YOKOHAMA MATHEMATICAL
JOURNAL VOL. 38, 1990

# ON LIMIT THEOREMS AND CATEGORY FOR DYNAMICAL SYSTEMS

By

DALIBOR VOLNÝ

(Received Sept. 18, 1989; Revised Jan. 20, 1990)

**Abstract.** Let a probability space and a 1-1 bimeasurable and measure preserving transformation be given. For a measurable function $f$, the process $(f\circ T^{i})$ is strictly stationary. Let us consider $L^{p}$ spaces of functions $f$ and their subsets which are determined by the limit behavior of the process $(f\circ T^{i})$ from the point of view of the central limit problem and the speed of convergence in the Ergodic Theorem. It is shown that the set of processes (functions $f$) with highly irregular behavior is of second category.

**Introduction and results.** Given a dynamical system $(\Omega, \mathcal{A}, T, \mu)$ where $(\Omega, \mathcal{A}, \mu)$ is a probability space and $T$ is a 1-1 bimeasurable and measure preserving mapping of $\Omega$ onto itself we are often interested in limit behavior of the process $(f\circ T^{i}, i\in \boldsymbol{Z})$, $f$ measurable. Each sequence of functions $f\circ T^{i}$, $i\in \boldsymbol{Z}$ is strictly stationary and each strictly stationary sequence of random variables can be represented in this way.

The laws of large numbers and the central limit theorem belong to the most important limit theorems in probability theory. In 1932 and 1931 J. von Neumann's and G. Birkhoff's ergodic theorems were proved (see [1]): For $f$ integrable,

$$\frac{1}{n}\sum_{j=1}^{n} f\circ T^{j} \longrightarrow E(f\mid \mathcal{I})$$

almost surely, and in $L^{1}$; if $f$ is square integrable then the convergence holds in $L^{2}$, too. By $\mathcal{I}$ we denote the $\sigma$-algebra of all invariant sets $A\in\mathcal{A}$, i.e. such that $T^{-1}A=A$. If each element of $\mathcal{I}$ has measure 0 or 1, we say that $\mu$ is ergodic.

In the last three decades an increasing attention has been devoted to the central limit theorem for stationary processes. By the central limit theorem we understand weak convergence of distributions of

---

American Mathematical Society 1980 subject classification: Primary 60F05, 60F15, 28D05; secondary 60G10.

Key words and phrases: the central limit problem for strictly stationary processes, speed of convergence in the Ergodic Theorem.

$$\frac{1}{a_n}\sum_{j=1}^{n} f\circ T^j, \qquad a_n\to\infty$$

to a probability law. From the ergodic theorems it follows that if $\liminf a_n/n>0$, the problem is trivial. In the sequel we shall suppose that $a_n/n\to 0$ holds. From the ergodic theorem and from $a_n\to\infty$ it follows that for almost every $\omega\in\{|E(f|\mathcal{I})|>0\}$, $|(a_n/n)\sum_{j=1}^{n} f(T^j\omega)|\to\infty$. So, we shall restrict ourselves to the Hilbert space $\mathcal{H}$ of all square integrable functions $f$ for which $E(f|\mathcal{I})=0$ *a.s.*, i.e. to the orthocomplement of the subspace of invariant functions in $L^2$. In the central limit problem, many results concerning special processes have been reached. In 1987, R. Burton and M. Denker ([4]) proved that for $\mu$ aperiodic and ergodic, the set of $f$ such that for $(f\circ T^i)$ the central limit theorem holds is dense in $\mathcal{H}$. (We say that $\mu$ is aperiodic if for each positive integer $n$ and set $A\in\mathcal{A}$ of positive measure there exists a measurable set $B\subset A$ with $\mu(B\cap T^{-n}B)<\mu(B)$.) Unlike the strong law of large numbers the central limit theorem does not hold for stationary processes generally. R.C. Bradley even showed that the set of limit points of distributions of $(a_n/n)\sum_{j=1}^{n} f\circ T^j$ can comprise the set of all infinitely divisible laws for a process which satisfies rather strong mixing conditions ([3]). We shall see that the set of $f\in\mathcal{H}$ such that for the sequence $(f\circ T^i)$ the central limit theorem does not hold, is of second category.

It is well known that the central limit problem is related to the rate of convergence in the law of large numbers. In 1978 U. Krengel showed that in von Neumann's and Birkhoff's ergodic theorems the rate of convergence can be arbitrarily slow. Using the results in the central limit problem we shall show that the slow rate is in some sense typical (see also [5]).

Let us formulate the results exactly.

**Theorem 1.** *Let $\mu$ be an aperiodic measure and $a_n\to\infty$, $a_n/n\to 0$. Then there exists a dense $G_\delta$ set of functions $f\in\mathcal{H}$ with the property that for each probability distribution $\nu$ on the real line there exists an increasing sequence of positive integers $n_k$ such that the distributions of*

$$\frac{1}{a_{n_k}}\sum_{j=1}^{n_k} f\circ T^j$$

*weakly converge to $\nu$.*

In the central limit thecrem we often normalize the sums by $\|\sum_{j=1}^{n} f\circ T^j\|_2$ where $\|f\|_2=(Ef^2)^{1/2}$; in this setting we have:

**Theorem 2.** *Under the assumptions of Theorem* 1 *for each $f$ from a dense $G_\delta$ subset of $\mathcal{H}$ the set of all weak limit points of distributions of*

$$\sum_{j=1}^{n} f\circ T^j \Big/ \Big\| \sum_{j=1}^{n} f\circ T^j \Big\|_2$$

*is equal to the set of all probability measures* $\nu$ *on the real line such that* $\int x^2 d\nu(x) \leqq 1$ *and* $\int x d\nu(x)=0$.

**Theorem 3.** *Let* $\mu$ *be an aperiodic measure,* $\alpha_n \to \infty$. *Then for each* $f$ *from a dense* $G_\delta$ *subset of* $\mathcal{H}$

(1) $$\limsup_{n\to\infty} \alpha_n \left| \frac{1}{n} \sum_{j=1}^{n} f\circ T^j \right| = \infty \qquad a.s.$$

*and*

(2) $$\limsup_{n\to\infty} \alpha_n \left\| \frac{1}{n} \sum_{j=1}^{n} f\circ T^j \right\|_2 = \infty .$$

**Remark.** Let $1 \leqq p < \infty$. If we define $\mathcal{H}$ as the Banach space of functions $f \in L^p$ with $E(f|\mathcal{I})=0$ *a.s.* and use $L^p$-norm instead of $L^2$-norm, Theorem 1 and Theorem 3 would remain valid. In Theorem 2 we would have to consider the set of probability measures $\nu$ with $\int |x|^p d\nu(x) \leqq 1$, $\int x d\nu(x)=0$. The proofs remain almost the same as in the $L^2$ case. An $L^\infty$ version of Theorem 3 can be found in [11] (see also [5]).

**Proofs.** First, we shall prove a simple lemma.

**Lemma.** *Let* $U$ *be an operator on* $\mathcal{H}$ *defined by* $Uf=f\circ T$, $I$ *be the identity operator. Then the image of* $I-U$ *is dense in* $\mathcal{H}$.

**Proof.** Can be found e.g. in [8], p. 40. We shall give it for $L^p$ spaces where $1 \leqq p < \infty$. Clearly, $(I-U)\mathcal{H}$ is a linear space. Let there exist $f \in \mathcal{H}$ which does not belong to the closure of $(I-U)\mathcal{H}$. Then there exists a function $g \in L^q$, $\frac{1}{p}+\frac{1}{q}=1$, such that $\int g\cdot(U-I)h d\mu=0$ for each $h \in \mathcal{H}$ and $\int gf d\mu \neq 0$. From the first property it follows that $g=U^{-1}g$, hence $g$ is $\mathcal{I}$-measurable (see [1]). As $E(f|\mathcal{I})=0$ *a.s.*, we have $\int fg d\mu=0$ which contradicts our supposition, q.e.d.

The Lemma does not hold in $L^\infty$. There exists a finitely additive invariant measure $\nu$ which is zero on the sets of $\mu$-measure zero and which is not countably additive. In the ergodic case this follows from [12], Theorem 11.8; the general aperiodic case can be derived in the same way with the help of Lebesgue measure space techniques given in [9], [10]. (We have to use a nonergodic version of a result of del Junco and Rosenblatt; it is proved in [11].)

For $g=f-Uf$ where $f\in L^{\infty}$ we have $\int g d\nu=0$ and there exists $h\in L^{\infty}$ with $0=Eh=E(h\,|\,\mathcal{I})$, $\int h d\nu\neq 0$.

**Proof of Theorem 1.** We shall find a countable dense subset $\Gamma$ of the space of all probability measures on the real line with the topology of weak convergence such that for each $\nu\in\Gamma$ and for each $f\in\mathcal{H}$ up to a set of the first category it holds:

(3) In the sequence of distributions of $\frac{1}{a_n}S_n(f)$ there exists a subsequence weakly converging to $\nu$; $S_n$ denotes $\sum_{j=1}^{n} f\circ T^j$.

For any fixed function $f\in\mathcal{H}$ the set of distributions $\nu$ such that (3) holds is closed with respect to the weak convergence. Thus, it is sufficient to prove that for each $\nu\in\Gamma$ the set of functions $f\in\mathcal{H}$ satisfying (3) is residual in $\mathcal{H}$.

First, we shall find $\Gamma$. Measure $\mu$ is nonatomic, hence for each probability distribution on $\boldsymbol{R}$ there exists a measurable function $h$ on $\Omega$ such that $\nu=\mu\circ h^{-1}$. The space of all probability distributions on $\boldsymbol{R}$ is separable so we need only countably many functions $h$. For each measurable function $h$ there exists a sequence $h_1, h_2, \cdots$ of functions from $\mathcal{H}$ such that $h_n\to h$ *a.s.* The measures $\mu\circ h_n^{-1}$ then weakly converge to $\mu\circ h^{-1}$. For $\Gamma$ we can thus choose a countable set of measures $\mu\circ g^{-1}$ where functions $g$ form a dense subset of $\mathcal{H}$.

Let $\nu\in\Gamma$ be given, $g\in\mathcal{H}$, $\nu=\mu\circ g^{-1}$. We can assume that $\|g\|_2>0$. We shall prove that the set of functions $f\in\mathcal{H}$ which satisfy (3), is a dense $G_\delta$.

Following [6], p. 200, for each $\varepsilon>0$ and a positive integer $k$ there exists a Rohlin tower $(F, n, \varepsilon)$ where $F\in\mathcal{A}$, the sets $F, T^{-1}F, \cdots, T^{-n+1}F$ are mutually disjoint, and

$$n\mu(F)>1-\varepsilon,\quad k<n,\quad k\mu(F)<\varepsilon.$$

Let us denote $\mu_F=\mu(\cdot\,|\,F)$. There exists a measurable function $f$ on $\Omega$ such that

$$\mu_F(mf)^{-1}=\nu \qquad \text{where} \quad m=\frac{k}{a_k},$$

$$f(\omega)=f(T^j\omega) \qquad \text{for} \quad \omega\in T^{-j}F,\quad 1\leqq j\leqq n-1,$$

$$f(\omega)=0 \qquad \text{for} \quad \omega\in\Omega\setminus\bigcup_{i=0}^{n-1}T^{-i}F.$$

On $\Omega\setminus G$ where $G=\Omega\setminus\bigcup_{j=k}^{n-1}T^{-j}F$ it holds $S_k(f)=kf$ hence for $k\leqq j\leqq n-1$ we have

$$\int_{T^{-j}F}\exp\Big(\frac{it}{a_k}S_k(f)\Big)d\mu=\int_{T^{-j}F}\exp(itmf)\,d\mu=\mu(F)E\exp(itg)$$

Therefore,

$$E\exp\Big(\frac{it}{a_k}S_k(f)\Big)=\int_G \exp\Big(\frac{it}{a_k}S_k(f)\Big)d\mu+(n-k)\mu(F)E\exp(itg)$$

hence

$$\Big|E\exp\Big(\frac{it}{a_k}S_k(f)\Big)-E\exp(itg)\Big|<4\varepsilon$$

as $1-(n-k+1)\mu(F)<2\varepsilon$.

In this way we can find a sequence of functions $f_k\in\mathcal{H}$ sucn that $|E\exp((it/a_k)S_k(f_k))-\int e^{itx}d\nu(x)|<4\varepsilon_k$, $\varepsilon_k\to 0$. From the definition of $f$ it follows that $E(mf)^2=n\mu(F)Eg^2$, hence $\|f\|_2<(1/m)\|g\|_2$. As $(k/a_k)\to\infty$, it holds $\|f_k\|_2\to 0$. For each $k=1, 2, \cdots$ there exists a neighborhood $\mathcal{V}(f_k)$ of $f_k$ such that for each $h\in\mathcal{V}(f_k)$ it is $\|h-f_k\|_2<1$ and $|E_\mu\exp((it/a_k)S_k(f_k))-E_\mu\exp((it/a_k)S_k(h))|<(1/k)$ for $t\in[-k, k]$. Thus, there exists an open ball $\mathcal{K}=\{h\in\mathcal{H}: \|h\|_2<r\}$, $r<\infty$, $\mathcal{V}(f_k)\subset\mathcal{K}$ for each $k$. Let $c_1, c_2, \cdots$ be positive reals, $c_k\to\infty$ and $(c_k/a_k)\downarrow 0$. For each $k$ we define $\mathcal{V}_k=\{h\in\mathcal{H}: \|h\|_2\leqq c_k\}$. For $k=1, 2, \cdots$ we define

$$H_k=\bigcup_{n=k}^{\infty}\{\mathcal{V}(f_n)+(U-I)\mathcal{V}_n\}$$

where $U$ is defined by $Uf=f\circ T$ and we put

$$H=\bigcap_{k=1}^{\infty}H_k\,.$$

The sets $\mathcal{V}(f_n)+(U-I)\mathcal{V}_n$ consist of functions $h=h'+h''$ where $h'\in\mathcal{V}(f_n)$, $h''=Ug-g$ for some $g\in\mathcal{V}_n$. Hence, each $H_k$ is an open set. According to the Lemma the image of $U-I$ is dense in $\mathcal{H}$. As $\mathcal{V}(f_n)\subset\mathcal{K}$ and $(U-I)\mathcal{V}_n\uparrow(U-I)\mathcal{H}$, each $H_k$ is dense in $\mathcal{H}$. Hence, $H$ is a dense $G_\delta$ set.

Let $f\in H$. From $f\in H_k$ it follows that there exists $n\geqq k$ and functions $h'_n\in\mathcal{V}(f_n)$, $g_n\in\mathcal{V}_n$ such that $f=h'_n+Ug_n-g_n$. It holds

$$\|S_n(Ug_n-g_n)/a_n\|_2=\frac{1}{a_n}\Big\|\sum_{j=1}^{n}U^j(Ug_n-g_n)\Big\|_2$$

$$=\frac{1}{a_n}\|U^n g_n-g_n\|_2\leqq 2c_k/a_k\longrightarrow 0,$$

$$\Big|E\exp\Big(\frac{it}{a_n}S_n(h'_n)\Big)-\int e^{itx}d\nu(x)\Big|<4\varepsilon_n+\frac{1}{n}$$

for $t\in[-n, n]$. Thus there exists a sequence of positive integers $n_k\to\infty$ such that for each $k$, $f=h'_{n_k}+h''_{n_k}$, $\mu((1/a_{n_k})S_{n_k}(h'_{n_k}))^{-1}\to\nu$ weakly, $\|(1/a_{n_k})S_{n_k}(h''_{n_k})\|_2\to 0$. According to [2], Theorem 4.4, the distributions of $(1/a_{n_k})S_{n_k}(f)$ weakly converge to $\nu$, q.e.d.

**Proof of Theorem 2.** Let $G$ denote the set of all probability measures $\nu$ on the real line such that $\int x d\nu(x)=0$, $\int x^2 d\nu(x)=1$. Using the same approach as in the proof of Theorem 1 we can see that for all $f$ from a dense $G_\delta$ subset of $\mathscr{H}$, the set of the weak limit points of distributions of $\sum_{j=1}^n f\circ T^j/\|\sum_{j=1}^n f\circ T^j\|_2$ contains $G$. It remains to show that a probability measure $\nu$ belongs to the weak closure of $G$ if and only if $\int x d\nu(x)=0$, $\int x^2 d\nu(x)\leqq 1$ (the distributions of $\sum_{j=1}^n f\circ T^j/\|\sum_{j=1}^n f\circ T^j\|_2$ belong to $G$).

Let $\nu_n\in G$, $\nu_n\to\nu$ weakly. For each bounded continuous function $h$ it holds $\int h d\nu_n\to\int h d\nu$; from this we can easily derive that $\int x^2 d\nu(x)$ cannot exceed 1. Let us suppose that $\int x d\nu(x)=a$, $a\neq 0$. Without loss of generality we can assume $a>0$. For each $a>\varepsilon>0$ there exist numbers $m\in\boldsymbol{N}$ and $N>1/(a-\varepsilon)$ such that $\int_{-N}^{N} x d\nu_m(x)\in(a-\varepsilon, a+\varepsilon)$. As $\int x d\nu_m(x)=0$, it holds $\int_{-\infty}^{-N}|x|d\nu_m(x)>a-\varepsilon$. Hence, $\int x^2 d\nu_m(x)>N(a-\varepsilon)>1$, which contradicts $\nu_m\in G$. Thus, $\int x d\nu(x)=0$.

Let $\nu$ be a probability distribution on $\boldsymbol{R}$, $\int x d\nu(x)$ and $\int x^2 d\nu(x)\leqq 1$. For every natural number $k$ we can easily find a measure $\nu_k\in G$ and a Borel set $A_k$ such that $\nu(A_k)>1-1/k$ and $\nu_k(B)=\nu(B)$ for each Borel set $B\subseteqq A$. Measures $\nu_k$ weakly converge to $\nu$, which finishes the proof.

**Proof of Theorem 3.** Let $(\alpha_n)$ be a sequence of positive real numbers, $\alpha_n\to\infty$. Without loss of generality we can restrict ourselves to sequences which are increasing "sufficiently slowly", so we can assume that $\alpha_n/n\to 0$. Let us put $a_n=n/\alpha_n$. From Theorem 1 it follows that for each $f$ from a dense $G_\delta$ subset of $\mathscr{H}$ and real number $K$ there exist positive integers $n_k\uparrow\infty$,

$$\frac{1}{a_{n_k}}\sum_{j=1}^{n_k} f\circ T^j \longrightarrow K \quad a.s.$$

From this we get the statement of the theorem.

The author thanks Professor Anzelm Iwanik from Technical University Wroclaw and Professor Ivan Netuka from Charler University Prague for their kind and helpful advices.

## References

[1] Billingsley, P.: *Ergodic Theory and Information*, Wiley, New York, London, Sydney, 1965.

[2] Billingsley, P.: *Convergence of Probability Measures*, Wiley, New York, 1968.
[3] Bradley, R. C.: Information regularity and the central limit question, *Rocky Mountain Journal of Math.* **13** (1983), 77-97.
[4] Burton, R. and Denker, M.: On the central limit therem for dynamical systems, *Trans. American Math. Soc.* **302** No. **2** (1987), 715-726.
[5] del Junco, A. and Rosenblatt, J.: Counterexamples in ergodic theory and in number theory, *Math. Ann.* **245** (1979), 185-197.
[6] Kornfeld, I. P., Sinai, Ja. G. and Fomin, S. V.: *Ergodic Theory* (in Russian), Nauka, Moskva, 1980.
[7] Krengel, U.: On the speed of convergence in the ergodic theorem, *Mh. Math.* **86** (1978), 3-6.
[8] Lamperti, J.: *Stochastic Processes*, Springer, New York, 1973.
[9] Rohlin, V. A.: Selected topics in the theory of dynamical systems, *Uspekhi Mat.* **4** (1949), 57-128.
[10] Rohlin, V. I.: On the fundamental ideas of measure theory, *Mat. Sbornik* (N. S.) **25** (1949), 107-150.
[11] Volý, D.: *in preparation.*
[12] Wagon, S.: *The Banach-Tarski Paradox*, Cambridge Univ. Press, Cambridge, 1985.

Mathematical Institute of Charles University
Sokolovská 83
186 00 Praha 8
Czechoslovakia.