_l}(R_l), \mathrm{law}_{\mu}(R_l)) \to 0$ by Theorem 6.3 (We can assume w.l.o.g. that $\mu \in \mathfrak{K}$). □

In the specific situation of Theorems 3.6, 4.8, and 5.3 above, the $\nu_l = \mu_{A_l}$ concentrate on ever smaller sets, and hence can never stay inside a single compact set $\mathfrak{K}$, while suitable push-forwards $\overline{\nu}_l$ sometimes do.

6.6. Independent limits for pairs of asymptotically invariant sequences

To facilitate the analysis of distributional limits of processes involving several asymptotically invariant sequences, we provide a natural method of checking asymptotic independence. It relies on the following easy probability fact.

Lemma 6.6 (Independence of limit variables by conditioning). — *Let $(\nu_l)_{l \geqslant 1}$ be probability measures on $(X, \mathcal{A})$, and $R_l : X \to \mathfrak{E}$, $R'_l : X \to \mathfrak{E}'$, $l \geqslant 1$, Borel maps into the compact metric spaces $\mathfrak{E}$ and $\mathfrak{E}'$. Let (R, R') be a random element of $\mathfrak{E} \times \mathfrak{E}'$ such that*

$$(R_l, R'_l) \overset{\nu_l}{\Longrightarrow} (R, R') \quad \textit{as} \quad l \to \infty. \tag{6.26}$$

Assume that there is some π-system $\mathcal{B}_{\mathfrak{E}}^{\pi}$ generating $\mathcal{B}_{\mathfrak{E}}$ such that for all $E \in \mathcal{B}_{\mathfrak{E}}^{\pi}$ we have $\Pr[R \in \partial E] = 0$, and, in case $\Pr[R \in E] > 0$, convergence in law holds under the ν_l conditioned on $\{R_l \in E\}$,

$$R'_l \overset{\nu_{l,\{R_l \in E\}}}{\Longrightarrow} R' \quad \textit{as} \quad l \to \infty. \tag{6.27}$$

Then R and R' are independent.

Proof. — By [Bil99, Theorem 2.8], our assumption (6.26) is equivalent to

$$\begin{gathered}\nu_l\left(R_l \in E, R'_l \in E'\right) \longrightarrow \Pr[R \in E, R' \in E'] \quad \text{as} \quad l \to \infty \\ \text{whenever } \Pr\left[R \in \partial E\right] = \Pr\left[R' \in \partial E'\right] = 0.\end{gathered} \tag{6.28}$$

We show that for any such (E, E') this limit coincides with $\Pr[R \in E] \Pr[R' \in E']$. Applying the same theorem from [Bil99] again, then proves that (R_l, R'_l) converges to an independent pair with marginals R, R'.

Take any $E \in \mathcal{B}_{\mathfrak{E}}^{\pi}$. If $\Pr[R \in E] = 0$, the assertion is trivial. Assume therefore that $\Pr[R \in E] > 0$. Pick any $E' \in \mathcal{B}_{\mathfrak{E}'}$ with $\Pr[R' \in \partial E'] = 0$. Due to (6.27),

$$\nu_l\left(R_l \in E, R'_l \in E'\right) \longrightarrow \Pr[R \in E] \Pr\left[R' \in E'\right] \quad \text{as} \quad l \to \infty,$$

and therefore

$$\Pr\left[R \in E, R' \in E'\right] = \Pr[R \in E] \Pr\left[R' \in E'\right].$$

Fixing such an E', the standard uniqueness theorem for measures shows that

$$\Pr\left[R \in B, R' \in E'\right] = \Pr[R \in B] \Pr\left[R' \in E'\right] \quad \text{for all} \quad B \in \mathcal{B}_{\mathfrak{E}}.$$

In particular, this is true for all $B \in \mathcal{B}_{\mathfrak{E}}$ with $\Pr[R \in \partial B] = 0$. □

Combining this with the uniform distributional convergence principle, and the idea that admissible time delays may result in good measures, we obtain

Theorem 6.7 (Asymptotic independence of two sequences). — *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system and $(R_l)_{l \geqslant 1}$, $(R'_l)_{l \geqslant 1}$ sequences of Borel measurable maps, (R'_l) asymptotically T-invariant in measure, into compact metric spaces $(\mathfrak{E}, d_{\mathfrak{E}})$ and $(\mathfrak{E}', d_{\mathfrak{E}'})$, respectively. Suppose that ν_l, $l \geqslant 1$, are probabilities on $(X, \mathcal{A})$ such that*

$$(R_l, R'_l) \overset{\nu_l}{\Longrightarrow} (R, R') \quad \text{as} \quad l \to \infty \tag{6.29}$$

for some random element (R, R') of $\mathfrak{E} \times \mathfrak{E}'$, and let $\mathcal{B}^{\pi}_{\mathfrak{E}}$ be some π-system generating $\mathcal{B}_{\mathfrak{E}}$ such that $\mathfrak{E} \in \mathcal{B}^{\pi}_{\mathfrak{E}}$ while $\Pr[R \in \partial E] = 0$ for all $E \in \mathcal{B}^{\pi}_{\mathfrak{E}}$.

Assume that for every $E \in \mathcal{B}^{\pi}_{\mathfrak{E}}$ with $\Pr[R \in E] > 0$ there is some compact set $\mathfrak{K}_E$ in $(\mathfrak{P}, d_{\mathfrak{P}})$, a sequence $(\nu_{l,E})_{l \geqslant 1}$ in $\mathfrak{P}$ with $d_{\mathfrak{P}}(\nu_{l,E}, \nu_{l,\{R_l \in E\}}) \to 0$, and a sequence $(\tau_{l,E})_{l \geqslant 1}$ of admissible delays for $(R'_l)_{l \geqslant 1}$ and $(\nu_{l,E})_{l \geqslant 1}$ such that $\overline{\nu}_{l,E} := T^{\tau_{l,E}}_* \nu_{l,E} \in \mathfrak{K}_E$ for $l \geqslant 1$. Then R and R' are independent.

(Note that (R_l) is not required to be asymptotically T-invariant.)

Proof. — By Lemma 6.6 it suffices to show that for every $E \in \mathcal{B}^{\pi}_{\mathfrak{E}}$ with $\Pr[R \in E] > 0$,

$$R'_l \overset{\nu_{l,\{R_l \in E\}}}{\Longrightarrow} R' \quad \text{as} \quad l \to \infty,$$

which, due to $d_{\mathfrak{P}}(\nu_{l,E}, \nu_{l,\{R_l \in E\}}) \to 0$ and (6.4), is equivalent to

$$R'_l \overset{\nu_{l,E}}{\Longrightarrow} R' \quad \text{as} \quad l \to \infty. \tag{6.30}$$

Take any such E. Since $\overline{\nu}_{l,E} \in \mathfrak{K}_E$ for all l, an application of Theorem 6.3 shows that $D_{\mathfrak{E}}(\mathrm{law}_{\mu}(R'_l), \mathrm{law}_{\overline{\nu}_{l,E}}(R'_l)) \to 0$, and since the $\tau_{l,E}$ are admissible delays for $(R'_l)_{l \geqslant 1}$ and $(\nu_{l,E})_{l \geqslant 1}$, we conclude that

$$D_{\mathfrak{E}}\left(\mathrm{law}_{\mu}\left(R'_l\right), \mathrm{law}_{\nu_{l,E}}\left(R'_l\right)\right) \longrightarrow 0 \quad \text{as} \quad l \to \infty. \tag{6.31}$$

The case $E := \mathfrak{E}$ yields $D_{\mathfrak{E}}(\mathrm{law}_{\mu}(R'_l), \mathrm{law}_{\nu_l}(R'_l)) \to 0$. Together with $R'_l \overset{\nu_l}{\Longrightarrow} R'$ and (6.31) this gives (6.30). $\square$

Remark 6.8 (The auxiliary measures $\nu_{l,E}$). — In the simplest cases, we can take $\nu_{l,E} := \nu_{l,\{R_l \in E\}}$. However, constructing suitable $\tau_{l,E}$ is sometimes easier if we use a slightly different sequence of measures, obtained as follows.

It is easily seen that if (B_l) and (B'_l) are sequences in $\mathcal{A}$ with $\mu(B_l \triangle B'_l) = o(\mu(B_l))$, then $d_{\mathfrak{P}}(\mu_{B_l}, \mu_{B'_l}) \to 0$. Therefore, if $\nu_l = \mu_{A_l}$ and the sets $B'_{l,E} \in \mathcal{A}$ are such that $\mu(B'_{l,E} \triangle (A_l \cap \{R_l \in E\})) = o(\mu(A_l \cap \{R_l \in E\}))$, then the measures $\nu_{l,E} := \mu_{B'_{l,E}}$ satisfy $d_{\mathfrak{P}}(\nu_{l,E}, \nu_{l,\{R_l \in E\}}) \to 0$.

7. Proofs for return- and hitting-time processes

7.1. Asymptotic invariance of hitting time processes

A sequence of hitting time processes for rare events, that is, a sequence (R_l) of variables $R_l = \mu(A_l)\Phi_{A_l}$, viewed through the single measure μ, is asymptotically

T-invariant in measure [Zwe07b, Corollary 6]. We provide a more precise statement in the next proposition.

Proposition 7.1 (Asymptotic invariance in measure of hitting-time processes). *Let $(X, \mathcal{A}, \mu, T)$ be a probability preserving system.*

(a) *For every set $A \in \mathcal{A}$ and integer $m \geqslant 0$,*

$$d_{[0,\infty]^{\mathbb{N}_0}}\left(\mu(A)\Phi_A \circ T^m, \mu(A)\Phi_A\right) \leqslant m\mu(A) \quad \text{on} \quad \{\varphi_A > m\}. \tag{7.1}$$

(b) *Suppose that $(A_l)_{l\geqslant 1}$ is a sequence of asymptotically rare events, and set $R_l := \mu(A_l)\Phi_{A_l} : X \to [0,\infty]^{\mathbb{N}_0}$. Then (R_l) is asymptotically T-invariant in measure.*

(c) *Likewise, if $\varphi_{A_l} < \infty$ a.e. and $R_l := \mu(A_l)\Phi_{A_l} \circ T_{A_l} : X \to [0,\infty]^{\mathbb{N}}$, then (R_l) is asymptotically T-invariant in measure.*

Proof. —

(a) For any $A \in \mathcal{A}$ and integer $m \geqslant 0$,

$$\Phi_A = \Phi_A \circ T^m + (m, 0, 0, \ldots) \quad \text{on} \quad \{\varphi_A > m\}, \tag{7.2}$$

and hence $d_{[0,\infty]^{\mathbb{N}_0}}(\mu(A)\Phi_A \circ T^m, \mu(A)\Phi_A) = d_{[0,\infty]}(\mu(A)(\varphi_A - m), \mu(A)\varphi_A)$ on that set. Since $d_{[0,\infty]}(s, s+\delta) = e^{-s}(1-e^{-\delta}) \leqslant \delta$ for all $s, \delta \in [0,\infty)$, (7.1) follows.

(b) Whenever $(A_l)_{l\geqslant 1}$ is a sequence of asymptotically rare events, the R_l satisfy

$$d_{[0,\infty]^{\mathbb{N}_0}}\left(R_l \circ T, R_l\right) \leqslant \mu(A_l) \quad \text{outside}\{\varphi_{A_l} \leqslant 1\} = T^{-1}A_l. \tag{7.3}$$

By assumption this upper bound for the distance tends to zero, and since $\mu(\varphi_{A_l} \leqslant 1) = \mu(A_l)$, so does the measure of the set on which the bound fails to apply.

(c) Analogous, using that $R_l \circ T = R_l$ outside $T^{-1}A_l$. □

The simple estimate (7.1) immediately leads to sufficient conditions for time delays τ_l to be admissible for a given sequence (ν_l) of initial densities.

Proposition 7.2 (Admissible time delays for return or hitting processes). — *Let T be a measure-preserving map on the probability space $(X, \mathcal{A}, \mu)$, (A_l) a sequence of asymptotically rare events, (ν_l) a sequence in $\mathfrak{P}$, and $\tau_l : X \to \mathbb{N}_0$, $l \geqslant 1$, measurable functions.*

(a) *$(\tau_l)_{l\geqslant 1}$ is an admissible delay sequence for the variables $R_l : X \to [0,\infty]^{\mathbb{N}_0}$ given by $R_l := \mu(A_l)\Phi_{A_l}$ and the measures ν_l provided that*

$$\mu(A_l)\, \tau_l \xrightarrow{\nu_l} 0 \quad \text{as} \quad l \to \infty, \tag{7.4}$$

and

$$\nu_l\left(\tau_l < \varphi_{A_l}\right) \longrightarrow 1 \quad \text{as} \quad l \to \infty. \tag{7.5}$$

(b) *Condition (7.5) alone is sufficient for $(\tau_l)_{l\geqslant 1}$ to be an admissible delay sequence for the variables $R_l : X \to [0,\infty]^{\mathbb{N}}$ given by $R_l := \mu(A_l)\Phi_{A_l} \circ T_{A_l}$ and the ν_l.*

Proof. —

(a) For arbitrary $A \in \mathcal{A}$ and any measurable $\tau : X \to \mathbb{N}_0$, we can apply (7.1) on each set $\{\tau = m\}$, $m \geqslant 0$, to see that

$$d_{[0,\infty]^{\mathbb{N}_0}}(\mu(A)\Phi_A \circ T^\tau, \mu(A)\Phi_A) \leqslant \tau\mu(A) \quad \text{on} \quad \{\varphi_A > \tau\}. \tag{7.6}$$

Now take any $\varepsilon > 0$. By the above we find that for every l,

$$\begin{aligned} d_{[0,\infty]^{\mathbb{N}_0}}(R_l \circ T^{\tau_l}, R_l) &\leqslant \tau_l\, \mu(A_l) \quad \text{on} \quad \{\varphi_{A_l} > \tau_l\} \\ &< \varepsilon \qquad\qquad \text{on} \quad \{\varphi_{A_l} > \tau_l\} \cap \{\mu(A_l)\, \tau_l < \varepsilon\}. \end{aligned} \tag{7.7}$$

But (7.4) and (7.5) ensure that $\nu_l(\{\varphi_{A_l} > \tau_l\} \cap \{\mu(A_l)\, \tau_l < \varepsilon\}) \to 1$ as $l \to \infty$, which proves our claim via the sufficient condition (6.23).

(b) Since, for any A and measurable τ we have $T_A \circ T^{\tau} = T_A$ on $\{\tau < \varphi_A\}$, we see that $R_l \circ T^{\tau_l} = R_l$ on $\{\tau_l < \varphi_{A_l}\}$, and the result follows. □

We can thus establish the first theorem advertised in this paper.

Proof of Theorem 3.2. — Conditions (3.6) and (3.7) guarantee, via Propositions 7.1 (b) and 7.2 (a), that (R_l) is asymptotically T-invariant in measure, and that (τ_l) is an admissible delay sequence for $(R_l) := (\mu(A_l)\Phi_{A_l})$ and (ν_l). Now (3.8) allows us to apply Proposition 6.5. □

7.2. Finite-dimensional marginals and distributional convergence

A sequence $s \in [0,\infty]^{\mathbb{N}_0}$ will be called *finite-valued* in case $s \in [0,\infty)^{\mathbb{N}_0}$. Let $\Phi = (\varphi^{(j)})_{j \geqslant 0}$ be a *random sequence in* $[0,\infty]$, that is, a random element of $[0,\infty]^{\mathbb{N}_0}$. We let $\Phi^{[d]} := (\varphi^{(0)}, \ldots, \varphi^{(d-1)})$ denote its initial piece of length d, $d \geqslant 1$. The (possibly degenerate) *distribution function* of the random vector $\Phi^{[d]}$ is $F^{[d]} : [0,\infty)^d \to [0,1]$, $F^{[d]}(t_0, \ldots, t_{d-1}) := \Pr[\varphi^{(0)} \leqslant t_0, \ldots, \varphi^{(d-1)} \leqslant t_{d-1}]$.

Assume that each Φ_l, $l \geqslant 1$, is a random sequence in $[0,\infty)$ with finite-dimensional distribution functions $F_l^{[d]} : [0,\infty)^d \to [0,1]$, $d \geqslant 1$. Abbreviating $\{F_l^{[d]}\} := \{F_l^{[d]}\}_{d \geqslant 1}$ we shall write

$$\left\{F_l^{[d]}\right\} \Longrightarrow \left\{F^{[d]}\right\} \quad \text{as} \quad l \to \infty, \tag{7.8}$$

if all $F_l^{[d]}$ converge weakly, as $l \to \infty$, to the corresponding distribution functions $F^{[d]}$ of Φ, that is, for every $d \geqslant 1$ we have $F_l^{[d]}(t_0, \ldots, t_{d-1}) \to F^{[d]}(t_0, \ldots, t_{d-1})$ at all continuity points $(t_0, \ldots, t_{d-1})$ of $F^{[d]}$.

Remark 7.3. — This is the mode of convergence studied in [Zwe16], where it was denoted by $\Phi_l \Longrightarrow \Phi$. It is closely related to the present meaning of $\Phi_l \Longrightarrow \Phi$ (distributional convergence of random elements of $[0,\infty]^{\mathbb{N}_0}$), which clearly implies $\{F_l^{[d]}\} \Longrightarrow \{F^{[d]}\}$. In fact, the two notions coincide in case the Φ_l and Φ are a.s. finite-valued, which is always the case for return-time processes Φ_{A_l} viewed through μ_{A_l}, and their limits $\widetilde{\Phi}$ (recall (3.3)).

7.3. The general duality between return- and hitting-time processes

It is a basic fact that for any sequence $(A_l)_{l \geqslant 1}$ of asymptotically rare events, its return-time statistics and its hitting-time statistics are intimately related to each other, as established in [HLV05]. This result has been extended to the level of processes in [Zwe16] (see also [Mar17]), where we proved

Theorem 7.4 (Hitting-time process versus return-time process [Zwe16]). — *Let* $(X, \mathcal{A}, \mu, T)$ *be an ergodic probability-preserving system, and* $(A_l)_{l \geqslant 1}$ *a sequence of asymptotically rare events. Let* $\{F_l^{[d]}\}$ *and* $\{\widetilde{F}_l^{[d]}\}$ *be the collections of finite-dimensional distribution functions of* $\mu(A_l)\Phi_{A_l}$ *under* μ *and the* μ_{A_l}*, respectively. Then*

$$\left\{F_l^{[d]}\right\} \Longrightarrow \left\{F^{[d]}\right\} \quad \text{for some process } \Phi \text{ in } [0,\infty] \text{ with d.f.s } \left\{F^{[d]}\right\} \tag{7.9}$$

iff

$$\left\{\widetilde{F}_l^{[d]}\right\} \Longrightarrow \left\{\widetilde{F}^{[d]}\right\} \quad \text{for some process } \widetilde{\Phi} \text{ in } [0,\infty] \text{ with d.f.s } \left\{\widetilde{F}^{[d]}\right\}. \tag{7.10}$$

In this case, the sub-probability distribution functions $F^{[d]}$ *and* $\widetilde{F}^{[d]}$ *of* $\Phi^{[d]}$ *and* $\widetilde{\Phi}^{[d]}$ *satisfy, for any* $d \geqslant 0$ *(where* $\widetilde{F}^{[0]} := 1$*) and* $t_j \geqslant 0$,

$$\int_0^{t_0} \left[\widetilde{F}^{[d]}(t_1, \ldots, t_d) - \widetilde{F}^{[d+1]}(s, t_1, \ldots, t_d)\right] ds = F^{[d+1]}(t_0, t_1, \ldots, t_d). \tag{7.11}$$

Through (7.11)*, the families* $\{F^{[d]}\}$ *and* $\{\widetilde{F}^{[d]}\}$ *uniquely determine each other.*

We will heavily rely on this duality.

7.4. The case of Poisson asymptotics

Below we will primarily be interested in the particular case where the limit process is an iid sequence of normalized exponentially distributed random variables Φ_{Exp}. If this particular limit occurs, then it automatically occurs both for the hitting-times and for the return times, because of Theorem 7.4 and

Proposition 7.5 (Characterizing Φ_{Exp} [Zwe16]). — *Let* Φ *be some stationary random sequence in* $[0,\infty)$*. Then* $\Phi = \Phi_{\mathrm{Exp}}$ *iff the finite-dimensional marginals have distribution functions* $F^{[d]}$ *satisfying*

$$F^{[d+1]}(t_0, t_1, \ldots, t_d) = \int_0^{t_0} \left[F^{[d]}(t_1, \ldots, t_d) - F^{[d+1]}(s, t_1, \ldots, t_d)\right] ds \tag{7.12}$$

whenever $d \geqslant 0$ *and* $t_j \geqslant 0$.

Combining the above with Theorem 3.2 we can now prove the abstract temporal Poisson limit theorem.

Proof of Theorem 3.6. —
(i) Let $R_l := \mu(A_l)\Phi_{A_l}$, $l \geqslant 1$, which gives an asymptotically invariant sequence of Borel measurable maps $R_l : X \to [0,\infty]^{\mathbb{N}_0}$. Due to compactness of $\mathfrak{M}([0,\infty]^{\mathbb{N}_0})$, it suffices to show that for every subsequence of indices $l_j \nearrow \infty$ along which

$$R_{l_j} \overset{\mu_{A_{l_j}}}{\Longrightarrow} \widetilde{\Phi} \quad \text{and} \quad R_{l_j} \overset{\mu}{\Longrightarrow} \Phi$$

for some random elements $\widetilde{\Phi}, \Phi$ of $[0,\infty]^{\mathbb{N}_0}$, both limits are iid with marginal $\widetilde{F}_{\mathrm{Exp}}$. Focusing on such a subsequence, we thus assume that

$$R_l \overset{\mu_{A_l}}{\Longrightarrow} \widetilde{\Phi} \quad \text{and} \quad R_l \overset{\mu}{\Longrightarrow} \Phi \quad \text{as} \quad l \to \infty. \tag{7.13}$$

We need to show that

$$\text{law}(\widetilde{\Phi}) = \text{law}(\Phi) = \text{law}\left(\widetilde{\Phi}_{\text{Exp}}\right). \tag{7.14}$$

(ii) Take any $\varepsilon > 0$ and choose $(\nu_l)_{l\geqslant 1}$ in $\mathfrak{P}$ with $d_{\mathfrak{P}}(\nu_l, \mu_{A_l}) < \varepsilon$ for all l, a compact set $\mathfrak{K} \subseteq \mathfrak{P}$, and measurable functions $\tau_l : X \to \mathbb{N}_0$ such that (3.15)-(3.17) hold. By Theorem 3.2 we have $D_{[0,\infty]^{\mathbb{N}_0}}(\text{law}_{\nu_l}(R_l), \text{law}_{\mu}(R_l)) \to 0$, so that

$$R_l \overset{\nu_l}{\Longrightarrow} \Phi \quad \text{as} \quad l \to \infty. \tag{7.15}$$

(iii) In view of (6.3), however, $D_{[0,\infty]^{\mathbb{N}_0}}(\text{law}_{\nu_l}(R_l), \text{law}_{\mu_{A_l}}(R_l)) < \varepsilon$ for all l, and letting $l \to \infty$, (7.13) and (7.15) allow us to conclude that $D_{[0,\infty]^{\mathbb{N}_0}}(\text{law}(\Phi), \text{law}(\widetilde{\Phi})) \leqslant \varepsilon$. But $\varepsilon > 0$ was arbitrary, and therefore

$$\text{law}(\Phi) = \text{law}\left(\widetilde{\Phi}\right). \tag{7.16}$$

For the finite dimensional distribution functions $\{F^{[d]}\}$ and $\{\widetilde{F}^{[d]}\}$ of Φ and $\widetilde{\Phi}$ this means that $\{F^{[d]}\} = \{\widetilde{F}^{[d]}\}$. Together with the fact that $\{F^{[d]}\}$ and $\{\widetilde{F}^{[d]}\}$ are related to each other as in (7.4) of Theorem 7.4, this shows that $\{F^{[d]}\}$ satisfies condition (7.12), and (7.14) follows by Proposition 7.5. ☐

7.5. Including a point mass at zero

The argument for convergence to $\widetilde{\Phi}_{(\text{Exp},\theta)}$ is similar to that for $\widetilde{\Phi}_{\text{Exp}}$. As a warm-up we characterize the one-dimensional distribution function $\widetilde{F}_{(\text{Exp},\theta)}$ through a generalization of the fixed point equation $\widetilde{F}(t) = \int_0^t [1 - \widetilde{F}(s)]\,ds$ distinguishing the exponential distribution function $F = \widetilde{F}_{(\text{Exp},1)}$. In a second step, we provide a characterization of $\widetilde{\Phi}_{(\text{Exp},\theta)}$ similar to Proposition 7.5.

Proposition 7.6 (Characterizing $\widetilde{F}_{(\text{Exp},\theta)}$ and $\widetilde{\Phi}_{(\text{Exp},\theta)}$). — *Take any $\theta \in (0, 1]$.*

(a) *If $\widetilde{F}$ is a probability distribution function on $[0,\infty)$, then $\widetilde{F} = \widetilde{F}_{(\text{Exp},\theta)}$ iff*

$$\widetilde{F}(t) = (1-\theta) + \theta \int_0^t \left[1 - \widetilde{F}(s)\right] ds \quad \text{for} \quad t \geqslant 0. \tag{7.17}$$

(b) *If $\widetilde{\Phi}$ is a stationary sequence of random variables in $[0,\infty)$ with finite-dimensional distribution functions $\widetilde{F}^{[d]}$ (where $\widetilde{F}^{[0]} := 1$), then $\text{law}(\widetilde{\Phi}) = \text{law}(\widetilde{\Phi}_{(\text{Exp},\theta)})$ iff*

$$\begin{aligned} \widetilde{F}^{[d+1]}(t_0, t_1, \ldots, t_d) &= (1-\theta)\,\widetilde{F}^{[d]}(t_1, \ldots, t_d) \\ &\quad + \theta \int_0^{t_0} \left[\widetilde{F}^{[d]}(t_1, \ldots, t_d) - \widetilde{F}^{[d+1]}(s, t_1, \ldots, t_d)\right] ds \end{aligned} \tag{7.18}$$

whenever $d \geqslant 0$ and $t_j \geqslant 0$.

Proof. —
(a) It is immediate that $\widetilde{F}_{(\mathrm{Exp},\theta)}$ from (3.19) satisfies (7.17). For the converse, assume (7.17) and let $F(t) := \int_0^t [1-\widetilde{F}(s)]\,ds$ for $t \geqslant 0$. Due to (7.17) we have $1-\widetilde{F}(s) = \theta[1-F(s)]$ and hence

$$F(t) = \theta \int_0^t [1-F(s)]\,ds \quad \text{for} \quad t \geqslant 0.$$

Therefore F is $\mathcal{C}^\infty$ on $(0,\infty)$, and $\overline{F}(t) := 1-F(t)$ satisfies $\overline{F}' = -\theta\overline{F}$ there. Consequently, $\overline{F}(t) = ce^{-\theta t}$, and since $F(0^+) = 0$ we have $c = 1$.
(b) Using (7.17) it is straightforward that the marginals of the iid sequence $\widetilde{\Phi}_{(\mathrm{Exp},\theta)}$ satisfy (7.18). For the converse, assume that $\widetilde{\Phi}$ satisfies (7.18).

The $d = 0$ case covered by part a) shows that $\widetilde{F}^{[1]} = \widetilde{F}_{(\mathrm{Exp},\theta)}$. Write $\widetilde{\Phi} = (\widetilde{\varphi}^{(j)})_{j\geqslant 0}$, then by stationarity, each $\widetilde{\varphi}^{(j)}$ has distribution $\widetilde{F}_{(\mathrm{Exp},\theta)}$. We need to prove that the $\widetilde{\varphi}^{(j)}$ are independent. Using stationarity again, we see that it suffices to check that

$$\text{for } d \geqslant 1, \text{the variable } \widetilde{\varphi}^{(0)} \text{ is independent of } \left\{\widetilde{\varphi}^{(1)}, \ldots, \widetilde{\varphi}^{(d)}\right\}. \tag{7.19}$$

Fix any $d \geqslant 1$, and take $(t_1, \ldots, t_d)$ such that $\widetilde{F}^{[d]}(t_1, \ldots, t_d) > 0$. Define

$$\begin{aligned}\widetilde{G}(s) &:= \Pr\left[\widetilde{\varphi}^{(0)} \leqslant s \,\middle|\, \widetilde{\varphi}^{(1)} \leqslant t_1, \ldots, \widetilde{\varphi}^{(d)} \leqslant t_d\right] \\ &= \widetilde{F}^{[d+1]}(s, t_1, \ldots, t_d) / \widetilde{F}^{[d]}(t_1, \ldots, t_d)\end{aligned}$$

for $s \geqslant 0$. Then (7.18) becomes

$$\widetilde{G}(t) = (1-\theta) + \theta \int_0^t \left[1-\widetilde{G}(s)\right] ds \quad \text{for } s \geqslant 0.$$

But since $\widetilde{\varphi}^{(0)}$ takes values in $[0,\infty)$, part a) ensures that $\widetilde{G} = \widetilde{F}_{(\mathrm{Exp},\theta)}$, meaning that $\Pr[\widetilde{\varphi}^{(0)} \leqslant s \mid \widetilde{\varphi}^{(1)} \leqslant t_1, \ldots, \widetilde{\varphi}^{(d)} \leqslant t_d] = \Pr[\widetilde{\varphi}^{(0)} \leqslant s]$ whenever the conditioning event has positive probability. This establishes (7.19). □

We are now ready for the proof of Theorem 3.8. The strategy is the same as in the case of Theorem 3.6, but we now split off the contribution of points which return within time τ_l.

Proof of Theorem 3.8. —
(i) Let $R_l := \mu(A_l)\Phi_{A_l}$, $l \geqslant 1$. As in the proof of Theorem 3.6 we can assume w.l.o.g. that

$$R_l \overset{\mu_{A_l}}{\Longrightarrow} \widetilde{\Phi} \quad \text{and} \quad R_l \overset{\mu}{\Longrightarrow} \Phi \quad \text{as} \quad l \to \infty. \tag{7.20}$$

We need to show that

$$\mathrm{law}\left(\widetilde{\Phi}\right) = \mathrm{law}\left(\widetilde{\Phi}_{(\mathrm{Exp},\theta)}\right). \tag{7.21}$$

(ii) Take any $\varepsilon > 0$ and choose $(\nu_l)_{l\geqslant 1}$ in $\mathfrak{P}$ with $d_{\mathfrak{P}}(\nu_l, \mu_{A_l}) < \varepsilon$ for all l, a compact set $\mathfrak{K} \subseteq \mathfrak{P}$, and measurable functions $\tau_l : X \to \mathbb{N}_0$ such that (3.21)-(3.25) hold. Abbreviate $\theta_l := \nu_l(A_l^\circ)$, so that

$$\nu_l = (1-\theta_l)\,\nu_l^\bullet + \theta_l \nu_l^\circ \quad \text{for} \quad l \geqslant 1. \tag{7.22}$$

Assumptions (3.21) to (3.23) allow us to apply Theorem 3.2 using the sequence (ν_l°) of measures, to see that $D_{[0,\infty]^{\mathbb{N}_0}}(\mathrm{law}_{\nu_l^\circ}(R_l), \mathrm{law}_\mu(R_l)) \to 0$. Consequently,

$$R_l \overset{\nu_l^\circ}{\Longrightarrow} \Phi \quad \text{as} \quad ł \to \infty. \tag{7.23}$$

To analyse the asymptotic distribution of R_l under $\nu_l^\bullet$ we observe first that due to (3.24) the assumption (3.21) implies that also

$$\mu(A_l)\,\varphi_{A_l} \overset{\nu_l^\bullet}{\longrightarrow} 0 \quad \text{as} \quad l \to \infty. \tag{7.24}$$

On the other hand,

$$\mathrm{law}_{\nu_l^\bullet}(R_l \circ T_{A_l}) = \mathrm{law}_{\overline{\nu}_l^\bullet}(R_l), \quad \text{where} \quad \overline{\nu}_l^\bullet := (T_{A_l})_* \nu_l^\bullet.$$

In view of (3.25) and (6.3), we thus have $D_{[0,\infty]^{\mathbb{N}_0}}(\mathrm{law}_{\nu_l^\bullet}(R_l \circ T_{A_l}), \mathrm{law}_{\mu_{A_l}}(R_l)) = D_{[0,\infty]^{\mathbb{N}_0}}(\mathrm{law}_{\overline{\nu}_l^\bullet}(R_l), \mathrm{law}_{\mu_{A_l}}(R_l)) \to 0$. Recalling (7.20) and the fact that $\widetilde{\Phi}$ is stationary (see (3.3)), this shows that

$$R_l \circ T_{A_l} \overset{\nu_l^\bullet}{\Longrightarrow} \boldsymbol{\sigma}\widetilde{\Phi} \quad \text{as} \quad l \to \infty, \tag{7.25}$$

where $\boldsymbol{\sigma}\widetilde{\Phi} := (\widetilde{\varphi}^{(1)}, \widetilde{\varphi}^{(2)}, \ldots)$ is the shifted version of $\widetilde{\Phi} = (\widetilde{\varphi}^{(0)}, \widetilde{\varphi}^{(1)}, \ldots)$. Since the limit in (7.24) is constant, and hence independent of all random variables, we can combine (7.24) and (7.25) to obtain

$$R_l = (\mu(A_l)\,\varphi_{A_l}, R_l \circ T_{A_l}) \overset{\nu_l^\bullet}{\Longrightarrow} \left(0, \boldsymbol{\sigma}\widetilde{\Phi}\right) = \left(0, \widetilde{\varphi}^{(1)}, \widetilde{\varphi}^{(2)}, \ldots\right) \quad \text{as} \quad l \to \infty. \tag{7.26}$$

Going back to (7.22) we can employ (7.23) and (7.26) to see that in $\mathfrak{M}([0,\infty]^{\mathbb{N}_0})$,

$$\begin{aligned} \mathrm{law}_{\nu_l}(R_l) &= (1-\theta_l)\,\mathrm{law}_{\nu_l^\bullet}(R_l) + \theta_l\,\mathrm{law}_{\nu_l^\circ}(R_l) \\ &\to (1-\theta)\,\mathrm{law}\left(0, \boldsymbol{\sigma}\widetilde{\Phi}\right) + \theta\,\mathrm{law}(\Phi) \quad \text{as} \quad l \to \infty. \end{aligned} \tag{7.27}$$

(iii) On the other hand, (6.3) guarantees that $D_{[0,\infty]^{\mathbb{N}_0}}(\mathrm{law}_{\nu_l}(R_l), \mathrm{law}_{\mu_{A_l}}(R_l)) < \varepsilon$ for all l, and hence by (7.20) that $D_{[0,\infty]^{\mathbb{N}_0}}(\mathrm{law}_{\nu_l}(R_l), \mathrm{law}(\widetilde{\Phi})) \leqslant \varepsilon$. Together with (7.27) this proves that

$$D_{[0,\infty]^{\mathbb{N}_0}}\left((1-\theta)\,\mathrm{law}\left(0, \boldsymbol{\sigma}\widetilde{\Phi}\right) + \theta\,\mathrm{law}(\Phi), \mathrm{law}\left(\widetilde{\Phi}\right)\right) \leqslant \varepsilon. \tag{7.28}$$

But $\varepsilon > 0$ was arbitrary, and therefore

$$\mathrm{law}(\widetilde{\Phi}) = (1-\theta)\,\mathrm{law}\left(0, \boldsymbol{\sigma}\widetilde{\Phi}\right) + \theta\,\mathrm{law}(\Phi). \tag{7.29}$$

For the finite dimensional distribution functions $\{F^{[d]}\}$ and $\{\widetilde{F}^{[d]}\}$ of Φ and $\widetilde{\Phi}$ this means that for all $d \geqslant 0$ and $t_0, t_1, \ldots, t_d \geqslant 0$,

$$\widetilde{F}^{[d+1]}(t_0, t_1, \ldots, t_d) = (1-\theta)\,\widetilde{F}^{[d]}(t_1, \ldots, t_d) + \theta F^{[d+1]}(t_0, t_1, \ldots, t_d). \tag{7.30}$$

However, because of (7.20), $\{F^{[d]}\}$ and $\{\widetilde{F}^{[d]}\}$ are related to each other as in (7.4) of Theorem 7.4. Together with (7.30) the latter shows that $\{\widetilde{F}^{[d]}\}$ satisfies condition (7.18), and (7.21) follows by Proposition 7.6 (b). □

8. Proofs for local processes

Throughout this section, the fixed compact metric space $(\mathfrak{Z}, d_{\mathfrak{Z}})$ is the state space for local observables.

8.1. The general asymptotic local process

To prepare the proof of Theorem 4.3, we first establish an approximation result.

Proposition 8.1 (Approximating d-dimensional marginals of a stationary sequence). — *Let T be an ergodic measure preserving map on the nonatomic probability space $(X, \mathcal{A}, \mu)$, and let $\widehat{\Psi}$ be an $\mathfrak{Z}$-valued stationary sequence which only assumes finitely many different values. Then, for any $d \geqslant 1$ and $\varepsilon > 0$, there is some measurable $\psi : X \to \mathfrak{Z}$ such that the sequence $\Psi := (\psi, \psi \circ T, \ldots)$ satisfies*

$$D_{\mathfrak{Z}^d}\left(\pi^d_*\left(\mathrm{law}_\mu(\Psi)\right), \pi^d_*\left(\mathrm{law}\left(\widehat{\Psi}\right)\right)\right) < \varepsilon. \tag{8.1}$$

Proof. —
(i) Write $\widehat{\Psi} = (\widehat{\psi}^{(0)}, \widehat{\psi}^{(1)}, \ldots)$ and let $F \subseteq \mathfrak{Z}$ be a finite set such that $\widehat{\psi}^{(0)} \in F$ a.s. Fix d and ε, and pick one particular element $y_* \in F$. It suffices to show that we can construct a local observable ψ and an arbitrarily large subset Y of X such that the d-dimensional marginal of Ψ, when conditioned on Y, coincides with the marginal of $\widehat{\Psi}$. That is, we prove that for every $\delta > 0$ there is some $Y \in \mathcal{A}$ with $\mu(Y^c) < \delta$, and a measurable $\psi : X \to \mathfrak{Z}$ such that

$$\pi^d_*\left(\mathrm{law}_{\mu_Y}(\Psi)\right) = \pi^d_*\left(\mathrm{law}\left(\widehat{\Psi}\right)\right). \tag{8.2}$$

(ii) Apply the classical Rokhlin Lemma (as in [Zwe16, Lemma 7.4]) to obtain a Rokhlin tower $(X_i)_{i=0}^I$ of height $I > 2d/\delta$ and with $\mu(X \setminus \bigcup_{i=0}^I X_i) < \delta/2$. This means that the X_i are pairwise disjoint and $X_i = T^{-(I-i)} X_I$ for $i \in \{0, \ldots, I\}$. Conditioning on the top level X_I of the tower we obtain the probability space $(X_I, X_I \cap \mathcal{A}, \mu_{X_I})$. Being nonatomic, it admits a partition into measurable sets,

$$X_I = \bigcup\nolimits_{(y_0, \ldots, y_I) \in F^{I+1}} X_I\left(y_0, \ldots, y_I\right) \quad \text{(disjoint)},$$

with $\mu_{X_I}(X_I(y_0, \ldots, y_I)) = \Pr[(\widehat{\psi}^{(0)}, \ldots, \widehat{\psi}^{(I)}) = (y_0, \ldots, y_I)]$. We define partitions of the other levels X_i, $i \in \{0, \ldots, I-1\}$,

$$X_i = \bigcup\nolimits_{(y_0, \ldots, y_I) \in F^{I+1}} X_i(y_0, \ldots, y_I) \quad \text{(disjoint)},$$

by setting

$$X_i\left(y_0, \ldots, y_I\right) := T^{-1} X_{i+1}\left(y_0, \ldots, y_I\right) = \ldots = T^{-(I-i)} X_I\left(y_0, \ldots, y_I\right).$$

Finally, define a measurable function $\psi : X \to \mathfrak{Z}$ through

$$\psi := \begin{cases} y_i & \text{on } X_i\left(y_0, \ldots, y_I\right), 0 \leqslant i \leqslant I, \\ y_* & \text{otherwise.} \end{cases} \tag{8.3}$$

Then, for any $(y_0, \ldots, y_I) \in F^{I+1}$ and $j \in \{0, 1, \ldots I-i\}$,

$$\psi \circ T^j = y_{i+j} \quad \text{on} \quad X_i\left(y_0, \ldots, y_I\right). \tag{8.4}$$

(iii) As a consequence of (8.4) we get, for $i \in \{0, \ldots, I-d+1\}$, $(z_0, \ldots, z_{d-1}) \in F^d$, a decomposition

$$X_i \cap \left\{\left(\psi, \ldots, \psi \circ T^{d-1}\right) = (z_0, \ldots, z_{d-1})\right\}$$
$$= \bigcup_{\substack{(y_0, \ldots, y_I) \in F^{I+1}: \\ (y_i, \ldots, y_{i+d-1}) = (z_0, \ldots, z_{d-1})}} X_i\left(y_0, \ldots, y_I\right)$$
$$= \bigcup_{\substack{(y_0, \ldots, y_I) \in F^{I+1}: \\ (y_i, \ldots, y_{i+d-1}) = (z_0, \ldots, z_{d-1})}} T^{-(I-i)} X_I\left(y_0, \ldots, y_I\right).$$

Therefore, as T preserves μ and $X_i = T^{-(I-i)} X_I$, we se that

$$\mu_{X_i}\left(\left(\psi, \ldots, \psi \circ T^{d-1}\right) = (z_0, \ldots, z_{d-1})\right) =$$
$$= \sum_{\substack{(y_0, \ldots, y_I) \in F^{I+1}: \\ (y_i, \ldots, y_{i+d-1}) = (z_0, \ldots, z_{d-1})}} \mu_{X_i}\left(T^{-(I-i)} X_I\left(y_0, \ldots, y_I\right)\right)$$
$$= \sum_{\substack{(y_0, \ldots, y_I) \in F^{I+1}: \\ (y_i, \ldots, y_{i+d-1}) = (z_0, \ldots, z_{d-1})}} \mu_{X_I}\left(X_I\left(y_0, \ldots, y_I\right)\right)$$
$$= \sum_{\substack{(y_0, \ldots, y_I) \in F^{I+1}: \\ (y_i, \ldots, y_{i+d-1}) = (z_0, \ldots, z_{d-1})}} \Pr\left[\left(\widehat{\psi}^{(0)}, \ldots, \widehat{\psi}^{(I)}\right) = (y_0, \ldots, y_I)\right]$$
$$= \Pr\left[\left(\widehat{\psi}^{(i)}, \ldots, \widehat{\psi}^{(i+d-1)}\right) = (z_0, \ldots, z_{d-1})\right]$$
$$= \Pr\left[\left(\widehat{\psi}^{(0)}, \ldots, \widehat{\psi}^{(d-1)}\right) = (z_0, \ldots, z_{d-1})\right],$$

meaning that

$$\text{(8.5)} \quad \operatorname{law}_{\mu_{X_i}}\left(\left(\psi, \ldots, \psi \circ T^{d-1}\right)\right) = \operatorname{law}\left(\left(\widehat{\psi}^{(0)}, \ldots, \widehat{\psi}^{(d-1)}\right)\right) \quad \text{for} \quad 0 \leqslant i \leqslant I-d+1.$$

Hence, taking $Y := \bigcup_{i=0}^{I-d+1} X_i$ we have

$$\mu(Y^c) = \mu\left(\bigcup_{i=I-d+2}^{I} X_i\right) + \mu\left(X \setminus \bigcup_{i=0}^{I} X_i\right) < d/I + \delta/2 < \delta$$

and

$$\operatorname{law}_{\mu_Y}\left(\left(\psi, \ldots, \psi \circ T^{d-1}\right)\right) = \operatorname{law}\left(\left(\widehat{\psi}^{(0)}, \ldots, \widehat{\psi}^{(d-1)}\right)\right),$$

as required in (8.2) above. □

We can now turn to the

Proof of Theorem 4.3. Let (A_l) and $\widetilde{\Psi} = (\widetilde{\psi}^{(0)}, \widetilde{\psi}^{(1)}, \ldots)$ be given. For every $l \geqslant 1$ there is some finite set $F_l \subseteq \mathfrak{Z}$ and a Borel measurable map $\theta_l : \mathfrak{Z} \to F_l$ such that $d_{\mathfrak{Z}}(\mathrm{Id}_{\mathfrak{Z}}, \theta_l) < 1/l$ on $\mathfrak{Z}$. Setting $\widehat{\Psi}_l := (\theta_l \circ \widetilde{\psi}^{(0)}, \theta_l \circ \widetilde{\psi}^{(1)}, \ldots)$ we obtain a stationary sequence in $\mathfrak{Z}$ which only assumes finitely many values and satisfies $d_{\mathfrak{Z}}(\widehat{\Psi}_l, \widetilde{\Psi}) < 1/l$ on the underlying probability space. Therefore,

$$\text{(8.6)} \qquad D_{\mathfrak{Z}^{\mathbb{N}_0}}\left(\operatorname{law}\left(\widehat{\Psi}_l\right), \operatorname{law}\left(\widetilde{\Psi}\right)\right) \longrightarrow 0 \quad \text{as} \quad l \to \infty.$$

Due to (8.6) it suffices to construct ψ_{A_l}, $l \geqslant 1$, such that the corresponding local processes $\widetilde{\Psi}_{A_l}$ approximate the $\widehat{\Psi}_l$ and satisfy

$$\text{(8.7)} \qquad D_{\mathfrak{Z}^{\mathbb{N}_0}}\left(\operatorname{law}_{\mu_{A_l}}\left(\widetilde{\Psi}_{A_l}\right), \operatorname{law}\left(\widehat{\Psi}_l\right)\right) \longrightarrow 0 \quad \text{as} \quad l \to \infty,$$

or, equivalently, that for every $d \geqslant 1$,

$$D_{\mathfrak{Z}^d}(\pi_*^d(\mathrm{law}_{\mu_{A_l}}(\widetilde{\Psi}_{A_l})), \pi_*^d(\mathrm{law}(\widehat{\Psi}_l))) \longrightarrow 0 \quad \text{as} \quad l \to \infty. \tag{8.8}$$

For each $l \geqslant 1$ apply Proposition 8.1 to $(A_l, A_l \cap \mathcal{A}, \mu_{A_l}, T_{A_l})$, $d := l$, $\varepsilon := 1/l$, and $\widehat{\Psi} := \widehat{\Psi}_l$ to obtain an local observable $\psi_{A_l} : A_l \to \mathfrak{Z}$ for A_l for which

$$D_{\mathfrak{Z}^d}\left(\pi_*^l\left(\mathrm{law}_{\mu_{A_l}}\left(\widetilde{\Psi}_{A_l}\right)\right), \pi_*^l\left(\mathrm{law}\left(\widehat{\Psi}_l\right)\right)\right) < 1/l. \tag{8.9}$$

Then (8.8) follows since $D_{\mathfrak{Z}^d}(\pi_*^d(\mathrm{law}(\Psi'')), \pi_*^d(\mathrm{law}(\Psi')))$ is non-decreasing in d for all random sequences Ψ', Ψ'' in $\mathfrak{Z}$. □

8.2. Towards specific limit processes

To get started, we record some basic properties of local processes.

Proposition 8.2 (Asymptotic invariance and admissible delays for (Ψ_{A_l})). — *Let T be a measure-preserving map on the probability space $(X, \mathcal{A}, \mu)$.*

(a) *For every set $A \in \mathcal{A}$, any local process Ψ_A on A, and any $m \geqslant 0$,*

$$\Psi_A \circ T^m = \Psi_A \quad \text{on} \quad \{\varphi_A > m\}\,. \tag{8.10}$$

(b) *Suppose that $(A_l)_{l\geqslant 1}$ is a sequence of asymptotically rare events, and $(\psi_{A_l})_{l\geqslant 1}$ a sequence of local observables for the A_l, with corresponding local processes Ψ_{A_l}. Set $R_l := \Psi_{A_l} : X \to \mathfrak{Z}^{\mathbb{N}}$. Then (R_l) is asymptotically T-invariant in measure.*

(c) *Let (ν_l) be a sequence in $\mathfrak{P}$, and let the measurable maps $\tau_l : X \to \mathbb{N}_0$ satisfy*

$$\nu_l\left(\tau_l < \varphi_{A_l}\right) \longrightarrow 1 \quad \text{as} \quad l \to \infty. \tag{8.11}$$

Then (τ_l) is an admissible delay sequence for (R_l) and (ν_l).

Proof. — Statement (a) is immediate from the fact that $T_A \circ T^m = T_A$ on $\{\varphi_A > m\}$ for every $m \geqslant 0$. Next, (A_l) being asymptotically rare means that $\mu(\varphi_{A_l} > m) \to 1$ as $l \to \infty$. In particular, $\mu(d_{\mathfrak{Z}^{\mathbb{N}}}(R_l \circ T, R_l) > 0) \leqslant \mu(\Psi_{A_l} \circ T \neq \Psi_{A_l}) \leqslant \mu(\varphi_{A_l} = 1) \to 0$, proving (b). Turning to (c) we note that (8.10) entails

$$\Psi_{A_l} \circ T^{\tau_l} = \Psi_{A_l} \quad \text{on} \quad \{\varphi_{A_l} > \tau_l\}\,,$$

whence $\nu_l(d_{\mathfrak{Z}^{\mathbb{N}}}(R_l \circ T^{\tau_l}, R_l) > 0) \leqslant \nu_l(\varphi_{A_l} \leqslant \tau_l) \to 0$, validating the sufficient condition (6.23). □

Statement (b) shows that Proposition 4.5 is a special case of Theorem 6.1 with $R_l := \Psi_{A_l}$. We can now supply the easy

Proof of Theorem 4.7. — Set $R_l := \Psi_{A_l}$, $l \geqslant 1$. By Proposition 8.2 (b) and (c) and condition (4.9), (R_l) is asymptotically T-invariant in measure, and (τ_l) is an admissible delay sequence for (R_l) and (ν_l). Now (4.10) allows us to apply Proposition 6.5. □

Next we turn to the

Proof of Theorem 4.8. —
(i) For every $k \geqslant 1$ choose a sequence $(\nu_l^{(k)})_{l\geqslant 1}$ in $\mathfrak{P}$ which satisfies the assumptions on $(\nu_l)_{l\geqslant 1}$ in the final paragraph of the theorem with $d_{\mathfrak{P}}(\nu_l^{(k)}, \mu_{A_l}) < 1/k$ for all l. By compactness of $(\mathfrak{M}(\mathfrak{Z}^{\mathbb{N}_0}), D_{\mathfrak{Z}^{\mathbb{N}_0}})$ and a diagonalization argument we may assume w.l.o.g. that we work with a subsequence along which we have distributional convergence for all measures involved. Specifically, assume that there are random sequences $\widetilde{\Psi} = (\psi^{(0)}, \psi^{(1)}, \ldots)$ and $\widetilde{\Psi}^{(k)} = (\psi^{(0,k)}, \psi^{(1,k)}, \ldots)$, $k \geqslant 1$, in $\mathfrak{Z}$ such that

$$\widetilde{\Psi}_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \widetilde{\Psi} \quad \text{and} \quad \widetilde{\Psi}_{A_l} \overset{\nu_l^{(k)}}{\Longrightarrow} \widetilde{\Psi}^{(k)} \quad \text{for all} \quad k \geqslant 1 \quad \text{as} \quad l \to \infty. \tag{8.12}$$

Due to (4.4) we know that $\widetilde{\Psi}$ is stationary, obviously with $\mathrm{law}(\psi^{(0)}) = \mathrm{law}(\psi)$. The main point is to show that $\widetilde{\Psi}$ is in fact iid.

To this end, write $\boldsymbol{\sigma}\widetilde{\Psi} := (\psi^{(1)}, \psi^{(2)}, \ldots)$ for the shifted version of $\widetilde{\Psi}$, and regard $\widetilde{\Psi}$ as the random element $(\psi^{(0)}, \boldsymbol{\sigma}\widetilde{\Psi})$ of $\mathfrak{Z} \times \mathfrak{Z}^{\mathbb{N}}$. Due to (8.12) we have

$$(\psi_{A_l}, \Psi_{A_l}) \overset{\mu_{A_l}}{\Longrightarrow} \left(\psi^{(0)}, \boldsymbol{\sigma}\widetilde{\Psi}\right) \quad \text{as} \quad l \to \infty. \tag{8.13}$$

Since $\widetilde{\Psi}$ is stationary, we know it is in fact iid as soon as

$$\left(\psi^{(0)}, \boldsymbol{\sigma}\widetilde{\Psi}\right) \text{ is an independent pair.} \tag{8.14}$$

(ii) For every $k \geqslant 1$ we can employ Theorem 6.7 with $\mathfrak{E} := \mathfrak{Z}$, $\mathfrak{E}' := \mathfrak{Z}^{\mathbb{N}}$, $\nu_l := \nu_l^{(k)}$, $R_l := \psi_{A_l}$, $R_l' := \Psi_{A_l}$, and $\mathcal{B}_{\mathfrak{E}}^{\pi} := \mathcal{B}_{\mathfrak{Z}}^{\pi}$. Indeed, by Proposition 8.2 b), the sequence (R_l') is asymptotically T-invariant in measure.

Fix k, take any $F \in \mathcal{B}_{\mathfrak{Z}}^{\pi}$ with $\Pr[\psi \in F] > 0$, and pick $(\nu_{l,F})$, $(\tau_{l,F})$ and $\mathfrak{K}_F$ for $(\nu_l^{(k)})$ as in the assumption of Theorem 4.8. Observe that via Proposition 8.2 (c) our assumption (4.13) ensures that $(\tau_{l,F})$ is always an admissible delay sequence for (R_l') and the $\nu_{l,F}$. Thus, Theorem 6.7 shows that for every $k \geqslant 1$,

$$\left(\psi^{(0,k)}, \boldsymbol{\sigma}\widetilde{\Psi}^{(k)}\right) \text{ is an independent pair.} \tag{8.15}$$

But since $d_{\mathfrak{P}}(\nu_l^{(k)}, \mu_{A_l}) < 1/k$ for all l, it is clear that

$$\widetilde{\Psi}^{(k)} = \left(\psi^{(0,k)}, \boldsymbol{\sigma}\widetilde{\Psi}^{(k)}\right) \Longrightarrow \left(\psi^{(0)}, \boldsymbol{\sigma}\widetilde{\Psi}\right) \quad \text{as} \quad k \to \infty. \tag{8.16}$$

Together with (8.15) this proves (8.14).
(iii) The above shows that $\widetilde{\Psi}_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \Psi^*$, and hence also $\Psi_{A_l} = \boldsymbol{\sigma}\widetilde{\Psi}_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \boldsymbol{\sigma}\Psi^* \overset{d}{=} \Psi^*$. From (8.12) we see that

$$\Psi_{A_l} \overset{\nu_l^{(k)}}{\Longrightarrow} \boldsymbol{\sigma}\widetilde{\Psi}^{(k)}$$

for each k, and applying Theorem 4.7 with

$$\nu_l := \nu_l^{(k)} = \left(\nu_l^{(k)}\right)_{\{\psi_{A_l} \in \mathfrak{Z}\}} \quad \text{and} \quad \tau_l := \tau_{l,\mathfrak{Z}},$$

we conclude that $\Psi_{A_l} \overset{\mu}{\Longrightarrow} \boldsymbol{\sigma}\widetilde{\Psi}^{(k)}$ for every k. But then all these limit processes have the same law, and hence also the same law as their distributional limit $\boldsymbol{\sigma}\widetilde{\Psi} \overset{d}{=} \Psi^*$. This proves $\Psi_{A_l} \overset{\mu}{\Longrightarrow} \Psi^*$. □

8.3. Robustness

To conclude this section, we provide a

Proof of Theorem 4.10. — We assume w.l.o.g. that $A_l' \subseteq A_l$ for all l. (Otherwise we can apply this partial result to compare either of (ψ_{A_l}) and $(\psi'_{A_l'})$ to $(\psi_{A_l}\,|_{A_l \cap A_l'})$.) Our goal is to prove that

$$D_{3^{\mathbb{N}_0}}\left(\mathrm{law}_{\mu_{A_l}}\left(\widetilde{\Psi}_{A_l}\right), \mathrm{law}_{\mu_{A_l'}}\left(\widetilde{\Psi}'_{A_l'}\right)\right) \longrightarrow 0 \quad \text{as} \quad l \to \infty. \tag{8.17}$$

(i) We first note that our asumptions guarantee

$$D_3\left(\mathrm{law}_{\mu_{A_l}}\left(\psi_{A_l}\right), \mathrm{law}_{\mu_{A_l'}}\left(\psi'_{A_l'}\right)\right) \longrightarrow 0 \quad \text{as} \quad l \to \infty. \tag{8.18}$$

For this, decompose

$$\mathrm{law}_{\mu_{A_l}}\left(\psi_{A_l}\right) = \mu_{A_l}\left(A_l'\right)\,\mathrm{law}_{\mu_{A_l'}}\left(\psi_{A_l}\right) + \mu_{A_l}\left(A_l \setminus A_l'\right)\,\mathrm{law}_{\mu_{A_l \setminus A_l'}}\left(\psi_{A_l}\right)$$

and

$$\mathrm{law}_{\mu_{A_l'}}\left(\psi'_{A_l'}\right) = \mu_{A_l}\left(A_l'\right)\,\mathrm{law}_{\mu_{A_l'}}\left(\psi'_{A_l'}\right) + \mu_{A_l}\left(A_l \setminus A_l'\right)\,\mathrm{law}_{\mu_{A_l'}}\left(\psi'_{A_l'}\right).$$

By (6.2),

$$D_3\left(\mathrm{law}_{\mu_{A_l}}\left(\psi_{A_l}\right), \mathrm{law}_{\mu_{A_l'}}\left(\psi'_{A_l'}\right)\right) \leqslant D_3\left(\mathrm{law}_{\mu_{A_l'}}\left(\psi_{A_l}\right), \mathrm{law}_{\mu_{A_l'}}\left(\psi'_{A_l'}\right)\right) + \mu_{A_l}\left(A_l \setminus A_l'\right),$$

and this bound tends to zero as $l \to \infty$, recall (6.5).

(ii) To prepare for the process version (8.17) of (8.18), we show, for every $j \geqslant 0$,

$$\mu_{A_l'}\left(T_{A_l}^j \neq T_{A_l'}^j\right) \longrightarrow 0 \quad \text{as} \quad l \to \infty. \tag{8.19}$$

Indeed, whenever $A' \subseteq A$, then $A' \cap \{T_A^j = T_{A'}^j\} \supseteq \bigcap_{i=0}^j T_A^{-i} A'$. Therefore,

$$\begin{aligned}\mu_A\left(A' \cap \left\{T_A^j \neq T_{A'}^j\right\}\right) &\leqslant \mu_A\left(\bigcup_{i=0}^j T_A^{-i}\left(A \setminus A'\right)\right) \\ &\leqslant \sum_{i=0}^j \mu_A\left(T_A^{-i}\left(A \setminus A'\right)\right) = (j+1)\mu_A\left(A \setminus A'\right),\end{aligned}$$

since T_A preserves μ_A. Applying this to $A_l' \subseteq A_l$ yields (8.19), as $\mu_{A_l}(A_l \setminus A_l') \to 0$.

We use this to check that for every $j \geqslant 0$,

$$d_3\left(\psi_{A_l} \circ T_{A_l}^j, \psi'_{A_l'} \circ T_{A'}^j\right) \xrightarrow{\mu_{A_l'}} 0 \quad \text{as} \quad l \to \infty. \tag{8.20}$$

Take any $\eta > 0$, then

$$\begin{aligned}&A_l' \cap \left\{d_3\left(\psi_{A_l} \circ T_{A_l}^j, \psi'_{A_l'} \circ T_{A'}^j\right) \geqslant \eta\right\} \\ &\qquad \subseteq \left(A_l' \cap \left\{T_{A_l}^j \neq T_{A_l'}^j\right\}\right) \cup \left(A_l' \cap T_{A'}^{-j}\left\{d_3\left(\psi_{A_l}, \psi'_{A_l'}\right) \geqslant \eta\right\}\right).\end{aligned}$$

As a consequence of this and the fact that $T_{A_l'}$ preserves $\mu_{A_l'}$ we find that

$$\mu_{A_l'}\left(d_3\left(\psi_{A_l} \circ T_{A_l}^j, \psi'_{A_l'} \circ T_{A'}^j\right) \geqslant \eta\right) \leqslant \mu_{A_l'}\left(T_{A_l}^j \neq T_{A_l'}^j\right) + \mu_{A_l'}\left(d_3\left(\psi_{A_l}, \psi'_{A_l'}\right) \geqslant \eta\right),$$

which tends to zero due to (8.19) and assumption (4.17). This proves (8.20).

(iii) Now recall that we can regard the local processes $\widetilde{\Psi}_{A_l}$ and $\widetilde{\Psi}'_{A'_l}$ as local observables taking values in $\mathfrak{Z}^{\mathbb{N}_0}$. Therefore our assertion (8.17) will follow from the weaker version (8.18) of the present theorem which has already been established in step (i), as soon as we validate

$$d_{\mathfrak{Z}^{\mathbb{N}_0}}\left(\widetilde{\Psi}_{A_l}, \widetilde{\Psi}'_{A'_l}\right) \xrightarrow{\mu_{A'_l}} 0 \quad \text{as} \quad l \to \infty. \tag{8.21}$$

Given $\varepsilon > 0$ choose $J \geqslant 1$ so large that $2^{-J}\text{diam}(\mathfrak{Z}) < \varepsilon/2$. Then, for all $l \geqslant 1$,

$$\begin{aligned} &d_{\mathfrak{Z}^{\mathbb{N}_0}}\left(\widetilde{\Psi}_{A_l}, \widetilde{\Psi}'_{A'_l}\right) \\ &\quad = \sum_{j\geqslant 0} 2^{-(j+1)} d_{\mathfrak{Z}}\left(\psi_{A_l} \circ T^j_{A_l}, \psi'_{A'_l} \circ T^j_{A'}\right) \leqslant \sum_{j=0}^{J} d_{\mathfrak{Z}}\left(\psi_{A_l} \circ T^j_{A_l}, \psi'_{A'_l} \circ T^j_{A'}\right) + \varepsilon/2, \end{aligned}$$

which in view of (8.20) leads to (8.21) and thus gives (8.17). □

9. Proofs for joint processes

We start with the easy

Proof of Proposition 5.1. — Since both $(\mu(A_l)\Phi_{A_l})_{l\geqslant 1}$ and $(\Psi_{A_l})_{l\geqslant 1}$ are asymptotically T-invariant in measure as sequences in $[0,\infty]^{\mathbb{N}}$ and $\mathfrak{Z}^{\mathbb{N}}$ respectively (Proposition 7.1 (b) and Proposition 8.2 (b)), we immediately see that $(\mu(A_l)\Phi_{A_l}, \Psi_{A_l})_{l\geqslant 1}$ is asymptotically T-invariant in measure as a sequence in $[0,\infty]^{\mathbb{N}} \times \mathfrak{Z}^{\mathbb{N}}$. Now use Theorem 6.1. □

Now compare the laws of joint processes under μ and under the μ_{A_l}.

Proof of Theorem 5.2. —

(i) Due to stationarity of $(\mu(A_l)\Phi_{A_l}, \widetilde{\Psi}_{A_l})$ under μ_{A_l}, it is easy to see that (5.4) is actually equivalent to the formally weaker statement (obtained by forgetting about the first entry of $\widetilde{\Psi}_{A_l}$)

$$R_l \overset{\mu_{A_l}}{\Longrightarrow} (\Phi_{\text{Exp}}, \Psi^*) \quad \text{as} \quad l \to \infty. \tag{9.1}$$

where $R_l := (\mu(A_l)\Phi_{A_l}, \Psi_{A_l}) : X \to [0,\infty]^{\mathbb{N}_0} \times \mathfrak{Z}^{\mathbb{N}} =: \mathfrak{E}$ (equipped with $d_{\mathfrak{E}} := d_{[0,\infty]^{\mathbb{N}_0}} + d_{\mathfrak{Z}}$). The standard subsequence argument based on compactness of $\mathfrak{E}$ shows that we can assume w.l.o.g. that there are random elements (Φ, Ψ) and $(\overline{\Phi}, \overline{\Psi})$ of $\mathfrak{E}$, with $\Phi = (\varphi^{(i)})_{i\geqslant 0}$, $\overline{\Phi} = (\overline{\varphi}^{(i)})_{i\geqslant 0}$, $\Psi = (\psi^{(i)})_{i\geqslant 0}$, and $\overline{\Psi} = (\overline{\psi}^{(i)})_{i\geqslant 0}$, such that

$$R_l \overset{\mu}{\Longrightarrow} (\Phi, \Psi) \quad \text{and} \quad R_l \overset{\mu_{A_l}}{\Longrightarrow} \left(\overline{\Phi}, \overline{\Psi}\right) \quad \text{as} \quad l \to \infty. \tag{9.2}$$

To prove the theorem, we now assume that

$$\text{one of } (\Phi, \Psi) \text{ and } \left(\overline{\Phi}, \overline{\Psi}\right) \text{ has the law of } (\Phi_{\text{Exp}}, \Psi^*). \tag{9.3}$$

In view of (9.2) and (5.1) & (5.2), Theorem 4.7 then shows that $\Psi \overset{d}{=} \overline{\Psi} \overset{d}{=} \Psi^*$. Similarly, Theorem 7.4 and Proposition 7.5 together show that if one of Φ and $\overline{\Phi}$ has the distribution of Φ_{Exp}, then so does the other. Hence, $\Phi \overset{d}{=} \overline{\Phi} \overset{d}{=} \Phi^*$ as well, but it is not immediate that both (Φ, Ψ) and $(\overline{\Phi}, \overline{\Psi})$ are independent pairs.

To prove that in fact $(\Phi,\Psi)\stackrel{d}{=}(\overline{\Phi},\overline{\Psi})\stackrel{d}{=}(\Phi_{\mathrm{Exp}},\Psi^*)$, we will first check that

$$(9.4)\qquad \begin{gathered}(\mu(A_l)\Phi_{A_l}\circ T_{A_l},\Psi_{A_l})\stackrel{\nu_l}{\Longrightarrow}(\Phi_{\mathrm{Exp}},\Psi^*)\quad\text{as}\quad l\to\infty\\ \text{holds for } (\nu_l)=(\mu) \text{ iff it holds for } (\nu_l)=(\mu_{A_l}).\end{gathered}$$

Once this is established, we show that the one component $\mu(A_l)\varphi_{A_l}$ of R_l missing in (9.4) is asymptotically independent of the rest: Writing $\boldsymbol{\sigma}\Phi:=(\varphi^{(i+1)})_{i\geqslant 0}$ and $\boldsymbol{\sigma}\overline{\Phi}:=(\overline{\varphi}^{(i+1)})_{i\geqslant 0}$ for the shifted versions of Φ and $\overline{\Phi}$ respectively, we claim that

$$(9.5)\qquad \text{if }\left(\overline{\Phi},\overline{\Psi}\right)\stackrel{d}{=}(\Phi_{\mathrm{Exp}},\Psi^*)\,,\text{ then }\varphi^{(0)}\text{ is independent of }(\boldsymbol{\sigma}\Phi,\Psi),$$

whereas

$$(9.6)\qquad \text{if }(\Phi,\Psi)\stackrel{d}{=}(\Phi_{\mathrm{Exp}},\Psi^*)\,,\text{ then }\overline{\varphi}^{(0)}\text{ is independent of }\left(\boldsymbol{\sigma}\overline{\Phi},\overline{\Psi}\right),$$

Together, assertions (9.4) - (9.6) prove our theorem.

(ii) Validating (9.4) is straightforward: Set $R'_l:=(\mu(A_l)\Phi_{A_l}\circ T_{A_l},\Psi_{A_l})$, $l\geqslant 1$. According to Proposition 7.1 (c) and Proposition 8.2 (b), the sequence (R'_l) is asymptotically T-invariant in measure. Recalling Proposition 7.2 (b) and 8.2 (c) , we see that (τ_l) is an admissible delay sequence for (R'_l). Now (5.2) allows us to appeal to Proposition 6.5 to complete the proof of (9.4).

(iii) Preparing for the proof of (9.5) and (9.6) we set, for $M\in\mathcal{B}_{\mathfrak{E}}$, $B_l(M):=\{(\mu(A_l)\Phi_{A_l},\widetilde{\Psi}_{A_l})\in M\}\in\mathcal{A}$. Observe then that due to $(\mu(A)\Phi_A\circ T_A,\Psi_A)=(\mu(A)\Phi_A,\widetilde{\Psi}_A)\circ T_A$, independence of $\varphi^{(0)}$ and $(\boldsymbol{\sigma}\Phi,\Psi)$ follows if we show that

$$(9.7)\qquad \begin{gathered}\mu\left(\{\mu(A_l)\varphi_{A_l}\leqslant t\}\cap T_{A_l}^{-1}B_l(M)\right)\longrightarrow\left(1-e^{-t}\right)\Pr\left[(\Phi_{\mathrm{Exp}},\Psi^*)\in M\right]\\ \text{for } t>0 \text{ and } M\in\mathcal{B}_{\mathfrak{E}} \text{ with } \Pr\left[(\Phi_{\mathrm{Exp}},\Psi^*)\in\partial M\right]=0.\end{gathered}$$

(Use a variant of [Bil99, Theorem 2.3] to argue as in the proof of [Bil99, Theorem 2.8].) Analogously, independence of $\overline{\varphi}^{(0)}$ and $(\boldsymbol{\sigma}\overline{\Phi},\overline{\Psi})$ is immediate if

$$(9.8)\qquad \begin{gathered}\mu_A\left(\{\mu(A_l)\varphi_{A_l}>s\}\cap T_{A_l}^{-1}B_l(M)\right)\longrightarrow e^{-s}\Pr\left[(\Phi_{\mathrm{Exp}},\Psi^*)\in M\right]\\ \text{for } s>0 \text{ and } M\in\mathcal{B}_{\mathfrak{E}} \text{ with } \Pr\left[(\Phi_{\mathrm{Exp}},\Psi^*)\in\partial M\right]=0.\end{gathered}$$

Now [Zwe16, Lemma 4.1] shows that for any $A,B\in\mathcal{A}$ and $t\geqslant 0$,

$$(9.9)\qquad \left|\int_0^t\mu_A\left(\{\mu(A)\varphi_A>s\}\cap T_A^{-1}B\right)\,ds-\mu\left(\{\mu(A)\varphi_A\leqslant t\}\cap T_A^{-1}B\right)\right|\leqslant\mu(A).$$

Hence, taking $A:=A_l$ and $B:=B_l(M)$ for an arbitrary $(\Phi_{\mathrm{Exp}},\Psi^*)$-continuity set $M\in\mathcal{B}_{\mathfrak{E}}$, we see that for every $t\geqslant 0$,

$$(9.10)\qquad \begin{aligned}&\int_0^t\mu_{A_l}\left(\{\mu(A_l)\varphi_{A_l}>s\}\cap T_{A_l}^{-1}B_l(M)\right)\,ds\\ &\qquad\qquad-\mu\left(\{\mu(A_l)\varphi_{A_l}\leqslant t\}\cap T_{A_l}^{-1}B_l(M)\right)\longrightarrow 0\quad\text{as}\quad l\to\infty.\end{aligned}$$

(iv) To validate (9.5), suppose that $(\overline{\Phi},\overline{\Psi})\stackrel{d}{=}(\Phi_{\mathrm{Exp}},\Psi^*)$. This entails

$$\mu_{A_l}\left(\{\mu(A_l)\varphi_{A_l}>s\}\cap T_{A_l}^{-1}B_l(M)\right)\longrightarrow e^{-s}\Pr\left[(\Phi_{\mathrm{Exp}},\Psi^*)\in M\right]$$

whenever $s > 0$ and M is a $(\Phi_{\mathrm{Exp}}, \Psi^*)$-continuity set. But then (9.7) follows via the crucial relation (9.10) by dominated convergence, ensuring independence of $\varphi^{(0)}$ and $(\boldsymbol{\sigma}\Phi, \Psi)$ as required.

Similarly, to prove (9.6), suppose that

$$(\Phi, \Psi) \stackrel{d}{=} (\Phi_{\mathrm{Exp}}, \Psi^*) .$$

Then,

$$\mu\left(\{\mu(A_l)\varphi_{A_l} \leqslant t\} \cap T_{A_l}^{-1} B_l(M)\right) \longrightarrow \left(1 - e^{-t}\right) \Pr\left[(\Phi_{\mathrm{Exp}}, \Psi^*) \in M\right]$$

for $t > 0$ and any $(\Phi_{\mathrm{Exp}}, \Psi^*)$-continuity set M. Fixing M and varying t, we can use (9.10) once again and apply [Zwe16, Lemma 4.2] to obtain (9.8), and hence the desired independence of $\overline{\varphi}^{(0)}$ and $(\boldsymbol{\sigma}\overline{\Phi}, \overline{\Psi})$. □

The argument for the joint limit theorem elaborates on a principle used before.

Proof of Theorem 5.3. —

(i) By compactness of $(\mathfrak{M}([0,\infty]^{\mathbb{N}_0} \times \mathfrak{Z}^{\mathbb{N}_0}), D_{[0,\infty]^{\mathbb{N}_0} \times \mathfrak{Z}^{\mathbb{N}_0}})$ we may assume w.l.o.g. that

$$\left(\mu(A_l)\Phi_{A_l}, \widetilde{\Psi}_{A_l}\right) \stackrel{\mu_{A_l}}{\Longrightarrow} \left(\Phi, \widetilde{\Psi}\right) \quad \text{as} \quad l \to \infty, \tag{9.11}$$

with $\Phi = (\varphi^{(0)}, \varphi^{(1)}, \ldots)$ and $\widetilde{\Psi} = (\psi^{(0)}, \psi^{(1)}, \ldots)$ random sequences in $[0,\infty]$ and $\mathfrak{Z}$, respectively. Theorem 4.8 shows that $\widetilde{\Psi}$ is iid with $\mathrm{law}(\psi^{(0)}) = \mathrm{law}(\psi)$ so that $\mathrm{law}(\widetilde{\Psi}) = \mathrm{law}(\Psi^*)$. Next, taking $s = 0$ and observing that $A_l \cap \{\mu(A_l)\varphi_{A_l} > 0\} = A_l$, we see that Theorem 3.6 guarantees $\mathrm{law}(\Phi) = \mathrm{law}(\Phi_{\mathrm{Exp}})$.

The main point is to show that Φ and $\widetilde{\Psi}$ are independent. Let $\boldsymbol{\sigma}\Phi := (\varphi^{(1)}, \varphi^{(2)}, \ldots)$ and $\boldsymbol{\sigma}\widetilde{\Psi} := (\psi^{(1)}, \psi^{(2)}, \ldots)$ denote the shifted versions of the individual limit processes. We will first show that

$$\psi^{(0)} \text{ is independent of } \left(\Phi, \boldsymbol{\sigma}\widetilde{\Psi}\right), \tag{9.12}$$

and then check that

$$\varphi^{(0)} \text{ is independent of } \left(\boldsymbol{\sigma}\Phi, \boldsymbol{\sigma}\widetilde{\Psi}\right). \tag{9.13}$$

Together these imply that $(\psi^{(0)}, \varphi^{(0)}, (\boldsymbol{\sigma}\Phi, \boldsymbol{\sigma}\widetilde{\Psi}))$ is an independent triple. But since convergence in (9.11) uses the T_{A_l}-invariant measures μ_{A_l} under which, for each $l \geqslant 1$, $(\mu(A_l)\Phi_{A_l}, \widetilde{\Psi}_{A_l})$ is a stationary sequence in $[0,\infty] \times \mathfrak{Z}$, we see that so is $(\Phi, \widetilde{\Psi})$, meaning that $\mathrm{law}(\Phi, \widetilde{\Psi}) = \mathrm{law}(\boldsymbol{\sigma}\Phi, \boldsymbol{\sigma}\widetilde{\Psi})$. Therefore the above can be iterated to show that for any $m \geqslant 1$,

$$\left(\psi^{(0)}, \varphi^{(0)}, \ldots, \psi^{(m-1)}, \varphi^{(m-1)}, \left(\boldsymbol{\sigma}^m\Phi, \boldsymbol{\sigma}^m\widetilde{\Psi}\right)\right)$$

is an independent tuple. Hence $\{\psi^{(j)}, \varphi^{(k)}\}_{j,k \geqslant 0}$ is indeed an independent family.

(ii) To prove (9.12) we are going to apply Theorem 6.7 with $\mathfrak{E} := \mathfrak{Z}$, $\mathfrak{E}' := [0,\infty]^{\mathbb{N}_0} \times \mathfrak{Z}^{\mathbb{N}}$, $R_l := \psi_{A_l}$, $R'_l := (\mu(A_l)\Phi_{A_l}, \Psi_{A_l})$, representing $(\mu(A_l)\Phi_{A_l}, \widetilde{\Psi}_{A_l})$ as the map $(R_l, R'_l) : X \to \mathfrak{E} \times \mathfrak{E}'$. The sequence (R'_l) is asymptotically T-invariant in measure since both $(\mu(A_l)\Phi_{A_l})$ and (Ψ_{A_l}) are (due to Propositions 7.1 (b) and 8.2 (b)).

For any $F \in \mathcal{B}^{\pi}_{\mathfrak{Z}}$ with $\Pr[\psi \in F] > 0$ pick $(\nu_{l,F})$, $(\tau_{l,F})$ and $\mathfrak{K}_F$ as in the statement of Theorem 5.3. By assumption, $\overline{\nu}_{l,F} := T_*^{\tau_{l,F}} \nu_{l,F} \in \mathfrak{K}_F$ for $l \geqslant 1$, so that (9.12) follows via Theorem 6.7 once we check that

(9.14) $\quad (\tau_{l,F})_{l\geqslant 1}$ is an admissible delay for $(\mu(A_l)\Phi_{A_l}, \Psi_{A_l})_{l\geqslant 1}$ and $(\nu_{l,F})_{l\geqslant 1}$.

Here it is enough to treat $(\mu(A_l)\Phi_{A_l})$ and (Ψ_{A_l}) separately. But in the first case (5.7) and (5.8) allow us to appeal to Proposition 7.2 (a), while (5.8) alone takes care of the second case via Proposition 8.2 (c).

(iii) We validate (9.13) analogously, this time applying Theorem 6.7 with $\mathfrak{E} := [0,\infty]$, $\mathfrak{E}' := [0,\infty]^{\mathbb{N}} \times \mathfrak{Z}^{\mathbb{N}_0}$, $R_l := \mu(A_l)\varphi_{A_l}$, $R'_l := (\mu(A_l)\Phi_{A_l} \circ T_{A_l}, \Psi_{A_l})$, representing $(\mu(A_l)\Phi_{A_l}, \Psi_{A_l})$ as the map $(R_l, R'_l) : X \to \mathfrak{E} \times \mathfrak{E}'$. Again, (R'_l) is asymptotically T-invariant in measure. We use $\mathcal{B}^{\pi}_{\mathfrak{E}} := \{(s,\infty) : s \geqslant 0\}$. For any $s \geqslant 0$ we have $\Pr[\varphi^{(0)} > s] > 0$ since we already know that $\varphi^{(0)}$ has an exponential distribution. Take $(\nu_{l,s})$, $(\tau_{l,s})$ and $\mathfrak{K}_s$ as in the statement of Theorem 5.3. By assumption, $\overline{\nu}_{l,s} := T_*^{\tau_{l,s}} \nu_{l,s} \in \mathfrak{K}_s$ for $l \geqslant 1$, and (9.13) follows via Theorem 6.7 if we check that

(9.15) $\quad (\tau_{l,s})_{l\geqslant 1}$ is an admissible delay for $(\mu(A_l)\Phi_{A_l} \circ T_{A_l}, \Psi_{A_l})_{l\geqslant 1}$ and$(\nu_{l,s})_{l\geqslant 1}$.

Condition (5.6) ensures that $(\tau_{l,s})_{l\geqslant 1}$ is admissible for $(\mu(A_l)\Phi_{A_l} \circ T_{A_l})$ via Proposition 7.2 (b), and also for (Ψ_{A_l}) by virtue of Proposition 8.2 (c).

(iv) The above shows that

$$\left(\mu(A_l)\Phi_{A_l}, \widetilde{\Psi}_{A_l}\right) \overset{\mu_{A_l}}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*),$$

hence

$$(\mu(A_l)\Phi_{A_l}, \Psi_{A_l}) = \left(\mu(A_l)\Phi_{A_l}, \boldsymbol{\sigma}\widetilde{\Psi}_{A_l}\right) \overset{\mu_{A_l}}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \boldsymbol{\sigma}\Psi^*) \overset{d}{=} (\Phi_{\mathrm{Exp}}, \Psi^*).$$

To prove that we also have $(\mu(A_l)\Phi_{A_l}, \Psi_{A_l}) \overset{\mu}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*)$ we appeal to Proposition 6.5:

Set $\mathfrak{E} := [0,\infty]^{\mathbb{N}_0} \times \mathfrak{Z}^{\mathbb{N}}$, $R_l := (\mu(A_l)\Phi_{A_l}, \Psi_{A_l})$ and $\nu_l := \mu_{A_l}$. As before, (R_l) is asymptotically T-invariant in measure. Take $\tau_l := \tau_{l,\mathfrak{Z}}$ and $\mathfrak{K} := \mathfrak{K}_{\mathfrak{Z}}$ as in assumption b) of Theorem 5.3, and note that $\nu_l = \nu_{l,\mathfrak{Z}} = \mu_{A_l \cap \{\psi_{A_l} \in \mathfrak{Z}\}}$. Finally, recall that we have already shown admissibility of this sequence (τ_l) for the present (R_l) and (ν_l) in step (9), see (9.14). □

10. Illustration in some easy standard situations

We illustrate the ease with which the above results can sometimes be applied by studying some basic piecewise invertible dynamical systems.

10.1. Piecewise invertible systems

We consider situations in which (X, d_X) is a metric space with Borel σ-field $\mathcal{A} = \mathcal{B}_X$, and where X comes with a partition ξ_0 (mod λ) into open *components* (e.g. X may be a union of disjoint open intervals in $\mathbb{R}$). Let λ be a σ-finite reference measure on $\mathcal{A}$.

A *piecewise invertible system* on X is a quintuple $(X, \mathcal{A}, \lambda, T, \xi)$, where $\xi = \xi_1$ is a (finite or) countable partition mod λ of X into open sets, refining ξ_0, such that each *branch* of T, i.e. its restriction to any of its *cylinders* $Z \in \xi$ is a homeomorphism onto TZ, *null-preserving* with respect to λ, that is, $\lambda\mid_Z \circ T^{-1} \ll \lambda$. If the measure is T-invariant, we denote it by μ and call $(X, \mathcal{A}, \mu, T, \xi)$ a *measure preserving* system. The system is called *uniformly expanding* if there is some $\rho \in (0,1)$ such that $d_X(x,y) \leqslant \rho \cdot d_X(Tx,Ty)$ whenever $x, y \in Z \in \xi$.

We let ξ_n denote the family of *cylinders of rank* n, that is, the sets of the form $Z = [Z_0, \ldots, Z_{n-1}] := \bigcap_{i=0}^{n-1} T^{-i} Z_i$ with $Z_i \in \xi$. Write $\xi_n(x)$ for the element of ξ_n containing x (which is defined a.e.). Each *iterate* $(X, \mathcal{A}, \mu, T^n, \xi_n)$, $n \geqslant 1$, of the system is again piecewise invertible. The *inverse branches* will be denoted $v_Z := (T^n\mid_Z)^{-1} : T^n Z \to Z$, $Z \in \xi_n$. All v_Z have Radon–Nikodym derivatives $v'_Z := d(\lambda \circ v_Z)/d\lambda$.

The system is *Markov* if $TZ \cap Z' \neq \varnothing$ for $Z, Z' \in \xi$ implies $Z' \subseteq TZ$.

10.2. Gibbs–Markov maps

One important basic class of piecewise invertible systems $(X, \mathcal{A}, \mu, T, \xi)$ is that of probability preserving *Gibbs–Markov maps* (*GM maps*). This means that $\operatorname{diam}(X) < \infty$, and μ is an invariant probability, that the system has a uniformly expanding iterate T^N, and satisfies the *big image property*, so that $\flat := \inf_{Z\in\xi} \mu(TZ) > 0$. Moreover, the v'_Z, $Z \in \xi$, have well behaved versions in that there exists some $r > 0$ such that $|v'_Z(x)/v'_Z(y) - 1| \leqslant r\, d_X(x,y)$ whenever $x, y \in TZ$, $Z \in \xi$ (see [AD01]). In this case r can be chosen in such a way that in fact

$$\left|\frac{v'_Z(x)}{v'_Z(y)} - 1\right| \leqslant r\, d_X(x,y) \quad \text{whenever } n \geqslant 1 \quad \text{and} \quad x, y \in T^n Z, Z \in \xi_n. \tag{10.1}$$

In this context, v'_Z will always denote such versions of the a.e. defined Radon–Nikodym derivatives $d(\mu \circ v_Z)/d\mu$.

We recall a few well known basic properties of such systems, all of which are obtained by elementary routine arguments. Let β be the partition generated by $T\xi$. By (10.1) the normalized image measures $T^n_* \mu_Z$ with $n \geqslant 1$ and $Z \in \xi_n$ have densities belonging to $\mathcal{U} := \{u \in \mathcal{D}(\mu) : |u(x)/u(y) - 1| \leqslant r\, d_X(x,y) \text{ whenever } x, y \in B \in \beta\}$, that is,

$$T^n_* \mu_Z \in \mathfrak{K} \quad \text{for all} \quad n \geqslant 1 \quad \text{and} \quad Z \in \xi_n, \tag{10.2}$$

where $\mathfrak{K} := \{\nu \in \mathfrak{P} : d\nu/d\mu \in \mathcal{U}\}$. But $\mathcal{U}$ is compact in $L_1(\mu)$ (Arzela–Ascoli, as in [Aar97, § 4.7]) and convex, so that $\mathfrak{K}$ is a compact convex set in $\mathfrak{P}$.

Property (10.1) also implies *bounded distortion* in that

$$\mu_Z\left(Z \cap T^{-n} A\right) = e^{\pm r} \mu_{T^n Z}(A) \quad \text{for all} \quad n \geqslant 1, Z \in \xi_n, \quad \text{and} \quad A \in \mathcal{A}. \tag{10.3}$$

In particular,

$$\mu\left(Z \cap T^{-n} A\right) \leqslant \flat^{-1} e^r \mu(Z)\mu(A) \quad \text{for all} \quad n \geqslant 1, \quad Z \in \xi_n, \quad \text{and} \quad A \in \mathcal{A}. \tag{10.4}$$

Also, an easy argument provides constants $\kappa \geqslant 1$ and $q \in (0,1)$ such that

$$\mu(Z) \leqslant \kappa q^n \quad \text{for all} \quad n \geqslant 1 \quad \text{and} \quad Z \in \xi_n. \tag{10.5}$$

10.3. Poisson asymptotics for (unions of) cylinders of GM-maps

To demonstrate how convenient the assumptions of our limit theorems are, we first illustrate their use in the setup of cylinders of GM-maps, re-proving the well-known

THEOREM 10.1 (Poisson asymptotics for shrinking cylinders of GM maps). — *Let $(X, \mathcal{A}, \mu, T, \xi)$ be an ergodic probability preserving Gibbs–Markov system. Let $x^* \in X$ be a point such that the cylinder $A_l := \xi_l(x^*)$ is defined for all $l \geqslant 1$.*

(a) *Assume that $x^* \in X$ is not periodic, then (A_l) exhibits Poisson asymptotics,*

$$\mu(A_l)\Phi_{A_l} \overset{\mu}{\Longrightarrow} \Phi_{\mathrm{Exp}} \quad \text{and} \quad \mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \Phi_{\mathrm{Exp}} \quad \text{as} \quad l \to \infty. \tag{10.6}$$

(b) *On the other hand, if $x^* \in X$ is periodic, $x^* = T^p x^*$ with $p \geqslant 1$ minimal, then*

$$\mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \Phi_{(\mathrm{Exp},\theta)} \quad \text{as} \quad l \to \infty, \tag{10.7}$$

where $\theta := 1 - v'_{A_p}(x^*) \in (0, 1)$.

Our proof via Theorems 3.6 and 3.8 will only employ the basic elementary facts about Gibbs–Markov maps mentioned before. The well-known strong mixing properties (valid in aperiodic situations) which follow (by more sophisticated arguments) from related observations are not used.

Proof. —
(a) We are going to apply Theorem 3.6, using $\nu_l := \mu_{A_l}$, $\mathfrak{K}$ as in (10.2), and the obvious delay times $\tau_l := l$. Set $\widetilde{\kappa} := \flat^{-1} e^r \kappa$.

Condition (3.15) is trivially satisfied since $\mu(A_l)$ is exponentially small, while condition (3.17) is taken care of by (10.2).

To validate (3.16), take any $\varepsilon > 0$. Choose $K \geqslant 1$ so large that $\widetilde{\kappa} q^K/(1-q) < \varepsilon$. Since x^* is not periodic, and T is continuous on cylinders, there is some l' such that $\varphi_{A_l} > K$ on A_l whenever $l \geqslant l'$ (recall that $\mathrm{diam}(A_l) \to 0$). For every $k \in \{1, \ldots, l\}$ we have $A_l \subseteq A_k = \xi_k(x^*) \in \xi_k$, and therefore

$$\mu_{A_l}\left(T^{-k}A_l\right) \leqslant \mu(A_l)^{-1}\mu\left(A_k \cap T^{-k}A_l\right) \leqslant \flat^{-1}e^r\mu(A_k) \leqslant \widetilde{\kappa} q^k \tag{10.8}$$

due to (10.4) and (10.5). For $l \geqslant l'$ we then find that

$$\begin{aligned} \mu_{A_l}(\varphi_{A_l} \leqslant l) &= \mu_{A_l}(K \leqslant \varphi_{A_l} \leqslant l) = \mu_{A_l}\left(\bigcup\nolimits_{k=K}^{l} T^{-k}A_l\right) \\ &\leqslant \sum\nolimits_{k=K}^{l} \mu_{A_l}\left(T^{-k}A_l\right) \leqslant \widetilde{\kappa} \sum\nolimits_{k=K}^{l} q^k < \varepsilon, \end{aligned} \tag{10.9}$$

and (3.16) follows because $\varepsilon > 0$ was arbitrary.
(b) We employ Theorem 3.8, using $\nu_l := \mu_{A_l}$ and $\mathfrak{K}$ as in (10.2). Define $A_l^\bullet := A_l \cap T^{-p}A_l = A_{l+p} \in \xi_{l+p}$ and $A_l^\circ := A_l \setminus A_l^\bullet$, so that

$$\mu\left(A_l^\bullet\right) = \mu\left(A_p \cap T^{-p}A_l\right) = T_*^p\left(\mu\mid_{A_p}\right)(A_l) = \mu(A_l) \cdot \textstyle\int_{A_l} v'_{A_p}\, d\mu_{A_l}.$$

Now $x^* \in A_l \subseteq T^p A_p$ for $l > p$, and $\mathrm{diam}(A_l) \searrow 0$. As v'_{A_p} is continuous on $T^p A_p$ with $v'_{A_p}(x^*) = 1 - \theta$, we get $\mu(A_l^\bullet) \sim (1-\theta)\mu(A_l)$ as $l \to \infty$, proving (3.20).

Observe that $\varphi_{A_l} = p$ on $A_l^\bullet$, and accordingly define $\tau_l := p$ on $A_l^\bullet$. Then (3.24) is clear. Moreover, $(T_{A_l})_*\mu_{A_l^\bullet} = \mu(A_l^\bullet)^{-1}\left(T_*^p(\mu\mid_{A_l})\right)\mid_{A_l}$, and this measure is given by the probability density $\mu(A_l^\bullet)^{-1}1_{A_l}v'_{A_l} =: h_l^\bullet$. Comparing these to the densities $\mu(A_l)^{-1}1_{A_l} =: h_l$ of the μ_{A_l} we obtain (3.25), because of $\operatorname{diam}(A_l) \searrow 0$ and (10.1).

Turning to the escaping part A_l°, note that it is ξ_{l+p}-measurable (mod μ). Define $\tau_l := l + p$ on A_l°, then (3.23) is immediate from (10.2) and convexity of $\mathfrak{K}$. We finally check (3.22). Up to a set of measure zero, $A_l^\circ = \bigcup_{W \in \xi_p \setminus \{A_p\}} A_l \cap T^{-l}W$, so that $\varphi_{A_l} > l$ on A_l°. Therefore,

$$\begin{aligned}\mu_{A_l^\circ}(\varphi_{A_l} \leqslant \tau_l) &= \mu_{A_l^\circ}(l < \varphi_{A_l} \leqslant \tau_l) \leqslant \textstyle\sum_{k=1}^p \mu_{A_l^\circ}\left(T^{-(l+k)}A_l\right) \\ &\leqslant \mu\left(A_l^\circ\right)^{-1} \textstyle\sum_{k=1}^p \mu\left(A_l \cap T^{-(l+k)}A_l\right) \\ &\leqslant \mu\left(A_l^\circ\right)^{-1} \textstyle\sum_{k=1}^p \sum_{V \in \xi_k} \mu\left(\left(A_l \cap T^{-l}V\right) \cap T^{-(l+k)}A_l\right) \\ &\leqslant \flat^{-1} e^r \mu_{A_l}\left(A_l^\circ\right)^{-1} \textstyle\sum_{k=1}^p \sum_{V \in \xi_k} \mu\left(A_l \cap T^{-l}V\right) \\ &= p\,\flat^{-1} e^r \mu_{A_l}\left(A_l^\circ\right)^{-1}\ \mu(A_l) \longrightarrow 0 \quad \text{as} \quad l \to \infty,\end{aligned}$$

where we have used $A_l = \bigcup_{V \in \xi_k} A_l \cap T^{-l}V$ (disjoint), (10.4), (10.5), and the fact that $\mu_{A_l}(A_l^\circ) \to \theta \in (0,1)$. □

Another simple situation is that of small sets A_l which consist of (fewer and fewer) rank-one cylinders. Partitioning the A_l into subsets of the same type, and recording which of those subsets an orbit hits, leads us to the study of basic discrete local processes.

Theorem 10.2 (Poisson asymptotics and local processes for unions of cylinders of GM maps). — *Let $(X, \mathcal{A}, \mu, T, \xi)$ be an ergodic probability preserving Gibbs–Markov system. Let (A_l) be a sequence of asymptotically rare events such that each A_l is ξ-measurable.*

(a) *Then (A_l) exhibits Poisson asymptotics,*

$$\mu(A_l)\Phi_{A_l} \overset{\mu}{\Longrightarrow} \Phi_{\text{Exp}} \quad \textit{and} \quad \mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \Phi_{\text{Exp}} \quad \textit{as} \quad l \to \infty. \tag{10.10}$$

(b) *Assume further that, for some integer $m \geqslant 2$, each A_l is partitioned into ξ-measurable subsets $A_l^{(1)}, \ldots A_l^{(m)}$ which satisfy $\mu_{A_l}(A_l^{(j)}) \to \vartheta_j \in (0,1)$ as $l \to \infty$. Define local observables by letting $\psi_{A_l}(x) := j$ if $x \in A_l^{(j)}$. Then,*

$$\begin{cases} \left(\mu(A_l)\Phi_{A_l}, \widetilde{\Psi}_{A_l}\right) \overset{\mu_{A_l}}{\Longrightarrow} \\ \left(\mu(A_l)\Phi_{A_l}, \Psi_{A_l}\right) \overset{\mu}{\Longrightarrow} \end{cases} (\Phi_{\text{Exp}}, \Psi^*) \quad \textit{as} \quad l \to \infty, \tag{10.11}$$

where $(\Phi_{\text{Exp}}, \Psi^*)$ is an independent pair with Ψ^* a $(\vartheta_1, \ldots, \vartheta_m)$-Bernoulli sequence.

This, too, is an easy consequence of the results above.

Proof. —
(a) We check that Theorem 3.6 applies with $\nu_l := \mu_{A_l}$, $\tau_l := 1$ and $\mathfrak{K}$ as in (10.2). Since (τ_l) is uniformly bounded, condition (3.15) is trivial. In view of (10.2) and ξ-measurability of the A_l, (3.17) is also satisfied. To validate (3.16), recall (10.4) to see that indeed

$$\begin{aligned}\mu_{A_l}(\varphi_{A_l}=1) &= \mu_{A_l}\left(T^{-1}A_l\right) = \mu(A_l)^{-1}\textstyle\sum_{Z\in\xi\cap A_l}\mu\left(Z\cap T^{-1}A_l\right)\\ &\leqslant \mu(A_l)^{-1}\textstyle\sum_{Z\in\xi\cap A_l}\flat e^r\mu(Z)\mu(A_l)\\ &\leqslant \flat^{-1}e^r\mu(A_l)\longrightarrow 0 \quad\text{as}\quad l\to\infty.\end{aligned}$$

(b) To establish the joint convergence asserted in (10.11), we will appeal to Theorem 5.3. The local observables ψ_{A_l} take their values in the compact discrete space $\mathfrak{Z} := \{1, \ldots, m\}$, and we are assuming that $\psi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \psi$ with $\Pr[\psi = j] = \vartheta_j$ for all $j \in \mathfrak{Z}$. We use the same $\mathfrak{K}$ as above.

To check condition (A) of Theorem 5.3, take any $s \in [0,\infty)$ and define $\tau_{l,s} := \lfloor s/\mu(A_l)\rfloor$. Then $A_l \cap \{\varphi_{A_l} > s/\mu(A_l)\} = A_l \cap \{\varphi_{A_l} > \tau_{l,s}\}$, so that (5.6) is trivially fulfilled. On the other hand, this set is $\xi_{\tau_{l,s}}$-measurable because A_l is ξ-measurable. Therefore, (10.2) and convexity of $\mathfrak{K}$ ensure that $T_*^{\tau_{l,s}}\mu_{A_l\cap\{\mu(A_l)\varphi_{A_l}>s\}} \in \mathfrak{K}$ for $l \geqslant 1$.

In order to validate condition (B) of Theorem 5.3, we use $\mathcal{B}_{\mathfrak{Z}}^{\pi} := \{\{j\} : j \in \mathfrak{Z}\}$ and, for arbitrary $F = \{j\}$, take $\tau_{l,F} := 1$, so that (5.7) is automatically satisfied. As $A_l \cap \{\psi_{A_l} \in F\} = A_l^{(j)}$ is ξ-measurable, (10.2) and convexity of $\mathfrak{K}$ immediately show that $T_*^{\tau_{l,F}}\mu_{A_l\cap\{\psi_{A_l}\in F\}} \in \mathfrak{K}$ for $l \geqslant 1$. Finally, (5.8) follows since

$$\begin{aligned}\mu_{A_l\cap\left\{\psi_{A_l}\in F\right\}}(\varphi_{A_l}\leqslant\tau_{l,F}) &= \mu_{A_l}\left(A_l^{(j)}\right)^{-1}\mu\left(A_l^{(j)}\cap T^{-1}A_l\right)\\ &= \mu_{A_l}\left(A_l^{(j)}\right)^{-1}\textstyle\sum_{Z\in\xi\cap A_l^{(j)}}\mu\left(Z\cap T^{-1}A_l\right)\\ &\leqslant \flat^{-1}e^r\,\mu_{A_l}\left(A_l^{(j)}\right)^{-1}\textstyle\sum_{Z\in\xi\cap A_l^{(j)}}\mu\left(Z\right)\mu\left(A_l\right)\\ &= \flat^{-1}e^r\,\mu\left(A_l\right)\longrightarrow 0\quad\text{as}\quad l\to\infty,\end{aligned}$$

where we used (10.4) again. □

We next provide some specific applications of this theorem in the context of continued fraction expansions. But sequences (A_l) as in our theorem do appear naturally in a variety of other situations. We mention one particular instance:

Remark 10.3. — Under the assumptions of Theorem 10.2 (a), it is immediate from Theorem 6.3 and Arzela–Ascoli that (10.10) implies

$$\mu(A_l)\Phi_{A_l} \overset{\nu_l}{\Longrightarrow} \Phi_{\mathrm{Exp}} \quad\text{as}\quad l\to\infty, \tag{10.12}$$

whenever (ν_l) is a sequence of probabilities with $\sup_{l\geqslant 1}\mathrm{Lip}_\xi(d\nu_l/d\mu) < \infty$, where $\mathrm{Lip}_\xi(w) := \sup_{Z\in\xi}\mathrm{Lip}_Z(w)$ with $\mathrm{Lip}_Z(w) := \sup_{x,y\in Z,x\neq y} \mid w(x)-w(y)\mid /d(x,y)$. This is a stronger (functional) version of [PT20, Proposition 3.13].

10.4. The continued fraction map. Variations on a theme of Doeblin

We will illustrate the use of Theorem 10.2 in the setup of a particularly prominent system. Set $X := [0,1]$, $\mathcal{A} := \mathcal{B}_X$, and let $T : X \to X$ be the *Gauss map* with $T0 := 0$ and

$$Tx := \frac{1}{x} - \left\lfloor\frac{1}{x}\right\rfloor = \frac{1}{x} - k \quad\text{for}\quad x \in \left(\frac{1}{k+1},\frac{1}{k}\right] =: I_k, k\geqslant 1, \tag{10.13}$$

which, since Gauss [Gau12], is known to preserve the probability density

$$h(x) := \frac{1}{\log 2}\frac{1}{1+x}, \quad x \in X. \tag{10.14}$$

The invariant *Gauss measure* μ on $\mathcal{A}$ defined by the latter, $\mu(A) := \int_A h(x)\,dx$, is exact (and hence ergodic). Iteration of T reveals the *continued fraction (CF) digits* of any $x \in X$, in that

$$x = \cfrac{1}{\mathsf{a}_1(x) + \cfrac{1}{\mathsf{a}_2(x) + \cdots}} \quad \text{with} \quad \mathsf{a}_n(x) = \mathsf{a} \circ T^{n-1}(x),\ n \geqslant 1, \tag{10.15}$$

where $\mathsf{a} : X \to \mathbb{N}$ is the *digit function* corresponding to $\xi := \{I_k : k \geqslant 1\}$, i.e. $\mathsf{a}(x) := \lfloor 1/x \rfloor = k$ for $x \in I_k$. It is a standard fact that the ergodic measure preserving piecewise invertible *CF-system* $(X, \mathcal{A}, \mu, T, \xi)$ is Gibbs–Markov.

We shall focus on the simple sequence of ξ-measurable asymptotically rare events given by $A_l := \{\mathsf{a} \geqslant l\} = \bigcup_{k \geqslant l} I_k$, which satisfy $\mu(A_l) \sim 1/(l \log 2)$ as $l \to \infty$, and allow to immediately apply Theorem 10.2 (a) to obtain Poisson asymptotics,

$$\mu(A_l)\Phi_{A_l} \overset{\mu}{\Longrightarrow} \Phi_{\mathrm{Exp}} \quad \text{and} \quad \mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \Phi_{\mathrm{Exp}} \quad \text{as} \quad l \to \infty. \tag{10.16}$$

This is a well-known classical fact with non-trivial history [Doe40, Ios77] and various extensions, see e.g. [IK02]. In the following we refine this statement, using two different sequences of local observables in order to obtain extra information on the distributions of the particular digits observed when the orbit hits A_l. In either case, Theorem 10.2 (b) applies without difficulties.

To get a better understanding of the actual size of those digits which happen to exceed some large l, we show that, asymptotically, whether these large digits are even of order l/ϑ (for some $\vartheta \in (0,1)$) is determined by an independent sequence of $(1-\vartheta, \vartheta)$-coin flips.

Proposition 10.4 (Just how large are large CF-digits?). — *Let $(X, \mathcal{A}, \mu, T, \xi)$ be the CF-system, and take any $\vartheta \in (0,1)$. Set $\psi_{A_l} := 1_{\{\mathsf{a} \geqslant l/\vartheta\}}$ on A_l, which identifies those digits $\geqslant l$ which are in fact $\geqslant l/\vartheta$. Then*

$$(\mu(A_l)\Phi_{A_l}, \Psi_{A_l}) \overset{\mu}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*) \quad \text{as} \quad l \to \infty, \tag{10.17}$$

where $(\Phi_{\mathrm{Exp}}, \Psi^*)$ is an independent pair with Ψ^* a $(1-\vartheta, \vartheta)$-Bernoulli sequence.

Proof. — Let $A'_l := \{\mathsf{a} \geqslant l/\vartheta\} = \bigcup_{k \geqslant l/\vartheta} I_k \subseteq A_l$, which is again ξ-measurable. Trivially, $\mu(A'_l) \sim \vartheta\mu(A'_l)$ as $l \to \infty$, and the assertion follows via Theorem 10.2 (b). □

Remark 10.5. — This result is closely related to the fact that ergodic sums of the digit function satisfy a functional stable limit theorem,

$$\left(\frac{\log 2}{n}\sum_{k=0}^{\lfloor nt \rfloor - 1} \mathsf{a} \circ T^k - t(\log n - \gamma)\right)_{t \geqslant 0} \overset{\mu}{\Longrightarrow} \mathrm{G} \quad \text{as} \quad n \to \infty, \tag{10.18}$$

where γ is Euler's constant, $\mathrm{G} = (\mathrm{G}_t)_{t \geqslant 0}$ is an α-stable motion with $\alpha = 1$ and skewness $\beta = 1$, and convergence takes place on the Skorohod space $\mathbb{D}[0,\infty)$ equipped with the J_1-topology. Proposition 10.4 can also be derived using information on the convergence (10.18), see the approach to that limit theorem developed in [TK10].

Turning to a different feature of individual CF-digits which cannot be extracted from (10.18), we now observe that the residue classes mod m of large CF-digits are asymptotically equidistributed and independent of each other (and of the waiting times).

Proposition 10.6 (Residue classes of large CF-digits). — *Let* $(X, \mathcal{A}, \mu, T, \xi)$ *be the CF-system, and take an integer* $m \geqslant 2$. *Define* $\psi : X \to \{0, \dots, m-1\}$ *by* $\psi(x) := j$ *if* $\mathsf{a}(x) \equiv j \pmod m$, *so that* $\psi \circ T^{n-1}$ *identifies the residue class mod* m *of the digit* a_n, *and set* $\psi_{A_l} := \psi \mid_{A_l}$. *Then*

$$(\mu(A_l)\Phi_{A_l}, \Psi_{A_l}) \overset{\mu}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*) \quad \text{as} \quad l \to \infty, \tag{10.19}$$

where $(\Phi_{\mathrm{Exp}}, \Psi^*)$ *is an independent pair with* Ψ^* *a* $(\frac{1}{m}, \dots, \frac{1}{m})$*-Bernoulli sequence.*

Proof. — Setting $I_0 := \varnothing$ we have $\{\psi = j\} = \bigcup_{i \geqslant 0} I_{im+j}$ for all j. From our explicit knowledge of μ and the I_k it is easily seen that, for each $j \in \{0, \dots, m-1\}$, the ξ-measurable sets $A_l^{(j)} := A_l \cap \{\psi = j\}$ satisfy $\mu_{A_l}(A_l^{(j)}) \to 1/m$ as $l \to \infty$. Now apply Theorem 10.2(b). □

10.5. Local processes of interval maps

We now turn to a basic situation in which the geometry of the underlying space suggests a natural way of describing the relative position inside small sets by specific local observables. Call $(X, \mathcal{A}, \mu, T, \xi)$ a (probability-preserving) *Gibbs–Markov interval map* provided that it is a GM-system as in the previous section, where X and each $Z \in \xi$ is an open interval, and the invariant probability μ is absolutely continuous w.r.t. one-dimensional Lebesgue measure λ. In this setup, we shall study asymptotically rare sequences (A_l) of subintervals, and take the normalizing interval charts $\psi_{A_l} : A_l \to [0,1]$ as our local observables.

Only using the elementary properties employed in the previous section, we are going to prove

Theorem 10.7 (Small intervals in GM interval maps). — *Let* $(X, \mathcal{A}, \mu, T, \xi)$ *be a probability-preserving ergodic Gibbs–Markov interval map,* (A_l) *an asymptotically rare sequence of subintervals, and* $\psi_{A_l} : A_l \to [0,1]$ *the corresponding normalizing interval charts, giving local processes* Ψ_{A_l}. *Assume that* $x^* \in X$ *is not periodic and such that each* $\xi_l(x^*)$ *is well defined, and the* A_l *are contained in neighbourhoods* I_l *of* x^* *with* $\mathrm{diam}(I_l) \to 0$. *Then,*

$$\left(\mu(A_l)\Phi_{A_l}, \widetilde{\Psi}_{A_l}\right) \overset{\mu_{A_l}}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*) \quad \text{as} \quad l \to \infty, \tag{10.20}$$

and

$$(\mu(A_l)\Phi_{A_l}, \Psi_{A_l}) \overset{\mu}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*) \quad \text{as} \quad l \to \infty, \tag{10.21}$$

where $(\Phi_{\mathrm{Exp}}, \Psi^*)$ *is an independent pair with* Ψ^* *an iid sequence of uniformly distributed elements of* $[0,1]$.

To establish this result we can essentially argue as in the preceding section, once we replace the intervals A_l by more convenient sets A'_l which are unions of cylinders of rank $\jmath(A_l)$, where

$$\jmath(A) := \frac{-2\log(\mu(A))}{-\log q}, \quad A \in \mathcal{A} \quad \text{with} \quad \mu(A) > 0, \tag{10.22}$$

where $q \in (0,1)$ is as in (10.5). Note that

$$\mu(A_l)\jmath(A_l) \longrightarrow 0 \quad \text{as} \quad l \to \infty, \tag{10.23}$$

whenever (A_l) is an asymptotically rare sequence.

Lemma 10.8 (Approximating intervals by cylinders). — *Let $(X, \mathcal{A}, \mu, T, \xi)$ be a probability-preserving ergodic Gibbs–Markov interval map and (A_l) an asymptotically rare sequence of subintervals. Define*

$$A'_l := \bigcup\nolimits_{W \in \xi_{\jmath(A_l)} : W \subseteq A_l} W, \quad l \geqslant 1, \tag{10.24}$$

then the $\xi_{\jmath(A_l)}$-measurable sets $A'_l \subseteq A_l$ satisfy $\mu(A_l \triangle A'_l) = o(\mu(A_l))$ as $l \to \infty$.

Proof. — If A is an interval, $\jmath \geqslant 1$, and $A' := \bigcup_{W \in \xi_\jmath : W \subseteq A_l} W$, then $A_l \setminus A'_l$ consists of at most two subintervals of measures not exceeding $\kappa q^\jmath$ (recall (10.5)). □

We are now ready for the

Proof of Theorem 10.7.

(i) In view of the Lemma and Remark 5.5, we can assume w.l.o.g. that (up to a set of measure zero) each A_l is $\xi_{\jmath(A_l)}$-measurable. (Note that $\jmath(A'_l) \geqslant \jmath(A_l)$.)

We prove our result by a direct application of Theorem 5.3. Note first that

$$\psi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \psi \quad \text{as} \quad l \to \infty, \tag{10.25}$$

where ψ is uniformly distributed in $\mathfrak{Z} := [0,1]$ (see the discussion following (4.6)).

(ii) We will validate condition (B) of Theorem 5.3 for every $F \in \mathcal{B}_{\mathfrak{Z}}^{\pi} := \{[a,b] : 0 \leqslant a \leqslant b \leqslant 1\}$ of positive measure. In the particular case of $F = \mathfrak{Z}$ this verifies the assumptions of Theorem 3.6, and hence implies Poisson asymptotics,

$$\mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Longrightarrow} \Phi_{\mathrm{Exp}} \quad \text{and} \quad \mu(A_l)\Phi_{A_l} \overset{\mu}{\Longrightarrow} \Phi_{\mathrm{Exp}} \quad \text{as} \quad l \to \infty. \tag{10.26}$$

Throughout, we use $\mathfrak{K}$ as in (10.2). Now fix any $F \in \mathcal{B}_{\mathfrak{Z}}^{\pi}$ with $\Pr[\psi \in F] = \lambda(F) > 0$.

The sets $B_{l,F} := A_l \cap \{\psi_{A_l} \in F\}$ are intervals. Define

$$B'_{l,F} := \bigcup\nolimits_{V \in \xi_{\jmath(B_{l,F})} : V \subseteq A_l} V, l \geqslant 1,$$

which satisfy $\mu(B_{l,F} \triangle B'_{l,F}) = o(\mu(B_{l,F}))$ as $l \to \infty$ (Lemma 10.8). Setting $\nu_{l,F} := \mu_{B'_{l,F}}$ we therefore have $d_{\mathfrak{P}}(\nu_{l,F}, \mu_{A_l \cap \{\psi_{A_l} \in F\}}) \to 0$.

Next, $\lambda(B_{l,F}) \sim \lambda(F)\lambda(A_l)$ as $l \to \infty$ by definition of $B_{l,F}$. Since the invariant density $d\mu/d\lambda =: h$ of T is continuous at x^* (x^* being in the interior of $\xi(x^*)$) with $h(x^*) > 0$, we see that also $\mu(B_{l,F}) \sim \lambda(F)\mu(A_l)$.

Let $\tau_{l,F} := \jmath(B_{l,F})$, $l \geqslant 1$, then it is immediate from (10.23) that $\mu(A_l)\tau_{l,F} \sim \lambda(F)^{-1}\mu(B_{l,F})\jmath(B_{l,F}) \to 0$, and hence (5.7). Also, since $B'_{l,F}$ is $\xi_{\jmath(B_{l,F})}$-measurable, it is clear from (10.2) and convexity of $\mathfrak{K}$ that $T_*^{\tau_{l,F}}\nu_{l,F} \in \mathfrak{K}$ for $l \geqslant 1$. It only remains to check (5.8) or, equivalently, that

$$\nu_{l,F}\left(\varphi_{A_l} \leqslant \tau_{l,F}\right) \longrightarrow 0 \quad \text{as} \quad l \to \infty. \tag{10.27}$$

This is done by an argument slightly extending that of Theorem 10.1. Take any $\varepsilon > 0$. Choose $K \geqslant 1$ so large that $\widetilde{\kappa} q^K/(1-q) < \varepsilon/2$ with $\widetilde{\kappa} := 2\lambda(F)^{-1}\flat^{-1}e^r\kappa$. For every $k \in \{1,\dots,l\}$ we have, using (10.4) and (10.5),

$$\begin{aligned}\mu_{B'_{l,F}}\left(T^{-k}A_l\right) &\leqslant \mu\left(B'_{l,F}\right)^{-1}\textstyle\sum_{Z\in\xi_k:\,Z\cap A_l\neq\varnothing}\mu\left(Z\cap T^{-k}A_l\right)\\ &\leqslant \flat^{-1}e^r\mu\left(B'_{l,F}\right)^{-1}\mu(A_l)\textstyle\sum_{Z\in\xi_k:\,Z\cap A_l\neq\varnothing}\mu(Z)\\ &\leqslant \flat^{-1}e^r\mu\left(B'_{l,F}\right)^{-1}\mu(A_l)\left(\mu(A_l)+2\kappa q^k\right)\\ &\leqslant \widetilde{\kappa}\left(\mu(A_l)+q^k\right) \quad \text{for} \quad l\geqslant l',\end{aligned} \tag{10.28}$$

where we note that at most two of the $Z \in \xi_k$ which intersect the interval A_l are not covered by A_l. Since x^* is not periodic, and a continuity point of each T^n, there is some l'' such that $\varphi_{A_l} > K$ on A_l whenever $l \geqslant l''$. As seen before, we also have $\mu(A_l)\tau_{l,F} < \varepsilon/(2\widetilde{\kappa})$ whenever $l \geqslant l'''$. We thus find that

$$\begin{aligned}\nu_{l,F}\left(\varphi_{A_l}\leqslant\tau_{l,F}\right) &= \mu_{B'_{l,F}}\left(K\leqslant\varphi_{A_l}\leqslant\tau_{l,F}\right) = \mu_{B'_{l,F}}\left(\textstyle\bigcup_{k=K}^{\tau_{l,F}}T^{-k}A_l\right)\\ &\leqslant \textstyle\sum_{k=K}^{\tau_{l,F}}\mu_{B'_{l,F}}\left(T^{-k}A_l\right) \leqslant \widetilde{\kappa}\left(\mu(A_l)\tau_{l,F}+\sum_{k=K}^{\tau_{l,F}}q^k\right)\\ &< \varepsilon \quad \text{for} \quad l\geqslant l'\vee l''\vee l''',\end{aligned} \tag{10.29}$$

and (10.27) follows as $\varepsilon > 0$ was arbitrary. Condition (B) of Theorem 5.3 is fulfilled.

(iii) Turning to condition (A) of Theorem 5.3, fix any $s \in [0,\infty)$. For $\theta > 0$ consider the sets $C_l(\theta) := A_l \cap \{\varphi_{A_l} > \theta\}$, $l \geqslant 1$. Since the A_l are $\xi_{\jmath(A_l)}$-measurable, each $C_l(\theta)$ is $\xi_{\jmath(A_l)+\theta}$-measurable.

We approximate the $B_l := A_l \cap \{\mu(A_l)\varphi_{A_l} > s\} = C_l(\theta_l)$ with $\theta_l := s/\mu(A_l)$ by the sets $B'_l := C_l(\theta'_l)$ with $\theta'_l := \theta_l - \jmath(A_l)$. It is clear from the definition of $\jmath(A)$ that $\theta'_l \sim \theta_l$ as $l \to \infty$. In view of step (ii) above, we can already use (10.26). The latter shows that $\mu(B_l \triangle B'_l) = o(\mu(B_l))$, and hence $d_{\mathfrak{P}}(\mu_{B'_l},\mu_{B_l}) \to 0$.

But for the B'_l condition (A) is very easy if we take $\tau_{l,s} := \theta_l$. Indeed, by (10.26), $\mu_{B'_l}(\varphi_{A_l} \leqslant \tau_{l,s}) = \mu_{B'_l}(s - \mu(A_l)\jmath(A_l) < \mu(A_l)\varphi_{A_l} \leqslant s) \to 0$ as $l \to \infty$. On the other hand, $T_*^{\tau_{l,s}}\mu_{B'_l} \in \mathfrak{K}$ for $l \geqslant 1$, because each B'_l is θ_l-measurable. □

11. Inducing and further examples

11.1. Induced versions of the processes

When studying specific systems, one often tries to find some good reference set $Y \in \mathcal{A}$ such that the first-return map $T_Y : Y \to Y$ is more convenient than T. In this case, it often pays to prove a relevant property first for T_Y, and to transfer it back to T afterwards.

In the following, we let $\varphi_A^Y : Y \to \overline{\mathbb{N}}$ denote the hitting time of $A \in \mathcal{A} \cap Y$ under the first-return map T_Y, that is,

$$\varphi_A^Y(x) := \inf\left\{j \geqslant 1 : T_Y^j x \in A\right\}, \quad x \in Y, \tag{11.1}$$

and write $\Phi_A^Y := (\varphi_A^Y, \varphi_A^Y \circ T_A, \varphi_A^Y \circ T_A^2, \ldots)$ on Y for the hitting-time process of A under T_Y. Since μ_Y is the natural invariant probability measure for T_Y, the canonical normalization for φ_A^Y and Φ_A^Y is $\mu_Y(A)$.

Given an $\mathfrak{Z}$-valued local observable on A with corresponding local process $\Psi_A = (\psi_A \circ T_A, \psi_A \circ T_A^2, \ldots)$, we can also consider the corresponding object for the first-return map, $\Psi_A^Y = (\psi_A \circ (T_Y)_A, \psi_A \circ (T_Y)_A^2, \ldots)$ on Y. But since the first-return maps on A respectively induced by T and T_Y coincide, $T_A = (T_Y)_A$, we have $\Psi_A^Y = \Psi_A\mid_Y$ and there is no need for this extra notation.

11.2. Relating original and induced processes

Inducing was first used to deal with limit laws for normalized return- or hitting times $\mu(A)\varphi_A$ in [BSTV03]. A more general abstract form of their result was given in [HWZ14], and [Zwe19] contains an even more flexible version. The theorem below confirms that the same strategy can also be employed when dealing with joint processes $(\mu(A)\Phi_A, \Psi_A)$ for small sets. The argument closely follows that of [HWZ14], and its process variant from [FFTV16], but compares the two hitting-time processes in probability rather than just in distribution, thus keeping track of their relation to the second process Ψ_A.

Theorem 11.1 (Joint limit processes under μ via inducing). — *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system, $Y \in \mathcal{A}$, (A_l) an asymptotically rare sequence in $\mathcal{A} \cap Y$, and $(\psi_{A_l})_{l \geqslant 1}$ a sequence of $\mathfrak{Z}$-valued local observables for the A_l with corresponding local processes Ψ_{A_l}.*

Assume that (Φ, Ψ) is a random element of $[0,\infty)^{\mathbb{N}} \times \mathfrak{Z}^{\mathbb{N}}$. Then, as $l \to \infty$,

$$(\mu(A_l)\Phi_{A_l}, \Psi_{A_l}) \overset{\mu}{\Longrightarrow} (\Phi, \Psi) \quad \textit{iff} \quad (\mu_Y(A_l)\Phi_{A_l}^Y, \Psi_{A_l}) \overset{\mu_Y}{\Longrightarrow} (\Phi, \Psi). \tag{11.2}$$

Proof. —
(i) According to Proposition 5.1, we can replace μ by μ_Y in the first convergence statement of (11.2). Therefore it suffices to show that for every $d \geqslant 1$,

$$\left(\mu(A_l)\Phi_{A_l}^{[d]}, \Psi_{A_l}^{[d]}\right) \overset{\mu_Y}{\Longrightarrow} \left(\Phi^{[d]}, \Psi^{[d]}\right) \quad \text{iff} \quad \left(\mu_Y(A_l)\Phi_{A_l}^{Y,[d]}, \Psi_{A_l}^{[d]}\right) \overset{\mu_Y}{\Longrightarrow} \left(\Phi^{[d]}, \Psi^{[d]}\right).$$

Since $\Phi^{[d]} = (\varphi^{(0)}, \ldots, \varphi^{(d-1)})$ is finite-valued by assumption, we do not lose information if instead we work with $\Phi^{\Sigma[d]} := (\varphi^{(0)}, \varphi^{(0)} + \varphi^{(1)}, \ldots, \varphi^{(0)} + \cdots + \varphi^{(d-1)})$. Define $\Phi_{A_l}^{\Sigma[d]}$ and $\Phi_{A_l}^{Y,\Sigma[d]}$ analogously as vectors of partial sums of $\Phi_{A_l}^{[d]}$ and $\Phi_{A_l}^{Y,[d]}$, respectively. Then,

$$\left(\mu(A_l)\Phi_{A_l}^{[d]}, \Psi_{A_l}^{[d]}\right) \overset{\mu_Y}{\Longrightarrow} \left(\Phi^{[d]}, \Psi^{[d]}\right) \quad \text{iff} \quad \left(\mu(A_l)\Phi_{A_l}^{\Sigma[d]}, \Psi_{A_l}^{[d]}\right) \overset{\mu_Y}{\Longrightarrow} \left(\Phi^{\Sigma[d]}, \Psi^{[d]}\right),$$

while

$$\left(\mu_Y(A_l)\Phi_{A_l}^{Y,[d]}, \Psi_{A_l}^{[d]}\right) \overset{\mu_Y}{\Longrightarrow} \left(\Phi^{[d]}, \Psi^{[d]}\right) \quad \text{iff} \quad \left(\mu_Y(A_l)\Phi_{A_l}^{Y,\Sigma[d]}, \Psi_{A_l}^{[d]}\right) \overset{\mu_Y}{\Longrightarrow} \left(\Phi^{\Sigma[d]}, \Psi^{[d]}\right).$$

Therefore our assertion (11.2) follows once we check that for every $d \geqslant 1$,

$$\left(\mu(A_l)\Phi_{A_l}^{\Sigma[d]}, \Psi_{A_l}^{[d]}\right) \overset{\mu_Y}{\Longrightarrow} \left(\Phi^{[d]}, \Psi^{[d]}\right) \quad \text{iff} \quad \left(\mu_Y(A_l)\Phi_{A_l}^{Y,\Sigma[d]}, \Psi_{A_l}^{[d]}\right) \overset{\mu_Y}{\Longrightarrow} \left(\Phi^{[d]}, \Psi^{[d]}\right).$$

The latter is immediate if we check that for every $d \geqslant 1$,

$$d_{[0,\infty]^d}\left(\mu(A_l)\Phi_{A_l}^{\Sigma[d]}, \mu_Y(A_l)\Phi_{A_l}^{Y,\Sigma[d]}\right) \overset{\mu_Y}{\longrightarrow} 0 \quad \text{as} \quad l \to \infty, \tag{11.3}$$

because this convergence in probability trivially carries over to the joint processes,

$$d_{[0,\infty]^d \times \mathfrak{Z}^d}\left(\left(\mu(A_l)\Phi_{A_l}^{\Sigma[d]}, \Psi_{A_l}^{[d]}\right), \left(\mu_Y(A_l)\Phi_{A_l}^{Y,\Sigma[d]}, \Psi_{A_l}^{[d]}\right)\right) \overset{\mu_Y}{\longrightarrow} 0 \quad \text{as} \quad l \to \infty.$$

(ii) To validate (11.3), we take some $d \geqslant 1$ and any $\varepsilon > 0$. Note that

$$\varkappa(\delta) := \sup_{t \geqslant 0} d_{[0,\infty]}\left(t, e^{\delta}t\right) \to 0 \quad \text{as} \quad \delta \searrow 0.$$

Now fix some $\delta > 0$ so small that $d\varkappa(\delta) < \varepsilon$. By the Ergodic theorem and Kac' formula, we have

$$m^{-1}\textstyle\sum_{j=0}^{m-1}\varphi_Y \circ T_Y^j \longrightarrow \mu(Y)^{-1} \quad \text{a.e. on} \quad Y, \tag{11.4}$$

so that the increasing sequence $(E_M)_{M \geqslant 1}$ of sets given by

$$E_M := \left\{\textstyle\sum_{j=0}^{m-1}\varphi_Y \circ T_Y^j = e^{\pm\delta}\mu(Y)^{-1}m \text{ for } m \geqslant M\right\} \in \mathcal{A} \cap Y$$

satisfies $\mu_Y(E_M^c) \to 0$ as $M \to \infty$. Now fix some M such that $\mu_Y(E_M^c) < \varepsilon/2$. Next, let $F_l := \{\varphi_{A_l}^Y \geqslant M\} \in \mathcal{A} \cap Y$, $l \geqslant 1$. Then, $F_l = Y \cap \bigcap_{j=1}^{M-1} T_Y^{-j} A_l^c$, and hence $\mu_Y(F_l^c) \leqslant \sum_{j=1}^{M-1} \mu_Y(T_Y^{-j}A_l) \leqslant M\,\mu_Y(A_l) \to 0$ as $l \to \infty$. Pick $L \geqslant 1$ so large that $\mu_Y(F_l^c) < \varepsilon/2$ for $l \geqslant L$.

(iii) Observe now that for $A \in \mathcal{A} \cap Y$ we have $\varphi_A = \sum_{j=0}^{\varphi_A^Y - 1} \varphi_Y \circ T_Y^j$ on Y, and, more generally, for any $i \geqslant 1$,

$$\varphi_A + \cdots + \varphi_A \circ T_A^{i-1} = \sum_{j=0}^{\varphi_A^Y + \cdots + \varphi_A^Y \circ T_A^{i-1} - 1} \varphi_Y \circ T_Y^j \quad \text{on} \quad Y. \tag{11.5}$$

By definition of E_M and F_l we have, for every $i \in \{1, \ldots, d\}$,

$$\varphi_{A_l} + \cdots + \varphi_{A_l} \circ T_{A_l}^{i-1} = e^{\pm\delta}\mu(Y)^{-1}\left(\varphi_{A_l}^Y + \cdots + \varphi_{A_l}^Y \circ T_{A_l}^{i-1}\right) \quad \text{on} \quad E_M \cap F_l.$$

Note that the left-hand expression is the ith component of $\Phi_{A_l}^{\Sigma[d]}$, while the sum on the right-hand side is the ith component of $\Phi_{A_l}^{Y,\Sigma[d]}$. Due to our choice of δ we thus find that

$$d_{[0,\infty]^d}\left(\mu(A_l)\Phi_{A_l}^{\Sigma[d]}, \mu_Y(A_l)\Phi_{A_l}^{Y,\Sigma[d]}\right) \leqslant d\varkappa(\delta) < \varepsilon \text{ on } E_M \cap F_l.$$

But since $\mu_Y((E_M \cap F_l)^c) < \varepsilon$ for $l \geqslant L$, this proves (11.3). □

We also provide an inducing principle for limits under the measures μ_{A_l}. Recall from (4.8) that transferring information about the asymptotics of spatiotemporal processes under one of $(\mu)_{l \geqslant 1}$ and $(\mu_{A_l})_{l \geqslant 1}$ to the other requires extra information. Therefore this principle is less general than Theorem 11.1, and we content ourselves with the case relevant for typical applications.

Proposition 11.2 (Joint limit processes under μ_{A_l} via inducing). — *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system, $Y \in \mathcal{A}$, (A_l) an asymptotically rare sequence in $\mathcal{A} \cap Y$, and $(\psi_{A_l})_{l \geqslant 1}$ a sequence of $\mathfrak{Z}$-valued local observables for the A_l with corresponding local processes Ψ_{A_l}. Assume there are measurable $\tau_l^Y : A_l \to \mathbb{N}_0$ and a compact set $\mathfrak{K} \subseteq \mathfrak{P}$ such that*

$$\mu_{A_l}\left(\tau_l^Y < \varphi_{A_l}^Y\right) \longrightarrow 1 \quad \text{as} \quad l \to \infty \quad \text{while} \quad (T_Y)_*^{\tau_l^Y} \mu_{A_l} \in \mathfrak{K} \ \text{ for } \ l \geqslant 1, \tag{11.6}$$

and that

$$\left(\mu_Y(A_l)\Phi_{A_l}^Y, \widetilde{\Psi}_{A_l}\right) \overset{\mu_{A_l}}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*) \quad \text{as} \quad l \to \infty, \tag{11.7}$$

with $(\Phi_{\mathrm{Exp}}, \Psi^*)$ an independent pair of iid sequences. Then

$$\left(\mu(A_l)\Phi_{A_l}, \widetilde{\Psi}_{A_l}\right) \overset{\mu_{A_l}}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*) \quad \text{as} \quad l \to \infty. \tag{11.8}$$

Proof. — Applying Theorem 5.2 to the first-return map T_Y we see that (11.6) and (11.7) together imply $(\mu_Y(A_l)\Phi_{A_l}^Y, \Psi_{A_l}) \overset{\mu_Y}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*)$. In view of Theorem 11.1 this entails

$$(\mu(A_l)\Phi_{A_l}, \Psi_{A_l}) \overset{\mu}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*) \quad \text{as} \quad l \to \infty. \tag{11.9}$$

Now define $\tau_l := \sum_{j=0}^{\tau_l^Y - 1} \varphi_Y \circ T_Y^j$ for $l \geqslant 1$. This ensures that $T^{\tau_l} = (T_Y)^{\tau_l^Y}$ on A_l and that $A_l \cap \{\tau_l < \varphi_{A_l}\} = A_l \cap \{\tau_l^Y < \varphi_{A_l}^Y\}$. Having thus found measurable $\tau_l : A_l \to \mathbb{N}_0$ satisfying

$$\mu_{A_l}(\tau_l < \varphi_{A_l}) \longrightarrow 1 \quad \text{as} \quad l \to \infty \quad \text{while} \quad T^{\tau_l}\mu_{A_l} \in \mathfrak{K} \text{ for } l \geqslant 1,$$

we can apply Theorem 5.2 again, this time to T, to derive (11.8) from (11.9). □

11.3. Application: Spatiotemporal Poisson limits for systems inducing GM maps

Using the results of this section, we can easily provide further examples of systems exhibiting spatiotemporal Poisson limits, including some with arbitrarily slow mixing rates. We first record

Theorem 11.3 (Small intervals via induced GM maps). — *Let $(X, \mathcal{A}, \mu, T)$ be a probability-preserving system, and $Y \in \mathcal{A}$ a set with $\mu(Y) > 0$ on which it induces a Gibbs-Markov interval map $(Y, \mathcal{A} \cap Y, \mu_Y, T_Y, \xi_Y)$. Let (A_l) an asymptotically rare sequence of subintervals of Y, and $\psi_{A_l} : A_l \to [0,1]$ the corresponding normalizing interval charts, giving local processes Ψ_{A_l}. Assume that $x^* \in Y$ is not periodic and such that each $\xi_{Y,l}(x^*)$ is well defined, and the A_l are contained in neighbourhoods I_l of x^* with $\operatorname{diam}(I_l) \to 0$. Then,*

$$\left(\mu(A_l)\Phi_{A_l}, \widetilde{\Psi}_{A_l}\right) \overset{\mu_{A_l}}{\Longrightarrow} (\,\Phi_{\mathrm{Exp}}, \Psi^*) \quad \text{as} \quad l \to \infty, \tag{11.10}$$

and

$$(\mu(A_l)\Phi_{A_l}, \Psi_{A_l}) \overset{\mu}{\Longrightarrow} (\,\Phi_{\mathrm{Exp}}, \Psi^*) \quad \text{as} \quad l \to \infty, \tag{11.11}$$

where $(\,\Phi_{\mathrm{Exp}}, \Psi^*)$ is an independent pair with Ψ^* an iid sequence of uniformly distributed elements of $[0,1]$.

Proof. — It is immediate from our assumptions that Theorem 10.7 applies to the induced system, thus proving

$$\left(\mu_Y(A_l)\Phi^Y_{A_l}, \Psi_{A_l}\right) \overset{\mu_Y}{\Longrightarrow} (\,\Phi_{\mathrm{Exp}}, \Psi^*) \quad \text{as} \quad l \to \infty, \tag{11.12}$$

and

$$\left(\mu_Y(A_l)\Phi^Y_{A_l}, \widetilde{\Psi}_{A_l}\right) \overset{\mu_{A_l}}{\Longrightarrow} (\Phi_{\mathrm{Exp}}, \Psi^*) \quad \text{as} \quad l \to \infty. \tag{11.13}$$

The first of these immediately implies (11.11) via Theorem 11.1. We then check that (11.10) can be derived from (11.13) using Proposition 11.2. In view of Remark 5.5 and Lemma 10.8 we can replace (A_l) by (A'_l) in both (11.10) and (11.13), where $A'_l := \bigcup_{W \in \xi_{Y,\jmath(A_l)} : W \subseteq A_l} W$, $l \geqslant 1$, (note the use of the partition ξ_Y of the induced system) with $\jmath(A_l)$ as in (10.22). To apply Proposition 11.2, recall that condition (11.6) for (A'_l) and the induced Gibbs–Markov interval map T_Y has already been validated in the proof of Theorem 10.7. □

We illustrate the use of this theorem by applying it to prototypical nonuniformly expanding interval maps with indifferent fixed points. (Everything said below generalized in a trivial way to Markovian interval maps with several indifferent fixed points satisfying the obvious analogous analytical conditions.)

Example 11.4 (Probability preserving intermittent interval maps). — Let (X, T, ξ) be piecewise increasing with $X = [0,1]$ and $\xi = \{(0,c),(c,1)\}$, mapping each $Z \in \xi$ onto $(0,1)$. Assume that $T\mid_{(c,1)}$ admits a uniformly expanding $\mathcal{C}^2$ extension to $[c,1]$, while $T\mid_{(0,c)}$ extends to a $\mathcal{C}^2$ map on $(0,c]$ and is expanding except for an indifferent fixed point at $x^* = 0$: for every $\varepsilon > 0$ there is some $\rho(\varepsilon) > 1$ such that $T' \geqslant \rho(\varepsilon)$ on $[\varepsilon, c]$, while $T0 = 0$ and $\lim_{x \searrow 0} T'x = 1$ with T' increasing on some $(0,\delta)$. Suppose also that

$$\begin{gathered}\text{there is a continuous decreasing function } g \text{ on } (0,c] \text{ with} \\ \int_0^c g(x)\,dx < \infty \quad \text{and} \quad |T''| \leqslant g \;\text{ on }\; (0,c].\end{gathered} \tag{11.14}$$

Let $c =: c_0 > c_1 > c_2 > \ldots > 0$ be the points with $Tc_m = c_{m-1}$ for $m \geqslant 1$. By an analytic argument which goes back to [Tha80], see for example [Zwe03, § 3] or [Tha05, § 4], assumption (11.14) ensures that the induced system on any $Y := Y^{(m)} := (c_m, 1)$ is Gibbs–Markov, and that T possesses a unique absolutely continuous (w.r.t. Lebesgue measure λ) invariant probability measure μ with density h strictly positive and continuous on $(0,1]$. This family of maps contains systems with very slow decay of correlations, see for example [Hol05].

Now take any $x^* \in X$ which is not periodic and such that each $\xi_j(x^*)$ is well defined, let $(A_l)_{l\geqslant 1}$ be a sequence of non-degenerate intervals contained in neighbourhoods I_l of x^* with $\operatorname{diam}(I_l) \to 0$, and let $\psi_{A_l} : A_l \to [0,1]$ denote the corresponding normalizing interval charts, giving local processes Ψ_{A_l}. Then Theorem 11.3 applies (with $Y = Y^{(m)}$ for m so large that $x^* \in Y$), so that (11.10) and (11.11) hold in the present situation. (We only need to observe that all cylinders $\xi_{Y,l}(x^*)$ of the induced system around this particular point are well defined since the cylinders of the original system are.)

12. Relation to the tail-σ-algebra

The abstract limit theorems of the present paper merely require the probability preserving system $(X, \mathcal{A}, \mu, T)$ to be ergodic. On top of this we only impose conditions on the specific asymptotically rare sequence (A_l) under consideration. These assumptions do not imply that the system is mixing, as is clear from the basic GM-map examples of Section 10, which can be taken to have a periodic structure. However, it is well-known that this is the only way in which an ergodic probability preserving GM-map can fail to be mixing, since it always has a discrete tail-σ-algebra.

We conclude by showing that if an ergodic probability preserving map T admits a (one-sided) generating partition (mod μ) such that the cylinders $\xi_l(x)$ around a.e. point $x \in X$ (the element of $\xi_l := \bigvee_{j=0}^{l-1} T^{-j}\xi$ containing x) satisfy a condition similar to that used above, with constant delay times and a common compact set of image measures, then it is exact up to a cyclic permutation.

Proposition 12.1 (Abundance of good cylinders implies discrete tail). — *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving system. Assume that there is a compact subset $\mathfrak{K}$ of $(\mathfrak{P}, d_{\mathfrak{P}})$ and a countable partition ξ of X (mod μ), with $\mathcal{A} = \sigma(\xi_n : n \geqslant 1)$ (mod μ) and the following property: For a.e. $x \in X$ the sequence $(\xi_l(x))_{l \geqslant 1}$ is well defined with $0 < \mu(\xi_l(x)) \searrow 0$, and it admits a sequence of constants $\tau_{x,l} \in \mathbb{N}_0$, such that*

$$T_*^{\tau_{x,l}} \mu_{\xi_l(x)} \in \mathfrak{K} \quad \text{for} \quad l \geqslant 1. \tag{12.1}$$

Then there are $p \in \mathbb{N}$ and $X_1, \ldots, X_p \in \mathcal{A}$ such that $T^{-1}X_{i+1} = X_i$. The tail-σ-algebra has the form $\bigcap_{n \geqslant 0} T^{-n}\mathcal{A} = \sigma(X_1, \ldots, X_p)$ (mod μ), and $T^p\mid_{X_i}$ is exact.

Remark 12.2. —

(a) The proof only requires $\mathfrak{K}$ to be *weakly* compact in $(\mathfrak{P}, d_{\mathfrak{P}})$ (with the latter identified with $(\mathcal{D}(\mu), \|\centerdot\|_{L_1(\mu)})$).
(b) Condition (12.1), and its generalization pointed out in a), can be interpreted as weak bounded distortion conditions. We exploit them through a Rokhlin type argument, see [Roh64].
(c) If we drop the assumption of ergodicity, the argument below still shows that the tail-σ-algebra is discrete.

Proof. — Abbreviate $\mathcal{A}_\infty := \bigcap_{n \geqslant 0} T^{-n}\mathcal{A}$. Our assertion follows by easy routine arguments once we show that there is some constant $\eta > 0$ such that

$$\mu(A) > \eta \quad \text{for all} \quad A \in \mathcal{A}_\infty \ \text{ with } \ \mu(A) > 0. \tag{12.2}$$

(i) Let $\mathfrak{U} \subseteq L_1(\mu)$ be the set of (probability) densities of all measures $\nu \in \mathfrak{K}$. Then $\mathfrak{U}$ is strongly compact in $L_1(\mu)$, and hence also uniformly integrable. Consequently, there is some $\eta > 0$ such that

$$\mu(h > 2\eta) > 2\eta \quad \text{ for } u \in \mathcal{U}. \tag{12.3}$$

(Otherwise there are $u_m \in \mathfrak{U}$ such that the sets $B_m := \{u_m > 2/m\}$ satisfy $\mu(B_m) \leqslant 2/m$ for $m \geqslant 1$. But $\mu(B_m) \to 0$ implies $\int_{B_m} u_m \, d\mu \to 0$ by uniform integrability, contradicting $\int u_m \, d\mu = 1$.) For probability densities u, v with $u \in \mathfrak{U}$ we then have

$$\mu(v > \eta) \geqslant \mu(u > 2\eta) - \mu\left(\{u > 2\eta\} \setminus \{v > \eta\}\right) > 2\eta - \eta^{-1}\|u - v\|_1 ,$$

so that

$$\mu(v > \eta) > \eta \quad \text{for probability densities } v \text{ with } \mathrm{dist}_{L_1(\mu)}(v, \mathfrak{U}) < \eta^2. \tag{12.4}$$

(ii) Now fix any $A \in \mathcal{A}_\infty$ (meaning that there are $A_n \in \mathcal{A}$ such that $A = T^{-n}A_n$ for $n \geqslant 0$) for which $\mu(A) > 0$. By the a.s. martingale convergence theorem (see [Bil86, Theorem 35.5 and Exercise 35.17]), we have

$$\mu_{\xi_l(x)}(\xi_l(x) \cap A) \longrightarrow 1 \quad \text{as} \quad l \to \infty \quad \text{for a.e. } x \in A. \tag{12.5}$$

Take a point $x \in A$ satisfying both (12.1) and (12.5). Abbreviating $\tau_l := \tau_{x,l}$, we consider the probability densities given by $u_l := \mu(\xi_l(x))^{-1}\widehat{T}^{\tau_l}1_{\xi_l(x)}$, $v_l := \mu(\xi_l(x) \cap A)^{-1}\widehat{T}^{\tau_l}1_{\xi_l(x)\cap A}$, and $w_l := \mu(\xi_l(x) \setminus A)^{-1}\widehat{T}^{\tau_l}1_{\xi_l(x)\setminus A}$, $l \geqslant 1$. Since

$$v_l = u_l + \frac{\mu(\xi_l(x) \setminus A)}{\mu(\xi_l(x) \cap A)}(u_l - w_l) \quad \text{for} \quad l \geqslant 1,$$

where $u_l \in \mathfrak{U}$ and $\mu(\xi_l(x) \setminus A)/\mu(\xi_l(x) \cap A) \to 0$, there is some $l_0 \geqslant 1$ such that

$$\mathrm{dist}_{L_1(\mu)}(v_l, \mathfrak{U}) < \eta^2 \quad \text{for} \quad l \geqslant l_0.$$

Because $\xi_l(x) \cap A \subseteq A = T^{-\tau_l}A_{\tau_l}$, we have $\{v_l > 0\} \subseteq A_{\tau_l}$ (mod μ). Therefore (12.4) entails

$$\mu(A_{\tau_l}) \geqslant \mu(v_l > 0) \geqslant \mu(v_l > \eta) > \eta \qquad \text{for } l \geqslant l_0.$$

But $\mu(A) = \mu(A_{\tau_l})$ for all l since T preserves μ. This proves (12.2). □

BIBLIOGRAPHY

[Aar97] Jon Aaronson, *An introduction to infinite ergodic theory*, Mathematical Surveys and Monographs, vol. 50, American Mathematical Society, 1997. ↑1400

[Aba04] Miguel Abadi, *Sharp error terms and necessary conditions for exponential hitting times in mixing processes*, Ann. Probab. **32** (2004), no. 1A, 243–264. ↑1363

[AD01] Jon Aaronson and Manfred Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs–Markov maps*, Stoch. Dyn. **1** (2001), no. 2, 193–237. ↑1400

[AS11] Miguel Abadi and Benoît Saussol, *Hitting and returning to rare events for all alpha-mixing processes*, Stochastic Processes Appl. **121** (2011), no. 2, 314–323. ↑1363, 1366

[Bil86] Patrick Billingsley, *Probability and Measure*, second ed., Wiley Series in Probability and Mathematical Statistics: Probability and mathematical statistics, John Wiley & Sons, 1986. ↑1413

[Bil99] ———, *Convergence of Probability Measures*, second ed., Wiley Series in Probability and Statistics, John Wiley & Sons, 1999. ↑1367, 1377, 1382, 1383, 1397

[BSTV03] Henk Bruin, Benoît Saussol, Serge E. Troubetzkoy, and Sandro Vaienti, *Return time statistics via inducing*, Ergodic Theory Dyn. Syst. **23** (2003), no. 4, 991–1013. ↑1408

[CC13] Jean-René Chazottes and Pierre Collet, *Poisson approximation for the number of visits to balls in non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dyn. Syst. **33** (2013), no. 1, 49–80. ↑1363

[Doe40] Wolfgang Doeblin, *Remarques sur la théorie métrique des fractions continues*, Compos. Math. **7** (1940), 353–371. ↑1404

[Eag76] G. K. Eagleson, *Some simple conditions for limit theorems to be mixing*, Teor. Veroyatn. Primen. **21** (1976), 653–660. ↑1379

[FFT12] Ana Cristina M. Freitas, Jorge M. Freitas, and Mike Todd, *The extremal index, hitting time statistics and periodicity*, Adv. Math. **231** (2012), no. 5, 2626–2665. ↑1363

[FFTV16] Ana Cristina M. Freitas, Jorge M. Freitas, Mike Todd, and Sandro Vaienti, *Rare events for the Manneville–Pomeau map*, Stochastic Processes Appl. **126** (2016), no. 11, 3463–3479. ↑1408

[Gau12] Carl F. Gauss, *Letter to Laplace*, Göttingen, January 30^{th}, 1812. ↑1404

[HLV05] Nicolai T. Haydn, Yves Lacroix, and Sandro Vaienti, *Hitting and return times in ergodic dynamical systems*, Ann. Probab. **33** (2005), no. 5, 2043–2050. ↑1366, 1386

[Hol05] Mark Holland, *Slowly mixing systems and intermittency maps*, Ergodic Theory Dyn. Syst. **25** (2005), no. 1, 133–159. ↑1411

[HP14] Nicolai T. Haydn and Y. Psiloyenis, *Return times distribution for Markov towers with decay of correlations*, Nonlinearity **27** (2014), no. 6, 1323–1349. ↑1363

[HSV99] Masaki Hirata, Benoît Saussol, and Sandro Vaienti, *Statistics of Return Times: A General Framework and New Applications*, Commun. Math. Phys. **206** (1999), no. 1, 33–55. ↑1362

[HWZ14] Nicolai T. Haydn, Nicole Winterberg, and Roland Zweimüller, *Return-time statistics, Hitting-time statistics and Inducing*, Ergodic Theory, Open Dynamics, and Coherent Structures, Springer Proceedings in Mathematics & Statistics, vol. 70, Springer, 2014, pp. 217–227. ↑1408

[HY16] Nicolai T. Haydn and Fan Yang, *Entry times distribution for mixing systems*, J. Stat. Phys. **163** (2016), no. 2, 374–392. ↑1363

[IK02] Marius Iosifescu and Cor Kraaikamp, *Metrical Theory of Continued Fractions*, Mathematics and its Applications (Dordrecht), vol. 547, Kluwer Academic Publishers, 2002. ↑1404

[Ios77] Marius Iosifescu, *A Poisson law for ψ-mixing sequences establishing the truth of a Doeblin's statement*, Rev. Roum. Math. Pures Appl. **22** (1977), 1441–1447. ↑1404

[Kre85] Ulrich Krengel, *Ergodic Theorems*, De Gruyter Studies in Mathematics, vol. 6, Walter de Gruyter, 1985. ↑1379

[Mar17] Jens Marklof, *Entry and return times for semi-flows*, Nonlinearity **30** (2017), no. 2, 810–824. ↑1386

[PS16] Françoise Pène and Benoît Saussol, *Poisson law for some non-uniformly hyperbolic dynamical systems with polynomial rate of mixing*, Ergodic Theory Dyn. Syst. **36** (2016), no. 8, 2602–2626. ↑1363

[PS20] ———, *Spatio-temporal Poisson processes for visits to small sets*, Isr. J. Math. **240** (2020), no. 2, 625–665. ↑1363

[PSZ17] Françoise Pène, Benoît Saussol, and Roland Zweimüller, *Return- and hitting-time limits for rare events of null-recurrent Markov maps*, Ergodic Theory Dyn. Syst. **37** (2017), no. 1, 244–276. ↑1377

[PT20] Françoise Pène and Damien Thomine, *Potential kernel, hitting probabilities and distributional asymptotics*, Ergodic Theory Dyn. Syst. **40** (2020), no. 7, 1894–1967. ↑1403

[Rén58] Alfréd Rényi, *On mixing sequences of sets*, Acta Math. Acad. Sci. Hung. **9** (1958), 215–228. ↑1379

[Res08] Sidney I. Resnick, *Extreme Values, Regular Variation and Point Processes*, Springer Series in Operations Research and Financial Engineering, Springer, 2008. ↑1367, 1377

[Roh64] Vladimir A. Rohlin, *Exact endomorphisms of a Lebesgue space*, Am. Math. Soc., Transl., II. Ser. **39** (1964), 1–36. ↑1412

[RZ20] Simon Rechberger and Roland Zweimüller, *Return- and hitting-time distributions of small sets in infinite measure preserving systems*, Ergodic Theory Dyn. Syst. **40** (2020), no. 8, 2239–2273. ↑1377

[Tha80] Maximilian Thaler, *Estimates of the invariant densities of endomorphisms with indifferent fixed points*, Isr. J. Math. **37** (1980), 303–314. ↑1411

[Tha05] ———, *Asymptotic distributions and large deviations for iterated maps with an indifferent fixed point*, Stoch. Dyn. **5** (2005), no. 3, 425–440. ↑1411

[TK10] Marta Tyran-Kamińska, *Weak convergence to Lévy stable processes in dynamical systems*, Stoch. Dyn. **10** (2010), no. 2, 263–289. ↑1405

[Whi02] Ward Whitt, *Stochastic-Process Limits. An introduction to stochastic-process limits and their application to queue*, Springer Series in Operations Research, Springer, 2002. ↑1367

[Yos38] Kosaku Yosida, *Mean ergodic theorem in Banach spaces*, Proc. Imp. Acad. Japan **14** (1938), 292–294. ↑1379

[Zwe03] Roland Zweimüller, *Stable limits for probability preserving maps with indifferent fixed points*, Stoch. Dyn. **3** (2003), no. 1, 83–99. ↑1411

[Zwe07a] ———, *Infinite measure preserving transformations with compact first regeneration*, J. Anal. Math. **103** (2007), 93–131. ↑1377

[Zwe07b] ———, *Mixing limit theorems for ergodic transformations*, J. Theor. Probab. **20** (2007), no. 4, 1059–1071. ↑1365, 1372, 1379, 1385

[Zwe16] ———, *The general asymptotic return-time process*, Isr. J. Math. **212** (2016), no. 1, 1–36. ↑1365, 1366, 1372, 1375, 1386, 1387, 1391, 1397, 1398

[Zwe19] ———, *Hitting-time limits for some exceptional rare events of ergodic maps*, Stochastic Processes Appl. **129** (2019), no. 5, 1556–1567. ↑1408

Manuscript received on 7th April 2020,
revised on 6th January 2022,
accepted on 11th March 2022.

Recommended by Editors S. Gouëzel and Y. Coudène.

eISSN: 2644–9463
This journal is a member of Centre Mersenne.

Roland ZWEIMÜLLER
Fakultät für Mathematik,
Universität Wien,
Oskar-Morgenstern-Platz 1,
1090 Vienna (Austria)
roland.zweimueller@univie.ac.at
http://www.mat.univie.ac.at/~zweimueller/