Mixing limit theorems for ergodic transformations

(to appear in *Journal of Theoretical Probability)*

Roland Zweimüller
Faculty of Mathematics, University of Vienna
Nordbergstrasse 15, 1090 Vienna, Austria
roland.zweimueller@univie.ac.at

November 15, 2006

Abstract

We show that distributional and weak functional limit theorems for ergodic processes often hold for arbitrary absolutely continuous initial distributions. This principle is illustrated in the setup of ergodic sums, renewal-theoretic variables, and hitting times for ergodic measure preserving transformations.

Keywords: weak invariance principle, functional limit theorem, ergodic transformations, mean ergodic theorem, strong distributional convergence, mixing limit theorem, hitting times

1 Introduction

Probability theory of dynamical systems has become a very active field of research. A *nonsingular* map T on a σ-finite measure space $(X, \mathcal{A}, m)$ (meaning that T is measurable and $m \circ T^{-1} \ll m$), generates, for any *initial distribution* $P \ll m$ according to which an *initial state* $\mathsf{X}_0 \in X$ is chosen, a sequence of random elements $\mathsf{X}_n := T^n \mathsf{X}_0$, $n \geq 0$, of X representing the consecutive states of the system. Despite their nontrivial dependence structure, such $(\mathsf{X}_n)_{n \geq 0}$, and derived processes like, for example, $\mathsf{Y}_n := f(\mathsf{X}_n)$, $n \geq 0$, for suitable measurable $f : X \to \mathbb{R}$, often share the regular asymptotic behavior of more classical types of random processes. Typically, such results depend on T being *ergodic* (i.e. $A \in \mathcal{A}$ and $T^{-1}A = A$ imply $0 \in \{m(A), m(A^c)\}$) and the existence of a σ-finite absolutely continuous *invariant measure* μ (meaning $\mu \circ T^{-1} = \mu$), plus some additional assumptions ensuring a form of asymptotic independence.

While results of this kind are usually stated in terms of the invariant measure (if finite), there is no reason to regard the latter as a natural initial distribution P representing, say, incomplete but nontrivial information about the actual initial state of a deterministic system. It is therefore most desirable to extend such results to a large class of nonstationary probability measures P. While almost sure statements trivially carry over to every $P \ll m$, the question is more interesting for convergence in distribution, where (for finite m) validity for all $P \ll m$ is equivalent to the limit theorem being *mixing in the sense of Rényi*, cf. [R], [E], and [AE], which means that it persists under conditioning on any event of positive probability. Besides its obvious interest from the point of view of modelling, the possibility of changing measures in weak limit theorems has recently become a useful tool for establishing distributional limit theorems in the first place, cf. [TZ].

A simple and very useful sufficient condition for scalar processes (called *asymptotic T-invariance* below) has been introduced in [E]. While still not as well known as it ought to be, it has entered the ergodic theory literature, see e.g. [A0], [A1], [T], and [MT]. The present semi-expository note attempts to popularize this principle by offering an ergodic theoretical approach to Eagleson's result, and provides a slightly more general formulation which also covers functional limit theorems. We then collect a couple of natural applications, that is, types of distributional and functional limit theorems for ergodic transformations which are always mixing in Rényi's sense.

Throughout (M,d) is a separable metric space with Borel σ-field $\mathcal{B}_M$. A sequence $(\nu_n)_{n\geq 1}$ of probability measures on $(M,\mathcal{B}_M)$ *converges weakly* to the probability measure ν on $(M,\mathcal{B}_M)$, written $\nu_n \Longrightarrow \nu$, if the integrals of bounded continuous function $\psi : M \to \mathbb{R}$ converge, i.e. $\int \psi \, d\nu_n \longrightarrow \int \psi \, d\nu$ as $n \to \infty$. If R_n, $n \geq 1$, are random variables on a probability space $(X, \mathcal{A}, P)$, taking values in $(M,\mathcal{B}_M)$, and R is another *random element* of M, not necessarily defined on X, then $(R_n)_{n\geq 1}$ *converges in distribution* to R if the distributions $P \circ R_n^{-1}$ of the R_n converge weakly to that of R. Explicitly specifying the underlying measure, we denote this by

$$R_n \overset{P}{\Longrightarrow} R.$$

We are interested in situations in which a distributional limit theorem $R_n \overset{P}{\Longrightarrow} R$ automatically carries over to a large collection of other probability measures: For measurable maps R_n, $n \geq 1$, of a σ-finite measure space $(X,\mathcal{A},m)$ into $(M,\mathcal{B}_M)$, *strong distributional convergence* (terminology taken from §3.6 of [A0]) to a random element R, written

$$R_n \overset{\mathcal{L}(m)}{\Longrightarrow} R,$$

means that $R_n \overset{Q}{\Longrightarrow} R$ for all probability measures $Q \ll m$. If m itself is a probability measure, this is equivalent to $R_n \overset{m}{\Longrightarrow} R$ (*mixing*), meaning that $R_n \overset{m_E}{\Longrightarrow} R$ for every $E \in \mathcal{A}$ with $m(E) > 0$, where $m_E(A) := m(E \cap A)/m(E)$ is the conditional measure on E.

The following result enables us to change initial distributions in a large variety of interesting situations.

Theorem 1 (Asymptotically invariant random elements of M) *Let R_n, $n \geq 1$, be random variables on the probability space $(X, \mathcal{A}, P)$, taking values in the separable metric space (M, d), and R a random element of M such that*

$$R_n \overset{P}{\Longrightarrow} R \qquad \text{as } n \to \infty.$$

Assume there is a σ-finite measure m on $(X, \mathcal{A})$ such that $P \ll m$, and an ergodic nonsingular transformation T on $(X, \mathcal{A}, m)$ under which $(R_n)_{n \geq 1}$ is asymptotically invariant in measure *in that*

$$d(R_n \circ T, R_n) \overset{m}{\longrightarrow} 0 \qquad \text{as } n \to \infty. \tag{1}$$

Then

$$R_n \overset{\mathcal{L}(m)}{\Longrightarrow} R \qquad \text{as } n \to \infty.$$

Remark 1 ***a)*** *Convergence in measure, $\overset{m}{\longrightarrow}$, w.r.t. a σ-finite measure is to be understood as $\overset{Q}{\longrightarrow}$ for all probability measures $Q \ll m$.*

b) *A parallel statement for continuous-parameter families $(R_t)_{t \geq 0}$ which satisfy $d(R_t \circ T, R_t) \overset{m}{\longrightarrow} 0$ as $t \to \infty$ follows by exactly the same argument.*

c) *If m happens to be infinite, we can w.l.o.g. pass to an equivalent probability measure (w.r.t. which T is nonsingular) without spoiling (1).*

d) *For real-valued R_n, the result is contained in [E] (but our proof is different). An example of a mixing functional limit theorem (i.e. for path-valued R_n), in the context of Donsker's invariance principle for independent and strongly (α-) mixing processes, is discussed in [B] (cf. §14 and §19 there), see also § VIII.5 of [JS].*

e) *A related result for a.s.-versions of distributional limit theorems is given in §4 of [ChG].*

2 Proof of Theorem 1

The *transfer operator* $\widehat{T} : L_1(m) \to L_1(m)$ of the nonsingular map T on $(X, \mathcal{A}, m)$ describes the evolution of probability densities under T, that is, $\widehat{T}u := d(P \circ T^{-1})/dm$, where P has density u w.r.t. m. Equivalently, $\int (g \circ T) \cdot u \, dm = \int g \cdot \widehat{T}u \, dm$ for all $u \in L_1(m)$ and $g \in L_\infty(m)$, i.e. $\widehat{T}$ is dual to $g \longmapsto g \circ T$. We let $\mathcal{D}(m)$ denote the set of probability densities w.r.t. m. The following classical companion of the mean ergodic theorem, due to Yosida [Y] (see also [Kr], Theorem 2.1.3), is central to our argument:

Theorem 2 (Yosida's Theorem) *Let T be a nonsingular map on a σ-finite measure space $(X, \mathcal{A}, m)$. Then T is ergodic if and only if*

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} \widehat{T}^k (u - u^*) \right\|_{L_1(m)} \longrightarrow 0 \qquad \text{for all } u, u^* \in \mathcal{D}(m). \tag{2}$$

We are now ready for the

Proof of Theorem 1. Fix any probability measure $Q \ll m$ on $(X, \mathcal{A})$. To prove $R_n \overset{Q}{\Longrightarrow} R$ it is enough to show that

$$\int \psi \, d(Q \circ R_n^{-1}) = \int \psi \circ R_n \, dQ \longrightarrow \mathbb{E}[\psi(R)] \quad \text{as } n \to \infty \tag{3}$$

whenever $\psi : M \to \mathbb{R}$ is bounded and (uniformly) Lipschitz. Fix such a ψ. Since, by assumption, the corresponding statement is true for P, (3) is equivalent to

$$\int \psi \circ R_n \, dP - \int \psi \circ R_n \, dQ \longrightarrow 0 \quad \text{as } n \to \infty. \tag{4}$$

Choose any $\varepsilon > 0$ and let $u := dP/dm$, and $u^* := dQ/dm$. Replacing P and Q by K-step arithmetic averages generated by T, with $K = K(\varepsilon)$ sufficiently large, we obtain, via (2),

$$\begin{aligned} &\left| \int \psi \circ R_n \, d\left(\frac{1}{K} \sum_{k=0}^{K-1} P \circ T^{-k} \right) - \int \psi \circ R_n \, d\left(\frac{1}{K} \sum_{k=0}^{K-1} Q \circ T^{-k} \right) \right| \\ &\qquad = \left| \int \psi \circ R_n \cdot \left(\frac{1}{K} \sum_{k=0}^{K-1} \widehat{T}^k (u - u^*) \right) dm \right| \\ &\qquad \le \sup_M |\psi| \cdot \left\| \frac{1}{K} \sum_{k=0}^{K-1} \widehat{T}^k (u - u^*) \right\|_{L_1(m)} < \frac{\varepsilon}{2} \quad \text{for all } n \ge 1. \end{aligned}$$

The assertion of the theorem therefore follows once we verify

$$\left| \int \psi \circ R_n \, d\left(\frac{1}{K} \sum_{k=0}^{K-1} P \circ T^{-k} \right) - \int \psi \circ R_n \, dP \right| < \frac{\varepsilon}{4} \quad \text{for } n \ge n_0(\varepsilon), \tag{5}$$

as well as the analogous statement with P replaced by Q.

Rewrite the left-hand side of (5) as $\left| \frac{1}{K} \sum_{k=0}^{K-1} \int (\psi \circ R_n \circ T^k - \psi \circ R_n) \, dP \right|$ to see that it suffices to prove that for every $j \ge 0$ and $v \in \mathcal{D}(m)$,

$$\int (\psi \circ R_n \circ T^{j+1} - \psi \circ R_n \circ T^j) \cdot v \, dm \longrightarrow 0 \quad \text{as } n \to \infty. \tag{6}$$

Since T is nonsingular, we may assume w.l.o.g. that $j = 0$. Let $\Psi_n := \psi \circ R_n$, and observe that due to Lipschitz continuity of ψ the sequence $(\Psi_n)_{n \ge 1}$ inherits asymptotic T-invariance in measure from $(R_n)_{n \ge 1}$, that is,

$$\Psi_n \circ T - \Psi_n \overset{m}{\longrightarrow} 0 \quad \text{as } n \to \infty. \tag{7}$$

Now take any $\varepsilon' > 0$, then

$$\int |\Psi_n \circ T - \Psi_n| \cdot v \, dm \le \varepsilon' + 2 \sup_M |\psi| \cdot \int_{\{|\Psi_n \circ T - \Psi_n| > \varepsilon'\}} v \, dm,$$

and because of (7) the right-most term tends to 0 as $n \to \infty$. This proves (6) and hence our theorem. ■

3 Applications

We collect a number of interesting situations in which Theorem 1 can, with little effort, be applied to give strong distributional convergence in distributional limit theorems and weak invariance principles for ergodic processes.

3.1 Ergodic sums and partial sum processes

The most obvious examples of asymptotically invariant sequences (R_n) for a conservative ergodic nonsingular map T on $(X, \mathcal{A}, m)$ are normalized ergodic sums $\sum_{k=0}^{n-1} f \circ T^k$, $n \geq 1$. For the reader's convenience we recall the most basic observation (cf. [E]):

Corollary 1 (Ergodic sums in $\mathbb{R}$) *Let T be a nonsingular conservative ergodic map on the probability space $(X, \mathcal{A}, m)$, $P \ll m$ some probability measure, and $f : X \to \mathbb{R}$ measurable. For constants $A_n \in \mathbb{R}$, $B_n > 0$, $n \geq 1$, consider the normalized ergodic sums*

$$S_n : X \to \mathbb{R},\ n \geq 1, \qquad S_n(x) := \frac{1}{B_n} \left(\sum_{k=0}^{n-1} f \circ T^k(x) - A_n \right).$$

Then, for any $\mathbb{R}$-valued random variable S,

$$S_n \overset{P}{\Longrightarrow} S \quad \textit{implies} \quad S_n \overset{\mathcal{L}(m)}{\Longrightarrow} S.$$

Proof. Asymptotic T-invariance $S_n \circ T - S_n \overset{m}{\longrightarrow} 0$ is obvious. ■

Turning to functional limit theorems, we denote functions to be regarded as elements of a path space M by x, y and their values at t by $\mathsf{x}_t, \mathsf{y}_t$. Similarly, random elements of M will be denoted by $\mathsf{S}, \mathsf{S}_n, \ldots$ with corresponding coordinate variables $\mathsf{S}_t, \mathsf{S}_{n,t}$ etc. As a simple warm-up, consider the space $\mathcal{C}[0,1]$ of continuous real functions on $[0,1]$, with the *uniform metric* $d_{\mathcal{C}}(\mathsf{x}, \mathsf{y}) := \sup_{t \in [0,1]} |\mathsf{x}_t - \mathsf{y}_t|$, $\mathsf{x}, \mathsf{y} \in \mathcal{C}[0,1]$. See [B] for an exposition of weak convergence theory for Borel probabilities on $\mathcal{C}[0,1]$. Most prominently, Donsker's invariance principle asserts distributional convergence in $\mathcal{C}[0,1]$ of (affinely interpolated) partial sum processes for iid sequences of square integrable random variables to a *Wiener process* $\mathsf{W} = (\mathsf{W}_t)_{t \in [0,1]}$ (understood to be a random element of $\mathcal{C}[0,1]$). Functional CLTs have been studied for a large variety of ergodic dynamical systems. We check that, for general ergodic stationary sequences, such a limit theorem, if valid, automatically holds for all absolutely continuous probabilities:

Corollary 2 (Ergodic sums in $\mathcal{C}$) *Let T be a measure preserving ergodic map on the probability space $(X, \mathcal{A}, m)$, $P \ll m$ some probability measure, $\sigma > 0$,* *and* $f \in L_2(m)$. *Consider the partial sum processes* $\mathsf{S}_n : X \to \mathcal{C}[0,1]$, $n \geq 1$, *given by*

$$\mathsf{S}_{n,t}(x) := \frac{1}{\sigma\sqrt{n}} \left(\sum_{k=0}^{\lfloor tn \rfloor -1} f \circ T^k(x) + \frac{t - \lfloor tn \rfloor / n}{1/n} \cdot (f \circ T^{\lfloor tn \rfloor}) \right).$$

Then

$$\mathsf{S}_n \overset{P}{\Longrightarrow} \mathsf{W} \text{ in } \mathcal{C}[0,1] \quad \text{implies} \quad \mathsf{S}_n \overset{\mathcal{L}(m)}{\Longrightarrow} \mathsf{W} \text{ in } \mathcal{C}[0,1].$$

Proof. In view of Theorem 1 we only need to check that $d_{\mathcal{C}}(\mathsf{S}_n \circ T, \mathsf{S}_n) \overset{m}{\longrightarrow} 0$ as $n \to \infty$, which is easily done:

Since we trivially have $d_{\mathcal{C}}(\mathsf{S}_n \circ T, \mathsf{S}_n) = \max_{0 \leq j \leq n} \left| (\mathsf{S}_n \circ T)_{j/n} - \mathsf{S}_{n,j/n} \right|$ and $\left| (\mathsf{S}_n \circ T)_{j/n} - \mathsf{S}_{n,j/n} \right| \leq (|f| + \left| f \circ T^j \right|)/\sigma\sqrt{n}$, we just need to show that

$$\frac{\max_{0 \leq j \leq n} \left| f \circ T^j \right|}{\sigma\sqrt{n}} \overset{m}{\longrightarrow} 0. \tag{8}$$

But for any $\varepsilon > 0$, we have

$$\begin{aligned} m\left(\left\{ \frac{\max_{0 \leq j \leq n} \left| f \circ T^j \right|}{\sigma\sqrt{n}} \geq \varepsilon \right\} \right) &\leq (n+1) \cdot m\left(\left\{ |f| \geq \varepsilon\sigma\sqrt{n} \right\} \right) \\ &\leq \frac{2}{(\sigma\varepsilon)^2} \int_{\left\{ |f| \geq \varepsilon\sigma\sqrt{n} \right\}} f^2 \, dm, \end{aligned}$$

and since $f \in L_2(m)$, the rightmost integral tends to zero as $n \to \infty$. ■

For more general limit theorems, consider processes with unbounded time interval and càdlàg paths: Let $\mathcal{D}[0,\infty)$ denote the spaces of right-continuous real functions $\mathsf{x} : [0,\infty) \to \mathbb{R}$ possessing left limits everywhere. We equip this space with the usual *Skorohod (J_1-) topology* (cf. [S]), which we introduce using a variant of the *Skorohod metric* given in [P], §VI.1. Given $\mathsf{x}, \mathsf{y} \in \mathcal{D}[0,\infty)$ we first define, for any $M > 0$, the distance $d_{\mathcal{D},M}(\mathsf{x},\mathsf{y})$ as the infimum of all $\delta > 0$ such that there exist some $k \in \mathbb{N}$ and points $0 = s_0 < s_1 < \ldots < s_k$ and $0 = t_0 < t_1 < \ldots < t_k$ with $s_k, t_k \geq M$, $|s_i - t_i| \leq \delta$ for $0 \leq i \leq k$, and

$$|\mathsf{x}_s - \mathsf{y}_t| \leq \delta \quad \text{if } s \in [s_i, s_{i+1}) \text{ and } t \in [t_i, t_{i+1}) \text{ for some } 0 \leq i < k.$$

The metric on $\mathcal{D}[0,\infty)$ we are going to use is given by

$$d_{\mathcal{D}}(\mathsf{x},\mathsf{y}) := \sum_{M \geq 1} 2^{-M} (1 \wedge d_{\mathcal{D},M}(\mathsf{x},\mathsf{y})), \quad \mathsf{x},\mathsf{y} \in \mathcal{D}[0,\infty).$$

For background information on the Skorohod space thus defined, we also refer to [B]. Extending the previous result, we show that distributional convergence of normalized ergodic sums in $\mathcal{D}[0,\infty)$ automatically implies strong distributional convergence:

Corollary 3 (Ergodic sums in $\mathcal{D}$) *Let T be a nonsingular conservative ergodic map on the σ-finite space $(X, \mathcal{A}, m)$, $P \ll m$ some probability measure, and $f : X \to \mathbb{R}$ measurable. For constants $A_n \in \mathbb{R}$, $B_n > 0$, $n \geq 1$, consider the normalized ergodic sum processes given by*

$$\mathsf{S}_n : X \to \mathcal{D}[0,\infty),\ n \geq 1, \quad \mathsf{S}_{n,t}(x) := \frac{1}{B_n}\left(\sum_{k=0}^{\lfloor tn \rfloor -1} f \circ T^k(x) - A_n\right).$$

Then, for any random element S of $\mathcal{D}[0,\infty)$,

$$\mathsf{S}_n \stackrel{P}{\Longrightarrow} \mathsf{S} \text{ in } \mathcal{D}[0,\infty) \quad \text{implies} \quad \mathsf{S}_n \stackrel{\mathcal{L}(m)}{\Longrightarrow} \mathsf{S} \text{ in } \mathcal{D}[0,\infty).$$

Proof. We check that $(\mathsf{S}_n)_{n\geq 1}$ is asymptotically T-invariant in measure, i.e. that

$$d_{\mathcal{D}}(\mathsf{S}_n \circ T, \mathsf{S}_n) \stackrel{m}{\longrightarrow} 0 \quad \text{as } n \to \infty. \tag{9}$$

Fix any $M \in \mathbb{N}$ and $n \geq 2$. To estimate $d_{\mathcal{D},M}(\mathsf{S}_n \circ T, \mathsf{S}_n)$, consider the points $s_i := i/n$ and $t_i := (i+1)/n$ for $1 \leq i \leq Mn =: k$, which clearly satisfy $|s_i - t_i| \leq 1/n$. Since, for $s \in [0, M]$,

$$(\mathsf{S}_n \circ T)_s = \frac{1}{B_n}\left(\sum_{k=1}^{\lfloor (s+\frac{1}{n})n \rfloor -1} f \circ T^k(x) - A_n\right) = \mathsf{S}_{n,s+\frac{1}{n}} - \frac{f}{B_n},$$

we see that

$$|(\mathsf{S}_n \circ T)_s - \mathsf{S}_{n,t}| \leq \frac{|f|}{B_n} \qquad \text{if } s \in [s_i, s_{i+1}) \text{ and } t \in [t_i, t_{i+1}) \text{ for some } 0 \leq i < k.$$

Given any $\delta > 0$, we therefore have $d_{\mathcal{D},M}(\mathsf{S}_n \circ T, \mathsf{S}_n) \leq \delta$ on $\{|f|/B_n \leq \delta\}$ for $n \geq 1/\delta$. As M does not show up in this estimate, the same is true for $d_{\mathcal{D}}$, that is,

$$d_{\mathcal{D}}(\mathsf{S}_n \circ T, \mathsf{S}_n) \leq \delta \quad \text{on } \left\{\frac{|f|}{B_n} \leq \delta\right\} \text{ for } n \geq 1/\delta.$$

By conservativity and our assumption $\mathsf{S}_n \stackrel{P}{\Longrightarrow} \mathsf{S}$, we have $B_n \to \infty$, showing that $m(\{|f|/B_n \leq \delta\}) \to 1$ as $n \to \infty$, and (9) follows. Now use Theorem 1 to conclude the proof. ■

Remark 2 ***a)*** *Using $d_{\mathcal{D},M}(\mathsf{x}, \mathsf{y}) \leq \sup_{[0,M]} |\mathsf{x} - \mathsf{y}|$ we could have tried to argue as before. However, the analogue of (8) breaks down, even in iid situations, if f has heavy tails: In fact, if the distribution of f is the one-sided stable law of order $\alpha \in (0,1)$, $M_n := \max_{0\leq j\leq n} |f \circ T^j|$, and $S_n := \sum_{k=0}^{n-1} f \circ T^k$, then S_n/M_n has a non-degenerate limit distribution, cf. [Da].*

b) *Due to the use of the clever metric from [P], the argument is as simple as it should be. It takes a less pleasant form if, instead, we work with the more common metric to be found e.g. in [B], §16 (which, of course, induces the same topology).*

In particular, weak convergence in $\mathcal{D}[0,\infty)$ of partial sum processes of stationary ergodic sequences to non-Gaussian stable processes automatically extends to all absolutely continuous measures. For a different example, take an infinite measure preserving transformation T satisfying a functional version of the Darling-Kac theorem for the occupation times $\sum_{k=0}^{n-1} 1_Y \circ T^k$ of some set with $0 < \mu(Y) < \infty$, cf. [A2], that is, convergence w.r.t. μ_Y in $\mathcal{D}[0,\infty)$ to a *Mittag-Leffler process* $\mathsf{M} = (\mathsf{M}_t)_{t\geq 0}$ of order $\alpha \in (0,1)$ (the inverse of the stable subordinator). The corollary shows that this functional convergence holds for every $Q \ll \mu$.

3.2 A renewal theoretic invariance principle

Here is another application to null-recurrent situations (i.e. maps T preserving an infinite measure μ): For a fixed set $Y \in \mathcal{A}$, $\mu(Y) > 0$, we define the $\mathbb{N}_0$-valued variables $Z_n(Y)$, $n \geq 0$, on X by $Z_n(Y)(x) := \max(\{0\} \cup \{1 \leq k \leq n : T^k x \in Y\})$, which gives the time of the last visit to Y up to time n. The Dynkin-Lamperti arcsine law for waiting times describes the asymptotic behaviour of these renewal-theoretic quantities in infinite measure preserving situations, see [Dy], [L1] for the classical setup, and [T], [TZ] for ergodic transformations and strong distributional convergence. In [L2] Lamperti has given a renewal-theoretic invariance principle which can be reformulated as a functional limit theorem for the Z_n. Here, too, our simple test applies:

Corollary 4 (Waiting times in $\mathcal{D}$) *Let T be a nonsingular conservative ergodic map on the σ-finite space $(X, \mathcal{A}, m)$, $P \ll m$ some probability measure, and $Y \in \mathcal{A}$, $m(Y) > 0$. Consider the normalized waiting-time processes given by*

$$\mathsf{Z}_n : X \to \mathcal{D}[0,\infty),\ n \geq 1, \quad \mathsf{Z}_{n,t}(x) := \frac{1}{n} Z_{\lfloor tn \rfloor}(Y).$$

Then, for any random element Z of $\mathcal{D}[0,\infty)$,

$$\mathsf{Z}_n \overset{P}{\Longrightarrow} \mathsf{Z} \text{ in } \mathcal{D}[0,\infty) \quad \text{implies} \quad \mathsf{Z}_n \overset{\mathcal{L}(m)}{\Longrightarrow} \mathsf{Z} \text{ in } \mathcal{D}[0,\infty).$$

Proof. We check asymptotic T-invariance of $(\mathsf{Z}_n)_{n\geq 1}$, $d_{\mathcal{D}}(\mathsf{Z}_n \circ T, \mathsf{Z}_n) \overset{m}{\longrightarrow} 0$ as $n \to \infty$. As in the proof of Corollary 3 it is easy to estimate $d_{\mathcal{D},M}(\mathsf{Z}_n \circ T, \mathsf{Z}_n)$ for $M \in \mathbb{N}$. Consider the same set of points, $s_i := i/n$ and $t_i := (i+1)/n$ for $1 \leq i \leq Mn =: k$. Abbreviating $Z_n := Z_n(Y)$, we note that

$$Z_n \circ T = Z_{n+1} - 1 \quad \text{on } \bigcup\nolimits_{j=1}^{n} T^{-j} Y.$$

Therefore, for $s \in [0, M]$, we have

$$(\mathsf{Z}_n \circ T)_s = \frac{Z_{\lfloor sn \rfloor} \circ T}{n} = \frac{Z_{\lfloor sn+1 \rfloor} - 1}{n} = \mathsf{Z}_{n,s+1/n} - \frac{1}{n} \quad \text{on } \bigcup\nolimits_{j=1}^{n} T^{-j} Y,$$

and hence

$$|(\mathsf{Z}_n \circ T)_s - \mathsf{Z}_{n,t}| \leq \frac{1}{n} \qquad \text{on } \textstyle\bigcup_{j=1}^{n} T^{-j} Y, \text{ if } s \in [s_i, s_{i+1}) \text{ and } t \in [t_i, t_{i+1}) \text{ for some } 0 \leq i < k.$$

As $\bigcup_{j=1}^{n} T^{-j}Y \nearrow X$ mod m, asymptotic T-invariance follows as before. ■

3.3 Hitting times and associated point processes

We turn to a different circle of questions which has recently attracted a lot of attention. Let T be a conservative ergodic nonsingular map on the σ-finite measure space $(X, \mathcal{A}, m)$. For sets $Y \in \mathcal{A}$ of non-zero measure define $\varphi_Y(x) := \inf\{n \geq 1 : T^n x \in A\}$, $x \in X$, which is finite a.e. When restricted to Y, φ_Y is the *return-time* function of Y, while it is usually called the *entrance-time* or *hitting-time* function when regarded as a function on X. The restriction of the map $T_Y : X \to Y$, $T_Y x := T^{\varphi_Y(x)}x$, to Y is the *first-return map* of Y. If $Y_n \in \mathcal{A}$, $n \geq 1$, are sets of positive measure with $Y_n \searrow \varnothing$ or $m(Y_n) \to 0$, we can think of $(Y_n)_{n\geq 1}$ as a sequence of *asymptotically rare events* and study, for some fixed initial probability $P \ll m$, the distributions of their hitting times φ_{Y_n} as $n \to \infty$. It has been shown that for a large variety of transformations with reasonable mixing conditions these distributions do converge (after normalization) to an exponential distribution. Relevant references include [H], [HSV], [AG], and [KL]. For situations with different limit laws see e.g. [BZ].

Again, results of this type are usually stated in terms of an invariant measure μ, or some other particular initial distribution P (Lebesgue measure on an interval, say), and sometimes extended to a family of very regular initial distributions $Q \ll m$ (see e.g. [CoG]). We show that they automatically hold for all $Q \ll m$ (for which Eagleson's result [E] suffices):

Corollary 5 (Hitting-times) *Let T be a conservative ergodic nonsingular map on the σ-finite space $(X, \mathcal{A}, m)$, and $P \ll m$ some probability measure. Let $Y_n \in \mathcal{A}$, and $\gamma_n > 0$, $n \geq 1$, with $\gamma_n \to \infty$. Assume that $Y_n \searrow \varnothing$ or that m is T-invariant with $m(Y_n) \to 0$. Consider the normalized hitting times*

$$R_n : X \to [0, \infty],\ n \geq 1, \qquad R_n(x) := \gamma_n^{-1} \cdot \varphi_{Y_n}(x).$$

Then, for any random variable R taking values in $[0, \infty]$,

$$R_n \overset{P}{\Longrightarrow} R \quad \textit{implies} \quad R_n \overset{\mathcal{L}(m)}{\Longrightarrow} R.$$

Proof. To check asymptotic T-invariance of $(R_n)_{n\geq 1}$, observe that for any set Y of positive measure,

$$\varphi_Y \circ T = \varphi_Y - 1 \quad \text{on } \{\varphi_Y > 1\} = T^{-1}Y^c. \tag{10}$$

Consequently, $|R_n \circ T - R_n| = 1/\gamma_n$ on $T^{-1}Y_n^c$, and by our assumptions on $(Y_n)_{n\geq 1}$ we have $Q(T^{-1}Y_n) \to 0$ as $n \to \infty$ for every probability measure $Q \ll m$, which proves $R_n \circ T - R_n \overset{m}{\longrightarrow} 0$, and hence Corollary 5. ■

Let $M := \mathbf{M}_p([0, \infty))$, the space of counting measures on $([0, \infty), \mathcal{B}_{[0,\infty)})$, that is, measures $\nu : \mathcal{B}_{[0,\infty)} \to \mathbb{N}_0$. Equipped with the topology of *vague convergence*, meaning that $\nu_n \to v$ in M iff $\nu_n(f) \to v(f)$ for every continuous function

$f : [0, \infty) \to \mathbb{R}$ with compact support, M is a Polish space (has a complete and separable metric d). A *point process* N on $[0, \infty)$ is a random element of M.

We are interested in the point processes N_n given by the successive visits to the target sets Y_n: The *interarrival times* between consecutive visits of to Y_n are

$$\varphi_{Y_n}\,,\ \varphi_{Y_n} \circ T_{Y_n}\,,\ \varphi_{Y_n} \circ T^2_{Y_n}\,,\ \ldots$$

and the actual *arrival times* are the $\varphi_{Y_n,k} := \sum_{j=0}^{k-1} \varphi_{Y_n} \circ T^j_{Y_n}$, $k \geq 1$. The number of visits to Y_n within the time set $B \in \mathcal{B}_{[0,\infty)}$ defines the point process $N_n : X \to \mathbf{M}_p([0, \infty))$, $N_n(B) := \sum_{k \geq 1} 1_B(\varphi_{Y_n,k})$. Under fairly general assumptions, functional versions of the hitting-time limit theorems mentioned above are available, asserting distributional convergence (after normalization) of the N_n to some limiting point process (usually Poisson), see e.g. [H].

Corollary 6 (Hitting-time processes) *Let T be a conservative ergodic non-singular map on the σ-finite space $(X, \mathcal{A}, m)$, and $P \ll m$ some probability measure. Let $Y_n \in \mathcal{A}$, and $\gamma_n > 0$, $n \geq 1$, with $\gamma_n \to \infty$. Assume that $Y_n \searrow \varnothing$ or that m is T-invariant with $m(Y_n) \to 0$. Consider the rescaled point processes*

$$\mathsf{N}_n : X \to \mathbf{M}_p([0, \infty)),\ n \geq 1, \quad \mathsf{N}_n(B) := \sum_{j \geq 0} 1_B \left(\frac{\varphi_{Y_n,j}(x)}{\gamma_n} \right),$$

Then, for any random element N of $\mathbf{M}_p([0, \infty))$,

$$\mathsf{N}_n \stackrel{P}{\Longrightarrow} \mathsf{N} \text{ in } \mathbf{M}_p([0, \infty)) \quad \text{implies} \quad \mathsf{N}_n \stackrel{\mathcal{L}(m)}{\Longrightarrow} \mathsf{N} \text{ in } \mathbf{M}_p([0, \infty)).$$

Proof. Recall that (according to standard results, see e.g. [Ka]) convergence $\mathsf{N}_n \stackrel{P}{\Longrightarrow} \mathsf{N}$ is equivalent to convergence of finite tuples of interarrival times, i.e. to the assertion that for every $K \geq 1$,

$$\gamma_n^{-1} \cdot \left(\varphi_{Y_n}, \varphi_{Y_n} \circ T_{Y_n}, \ldots, \varphi_{Y_n} \circ T^{K-1}_{Y_n} \right) \stackrel{P}{\Longrightarrow} (\tau_0, \tau_1, \ldots, \tau_{K-1})\,, \tag{11}$$

where the τ_j, $j \geq 0$, are the interarrival times of the limit process N. It is therefore sufficient to show that, for $K \geq 1$,

$$\gamma_n^{-1} \cdot \left(\varphi_{Y_n}, \varphi_{Y_n} \circ T_{Y_n}, \ldots, \varphi_{Y_n} \circ T^{K-1}_{Y_n} \right) \stackrel{\mathcal{L}(m)}{\Longrightarrow} (\tau_0, \tau_1, \ldots, \tau_{K-1})\,.$$

Since it is easily seen that for any Y of positive measure,

$$T^k_Y \circ T = T^k_Y \quad \text{for } k \geq 1 \text{ on } \{\varphi_Y > 1\} = T^{-1}Y^c,$$

we obtain, recalling (10) above,

$$\sup_{k \geq 1} \left| \varphi_Y \circ T^k_Y \circ T - \varphi_Y \circ T^k_Y \right| = 1 \quad \text{on } T^{-1}Y^c. \tag{12}$$

Now fix any $K \geq 1$ and consider the random vector $\mathsf{R}_n : X \to \mathbb{R}^K$ on the left-hand side of (11). Equipping $\mathbb{R}^K$ with the metric d corresponding to the sup-norm, we see from (12) that

$$d(\mathsf{R}_n \circ T, \mathsf{R}_n) = 1/\gamma_n \quad \text{on } T^{-1}Y_n^c,$$

and $d(\mathsf{R}_n \circ T, \mathsf{R}_n) \stackrel{m}{\longrightarrow} 0$ follows as in the proof of Corollary 5. ■

Remark 3 *Analogous statements (with analogous proofs) hold for continuous-parameter families $(Y_\epsilon)_{\epsilon>0}$ of measurable sets with $\mu(Y_\epsilon) \to 0$ or $Y_\epsilon \searrow \varnothing$ as $\epsilon \searrow 0$.*

3.4 Appendix: A compactness proof of Yosida's theorem

We finally take the opportunity to point out an alternative to the usual Hahn-Banach proof of the crucial characterization of ergodicity via (2). It is analogous to the proof of Lin's characterization of exactness, cf. [L] or Theorem 1.3.3 in [A0]. In the present context it is of interest for the following reason: The proof of Eagleson's result in [E] relies on a "compactness plus subsequence-in-subsequence" argument (for the sequence of characteristic functions of the variables R_n under consideration), and so does the argument given in Section 3.6 of [A0] (for sequences of continuous functions of the R_n). Due to the use of Yosida's theorem, which also gives a good intuitive understanding of "why the result is true", our proof of Theorem 1 avoids this type of reasoning. In a sense, however, the compactness argument has only been hidden behind the statement of Yosida's result:

Proof of Yosida's Theorem. The other direction being trivial, we assume that the nonsingular map T on the σ-finite measure space $(X, \mathcal{A}, m)$ is ergodic, and fix any $w \in L_1(m)$ with $\int_X w\, dm = 0$. Let $W_n := n^{-1} \sum_{k=0}^{n-1} \widehat{T}^k w$ and $G_n := \mathrm{sgn} W_n$, $n \in \mathbb{N}$, so that $\|G_n\|_\infty \leq 1$ and

$$\left\| \frac{1}{n} \sum_{k=0}^{n-1} \widehat{T}^k w \right\|_1 = \int_X G_n \cdot W_n \, dm = \int_X D_n \cdot w \, dm,$$

where the $D_n := n^{-1} \sum_{k=0}^{n-1} G_n \circ T^k$, $n \in \mathbb{N}$, satisfy $\|D_n\|_\infty \leq 1$. We need to prove

$$\int_X D_n \cdot w \, dm \longrightarrow 0 \quad \text{as } n \to \infty. \tag{13}$$

To do so, we choose some partition $\gamma \subseteq \mathcal{A}$ of X into sets of finite positive measure, and let $\mathcal{A}_w \subseteq \mathcal{A}$ denote the σ-field generated by γ, w, and the D_n, $n \in \mathbb{N}$. As $(X, \mathcal{A}_w, m\,|_{\mathcal{A}_w})$ is σ-finite and countably generated, $L_1(m\,|_{\mathcal{A}_w})$ is separable, so that (see, e.g. Proposition II.A.15 of [W]) the closed unit ball $\mathfrak{U}_\infty$ of its dual space $L_\infty(m\,|_{\mathcal{A}_w})$ is metrizable in the weak*-topology. Together with Alaoglu's theorem (Theorem II.A.9 of [W]), this shows that $\mathfrak{U}_\infty$ is weak* sequentially compact. Let D be any weak* limit point of the sequence $(D_n)_{n\in\mathbb{N}}$ in $\mathfrak{U}_\infty$, i.e. suppose there are $n_j \nearrow \infty$ such that

$$\int_X D_{n_j} \cdot f \, dm \longrightarrow \int_X D \cdot f \, dm \quad \text{for all } f \in L_1(m\,|_{\mathcal{A}_w}). \tag{14}$$

We are going to show that any such D is a.e. constant, implying in particular that

$$\int_X D_{n_j} \cdot w \, dm \longrightarrow D \cdot \int_X w \, dm = 0 \quad \text{as } j \to \infty.$$

A subsequence-in-subsequence argument based on sequential compactness of $\mathfrak{U}_\infty$ then yields (13).

By ergodicity, we need only check that any limit point D as above is T-invariant (mod m) to see that it is constant (mod m). According to (14) we also have

$$\int_X D_{n_j} \circ T \cdot f \, dm = \int_X D_{n_j} \cdot \widehat{T} f \, dm \longrightarrow \int_X D \cdot \widehat{T} f \, dm = \int_X D \circ T \cdot f \, dm$$

as $j \to \infty$, for all $f \in L_1(m \mid_{\mathcal{A}_w})$, but

$$\begin{aligned} \left| \int_X (D_{n_j} \circ T - D_{n_j}) \cdot f \, dm \right| &= \left| \frac{1}{n_j} \int_X \left(G_{n_j} - G_{n_j} \circ T^{n_j} \right) \cdot f \, dm \right| \\ &\leq \frac{2 \, \| f \|_1}{n_j} \to 0, \end{aligned}$$

so that the respective limits coincide, i.e. $\int_X D \circ T \cdot f \, dm = \int_X D \cdot f \, dm$ for all $f \in L_1(m \mid_{\mathcal{A}_w})$, proving $D \circ T = D$ a.e. as required. ■

Acknowledgement. This research was supported by an APART [Austrian programme for advanced research and technology] fellowship of the Austrian Academy of Sciences.

References

[A0] J. Aaronson: *An Introduction to Infinite Ergodic Theory.* AMS 1997.

[A1] J. Aaronson: *The asymptotic distributional behaviour of transformations preserving infinite measures.* Journal d'Analyse Mathematique **39** (1981), 203-234.

[A2] J. Aaronson: *Random f-expansions.* Ann. Probab. **14** (1986), 1037-1057.

[AG] M. Abadi, A. Galves: *Inequalities for the occurrence times of rare events in mixing processes. The state of the art.* Markov. Proc. Relat. Fields **7** (2001), 97-112.

[AE] D.J. Aldous, G.K. Eagleson: *On mixing and stability of limit theorems.* Ann. Probab. **6** (1978), 325-331.

[B] P. Billingsley: *Convergence of Probability Measures.* 2nd ed. Wiley 1999.

[BZ] X. Bressaud, R. Zweimüller: *Non exponential law of entrance times in asymptotically rare events for intermittent maps with infinite invariant measure.* Ann. H. Poincaré **2** (2001), 501-512.

[ChG] J.-R. Chazottes, S. Gouëzel: *On almost-sure versions of classical limit theorems for dynamical systems.* Probab. Theory Relat. Fields, *to appear.*

[CoG] P. Collet, A. Galves: *Statistics of close visits to the indifferent fixed point of an interval map.* J. Stat. Phys. **72** (1993), 459-478.

[Da] D.A. Darling: *The influence of the maximal term in the addition of independent random variables.* Trans. Amer. Math. Soc. **73** (1952), 95-107.

[DS] N. Dunford, J. T. Schwartz: *Linear operators I.* Reprint, Wiley 1988.

[Dy] E. B. Dynkin: *Some limit theorems for sums of independent random variables with infinite mathematical expectation.* Selected Transl. in Math. Statist. and Probability **1** (1961), 171-189.

[E] G. K. Eagleson: *Some simple conditions for limit theorems to be mixing.* Teor. Verojatnost. i Primenen. **21** (1976), 653-660.

[H] M. Hirata: *Poisson law for Axiom A diffeomorphisms.* Ergod. Th. & Dynam. Sys. **13** (1993), 533-556.

[HSV] M. Hirata, B. Saussol, S. Vaienti: *Statistics of return times: A general framework and new applications.* Commun. Math. Phys. **206** (1999), 33-55.

[JS] J. Jacod, A.N. Shiryaev: *Limit theorems for Stochastic Processes.* Springer 1987.

[Ka] A.F. Karr: *Point Processes and their Statistical Inference.* Marcel Dekker 1986.

[Kr] U. Krengel: *Ergodic Theorems.* de Gruyter 1985.

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

[L1] J. Lamperti: *Some limit theorems for stochastic processes.* J. Math. Mech. **7** (1958), 433-448.

[L2] J. Lamperti: *An invariance principle in renewal theory.* Ann. Math. Statist. **33** (1962), 685-696.

[L] M. Lin: *Mixing for Markov operators.* Z. Wahrsch. verw. Geb. **19** (1971), 231-243.

[MT] I. Melbourne, A. Török: *Statistical limit theorems for suspension flows.* Israel J. Math. **144** (2004), 191-209.

[P] D. Pollard: *Convergence of Stochastic Processes*. Springer 1984.

[R] A. Rényi: *On mixing sequences of sets*. Acta Math. Acad. Sci. Hung. **9** (1958), 3-4.

[S] A.V. Skorohod: *Limit theorems for random processes*. Teor. Verojatnost. i Primenen. **1** (1956), 289-319.

[T] M. Thaler: *The Dynkin-Lamperti Arc-Sine Laws for Measure Preserving Transformations.* Trans. Amer. Math. Soc. **350** (1998), 4593-4607.

[TZ] M. Thaler, R. Zweimüller: *Distributional limit theorems in infinite ergodic theory*. Probab. Theory Relat. Fields **135** (2006), 15-52.

[W] P. Wojtaszczyk: *Banach spaces for analysts*. Cambridge UP, 1991.

[Y] K. Yosida: *Mean ergodic theorem in Banach spaces*. Proc. Imp. Acad. Tokyo **14** (1938), 292-294.