Math. Slovaca, 43 (1993), No. 3, 277–292

MILD LAW OF LARGE NUMBERS AND ITS CONSEQUENCES

JAROSLAV MOHAPL

(*Communicated by Milan Medveď*)

ABSTRACT. Let (X,d) be a metric space and let $T\colon X \to X$ be a continuous mapping. If $\lim\limits_{n\to\infty} \frac{1}{n} \sum\limits_{k=1}^{n} f\circ T^k$ exists and defines a uniformly continuous function on X whenever f is uniformly continuous, we say that T satisfies the Mild Law of Large Numbers $(MLLN)$. Sufficient conditions under which $MLLN$ holds are given. Ergodic theorems for deterministic and stochastic dynamical systems with trajectories taking values in X following from $MLLN$ are derived. For stochastic systems existence of invariant measures is also discussed.

1. Introduction

The results of this paper concern ergodic theorems and laws of large numbers for deterministic as well as for stochastic dynamical systems. They will be formulated as a consequence of what we can call a mild law of large numbers $(MLLN)$. Although the nature of this law is non-probabilistic, it yields an interesting insight into the structure of stochastic systems and of the classical strong law of large numbers $(SLLN)$. The most important result guarantees a possibility to decompose the phase space into disjoint sets on which the system has a regular ergodic behaving provided the appropriate invariant σ-smooth measures supported by these sets exist. The reader interested in the application of this result is referred to [7], where the existence of such a decomposition is the basic hypothesis (cf. [8, Ch. 14]).

In order to present the results we introduce the notation used throughout the paper: (X,d) – a separable metric space, $C(X)$ – the Banach space of all real bounded continuous functions on (X,d) with the supremum norm, $U_d(X)$ and $L_d(X)$ – the subspaces of all uniformly continuous and Lipschitz functions from $C(X)$, respectively, $\mathcal{B}(X)$ – the algebra of Baire sets that is generated by

AMS Subject Classification (1991): Primary 28D05, 60F05, 60J05, 58F11.
Key words: Contractive map, Invariant measure, Markov process, Law of large numbers.

the open subsets of X, $M(X)$ – the class of all Baire measures on $\mathcal{B}(X)$ with finite variation representing the norm dual of $C(X)$, $M_\sigma(X)$ – the subspace of all σ-smooth measures in $M(X)$, w – the weak topology of the duality pair $(M_\sigma(X), C(X))$, (T^t) – a flow of continuous mappings from X into X.

We will prefer to write T instead T^1 and I (identity) for T^0. Sometimes instead (X,d), $x,y,\dots \in X$ symbols like (S,ϱ), $s,\bar{s}\dots \in S$ will be used and then we will write $C(S)$, $U_\varrho(S)$, $L_\varrho(S)$ etc. As to the used terminology the reader is referred to [1, 4, 15, 19].

2. *MLLN* for contractive mappings

In this section T will denote a contractive map. We call T *contractive* provided $d(T^n x, T^n y) \le c_T d(x,y)$ for a finite constant c_T that is common for all $x,y \in X$ and all natural numbers n. In the case when $c_T < 1$, the well-known assertion [5, Ch. 2, §4] states that there is just one $x_0 \in X$ such that $\lim T^n x = x_0$ for each $x \in X$ and, consequently, $\lim f \circ T^n x = f x_0$ for each $f \in C(X)$. Obviously, x_0 is the *fixed point* for T, i.e. $T x_0 = x_0$, and the relation $\lim f \circ T^n x = f x_0$ is stronger then the relation $\lim \frac{1}{n} \sum_{k=1}^{n} f \circ T^k x = f x_0$, the *SLLN*. The point (Dirac) measure δ_{x_0} with mass concentrated in x_0 is the unique invariant ergodic measure for T, that is, $\delta_{x_0} T^{-1} = \delta_{x_0}$ and $\delta_{x_0} J = 0$ or 1 for each $J \in \mathcal{B}(X)$ with the property $TJ = J$. These facts remain to be true if we replace (T^n) by any more general flow (T^t) and the operators $\frac{1}{n} \sum_{k=1}^{n}$ by $\frac{1}{t} \int_0^t \mathrm{d}s$. The situation $1 \le c_T < \infty$ is rather complicated as can be shown by numerous examples, but

THEOREM 2.1. *If T is a contractive map of (X,d) into (X,d), then for each $f \in U_d(X)$ $\left(\frac{1}{n} \sum_{k=1}^{n} f \circ T^k\right)$ is a Cauchy sequence in $U_d(X)$. If in addition T belongs to a flow (T^t), then $\left(\frac{1}{t} \int_0^t f \circ T^s \,\mathrm{d}s\right)$ is also a Cauchy sequence in $U_d(X)$.*

The assertion of Theorem 2.1 will be called the *MLLN* and, as it will be shown later, it can hold also for non-contractive systems.

Provided the *MLLN* holds, the map $f \to f^*$, where $f \in U_d(X)$ and $f^* = \lim \frac{1}{n} \sum_{k=1}^{n} f \circ T^k$ defines a bounded linear functional from $U_d(X)$ into $U_d(X)$. Using the Riesz representation theorem we can to each $x \in X$ relate a measure

$\delta_x^* \in M(X)$ (not necessarily Dirac or σ-smooth) such that $f^*x = \delta_x^* f$ for all $f \in U_d(X)$ (we write mf instead $\int fx\, m(\mathrm{d}x)$ for $m \in M(X)$ and measurable f).

THEOREM 2.2. *The relation $\equiv$ defined for each couple $x, y \in X$ by $x \equiv y$ if and only if $\delta_x^* = \delta_y^*$ is an equivalence. If $\mathcal{E}$ is the class of equivalences defined by $\equiv$, then the sets in $\mathcal{E}$ are measurable and invariant.*

Proof. $f^* \in U_d(X)$ whenever $f \in U_d(X)$ and therefore $\{x : f^*x = pf, x \in X\}$ is a closed set for each $p \in M(X)$ (perhaps empty). Therefore $\{x : \delta_x^* = p, \ x \in X\} = \bigcap_{f \in U_d(X)} \{x : f^*x = pf, \ x \in X\}$ is a closed, obviously invariant, set.

The class $\mathcal{E}$ forms the ergodic decomposition of X. Its main feature is that for each $f \in U_d(X)$ the restriction of f^* to $I \in \mathcal{E}$ is a constant function.

THEOREM 2.3. *If $p \in M_\sigma(X)$ and $I = \{x : \ \delta_x^* = p, \ x \in X\} = \emptyset$, then the restriction of p to $\mathcal{B}(I)$ is an ergodic invariant measure for T_I, $T_I \colon I \to I$, where T_I is the restriction of T to I.*

Proof. If $p \in M_\sigma(X)$, and if the $MLLN$ holds, then by Alexandrov's theorem [16, Ch. 7] $\limsup \frac{1}{n} \sum_{k=1}^{n} \chi_F \circ T^k x \le pF$ for each $x \in X$ and each closed set $F \subseteq I$. But then, using the invariantness of p and Fatou's lemma, we obtain that

$$pF = \lim \frac{1}{n} \sum_{k=1}^{n} \int_I \chi_F \circ T^k x \, p(\mathrm{d}x) \le \int_I \limsup \frac{1}{n} \sum_{k=1}^{n} \chi_F \circ T^k x \, p(\mathrm{d}x)$$

$$\le pFpI \le (pI)^2,$$

and as p is a regular measure, $pI = \sup\{pF : \ F \subseteq I, \ \ F \text{ closed}\} \le (pI)^2$ which can hold if and only if $pI = 0$, or $pI = 1$ $(0 \le p \le 1)$.

Since the procedure remains to be valid for any invariant subset of I, the ergodicity of p is proved.

Theorems 2.1 – 2.3 state, that if for each $x \in X$ the sequence $\left(\frac{1}{n} \sum_{k=1}^{n} \delta_x T^{-k}\right) \subset M_\sigma(X)$ is relatively compact in $\left(M_\sigma(X), w\right)$ then X can be decomposed into in some sense maximal invariant subsets on which the restriction of T has a unique ergodic invariant measure.

Note that here and latter we will use the important conjunction $\chi_E \circ Tx = \delta_x T^{-1} E$ that holds for each $x \in X$ and $E \in \mathcal{B}(X)$. In this notation $\frac{1}{n} \sum_{k=1}^{n} \delta_x T^{-1}$

is what M e y n in [10] calls the occupation probability or what can be called the empirical probability. The $MLLN$ states that these empirical probabilities always converge in some sense to a possibly finitely additive measure with $p\emptyset = 0$, $pX = 1$ and $0 \le p \le 1$. Value of this assertion can be well checked if we realize that it may be $\delta_x^* \ne \delta_y^*$ if $x \ne y$ (as an example take $T = I$). Immediately from [20, Ch. 13, Thm. 2] and from the $MLLN$ we can obtain this version of Birkhoff's ergodic theorem (cf. with [10, Proposition 4.2]).

THEOREM 2.4. *If T is a contractive map from (X, d) into (X, d), then the following conditions are equivalent:*

i) *there exists an invariant measure $p \in M(X)$ for T such that for each measurable function f $\left(\frac{1}{n} \sum_{k=1}^{n} f \circ T^k x\right)$ is a Cauchy sequence for p almost all $x \in X$;*

ii) *for at last one $x \in X$ $\left(\frac{1}{n} \sum_{k=1}^{n} \delta_x T^{-k}\right)$ is relatively compact in $\big(M_\sigma(X), C(X)\big)$.*

Now it is time to prove the basic Theorem 2.1. For this purpose we introduce the sets $L(c, d)$ that are defined by

$$L(c, d) = \big\{f : \ |fx - fy| \le c d(x, y) \ \text{ for all } \ x, y \in X \ \text{ and } \ \|f\| \le c\big\}.$$

Their specific feature is that they are sequentially compact in the topology of pointwise convergence on X.

P r o o f o f T h e o r e m 2.1. Each separable metric space X is due to the known Urysohn's theorem [4, Ch. 4, Thm. 16] homeomorphic with a dense subset $\overline{X}_0$ of a compact metric space $(\overline{X}, \overline{d})$. The homeomorphism H of $\overline{X}_0$ onto X defines a metric d_c on X by the relation $d_c(x, y) = \overline{d}(H^{-1}x, H^{-1}y)$. A continuous map from a metric compact into a metric space is uniformly continuous, hence,

$$\forall_{\varepsilon>0} \ \exists_{\delta>0} \ \forall_{x,y\in X} \ d_c(x, y) < \delta \implies d(x, y) < \varepsilon. \tag{$*$}$$

This follows from the equations $d_c(x, y) = \overline{d}(H^{-1}x, H^{-1}y) = \overline{d}(\overline{x}, \overline{y})$ and $d(H\overline{x}, H\overline{y}) = d(x, y)$ provided we put $x = H\overline{x}$ and $y = H\overline{y}$.

Let $f \in L(1, d)$. Then $\left(\frac{1}{n} \sum_{k=1}^{n} f \circ T^k\right) \subset L(c_T, d)$ has a pointwise convergent subsequence, denoted for simplicity again $\left(\frac{1}{n} \sum_{k=1}^{n} f \circ T^k\right)$, with a limit function $f^* \in L(c_T, d)$. Let $\varepsilon > 0$ be a given number and $\delta > 0$ be the number from

$(*)$. To δ we can find a finite sequence $(x_p) \subset X$ such that each x has any x_p lying in the circle $\{y : d_c(x, y) < \delta, \quad y \in X\}$. If we choose n_0 so that $\left|\frac{1}{n}\sum\limits_{k=1}^{n} f \circ T^k x_p - f^* x_p\right| < \varepsilon$ for all $x_p \in (x_p)$ and n greater than some n_0, then

$$\left|\frac{1}{n}\sum_{k=1}^{n} f \circ T^k x - f^* x\right| \le c_T d(x, x_p) + \left|\frac{1}{n}\sum_{k=1}^{n} f \circ T^k x_p - f^* x_p\right| + c_T d(x_p, x),$$

and the right-hand expression is by $(*)$ less than $2c_T\varepsilon + \varepsilon$ for all $n > n_0$ if x_p is sufficiently near to x. This shows that our subsequence converges uniformly to f^*.

If we define the operator U_T on $U_d(X)$ by $U_T f = f \circ T$, then U_T is linear and its operator norm is less or equal to one, hence, by [20, Ch. 8] $\mathcal{R}(I - U_T)^{\text{cl}} = \left\{f : \lim \frac{1}{n}\sum\limits_{k=1}^{n} f \circ T^k = 0, \quad f \in U_d(X)\right\}$ and using the previous proved result it can be verified that $\left(\frac{1}{n}\sum\limits_{k=1}^{n} f \circ T^k\right)$ is a Cauchy sequence for each $f \in L(1, d)$ (see[20, Ch. 8]).

Since each $f \in L_d(X)$ is a multiple of a function from $L(1, d)$ this assertion remains to be true for all $f \in L_d(X)$. But $L_d(X)$ is norm dense in $U_d(X)$ and this completes the first part of the proof.

In the continuous time case we can use a known trick. We fix some $x \in X$ and put $m_x E = \int\limits_0^1 \chi_E \circ T^s x \, \mathrm{d}s$. For each $f \in U_d(X)$

$$\lim \frac{1}{n}\int_0^n f \circ T^s x \, \mathrm{d}s = \lim \frac{1}{n}\sum_{k=1}^{n}\int_0^1 f \circ T^{s+k} x \, \mathrm{d}s$$
$$= \int \lim \frac{1}{n}\sum_{k=1}^{n} f \circ T^k y \int_0^1 \delta_x T^{-s}(\mathrm{d}y) \, \mathrm{d}s.$$

Applying the first part of this theorem and the bounded convergence theorem we obtain that the last expression is equal to $\int f^* y \, m_x(\mathrm{d}y) = \int\limits_0^1 f^* T^s x \, \mathrm{d}s$. By the way, m_x is non-negative and $m_x X = 1$, hence

$$\left|\int \frac{1}{n}\sum_{k=1}^{n}(f \circ T^k y - f^* y) \, m_x(\mathrm{d}y)\right| \le \left\|\frac{1}{n}\sum_{k=1}^{n} f \circ T^k - f^*\right\|.$$

Now it is clear that $\left(\frac{1}{t}\int\limits_0^t f\circ T^s\ \mathrm{d}s\right)$ is a Cauchy sequence in $U_d(X)$.

R e m a r k 2.5. For each $f\in U_d(X)$ f^* is a constant on each $I\in\mathcal{E}$ and $TI=I$. Therefore for each $x\in I$, $I\in\mathcal{E}$, $f^*\circ T^s x=f^*x$ whatever is s. So we can conclude that $\left(\frac{1}{n}\sum\limits_{k=1}^{n} f\circ T^k\right)$ and $\left(\frac{1}{t}\int\limits_0^t f\circ T^s\ \mathrm{d}s\right)$ have the same limit for each $f\in U_d(X)$.

Now we can widely generalize the results from [20, Ch. 13, §4].

EXAMPLE 2.6. *Let (S,ϱ) be a locally compact separable metric space. Let $C_0(S)$ be the Banach subspace of all functions from $C(S)$ that vanish at infinity and let (T^t) be a flow of continuous mappings from $C_0(S)$ into $C_0(S)$ with T contractive. Then to each $s\in S$ there exists a measure $\delta_s^*\in M_\sigma(S)$ such that*

$$\limsup_{x\in C_0(S)}\left|\frac{1}{t}\int\limits_0^t T^{\bar t}x(s)\ \mathrm{d}\bar t-\int\limits_S x(\bar s)\ \delta_s^*(\mathrm{d}\bar s)\right|=0\,.$$

The space S can be decomposed into a class $\mathcal{I}$ of disjoint measurable subsets so that to each $I\in\mathcal{I}$ except at most one we can relate a probability $\pi_I\in M_\delta(S)$ with the property $\int T^t x(s)\ \pi_I(\mathrm{d}s)=\int x(s)\ \pi_I(\mathrm{d}s)$ for each $x\in C_0(S)$ and $t\ge 0$ and I is the maximal set on which $\int\limits_S x(\bar s)\ \delta_s^*(\mathrm{d}\bar s)$ is constant for all $x\in C_0(S)$.

P r o o f. For each $s\in S$ $\delta_s\in M_\sigma(S)$ and by [19, Ch. 3] $M_\sigma(S)$ can be identified with X^*, the norm dual of X. As $X^*\subset U_d(X)$, where $d(x,y)=\sup\{|x(s)-y(s)|\,,\ s\in S\}$, the first assertion is now a consequence of Theorem 2.1.

The ergodic decomposition $\mathcal{I}$ of S can be defined like in Theorem 2.3 by means of the equivalence $\equiv$ defined for each $s,u\in S$ by $s\equiv u$ if and only if $\delta_s^*=\delta_u^*$. As to the measurability of $I\in\mathcal{I}$ note that if (x_p) is a countable dense subset of $C_0(S)$, then $\left\{x_p^*:\ x_p^*=\lim\frac{1}{n}\sum\limits_{k=1}^{n}T^k x_p\,,\ x_p\in(x_p)\right\}$ is dense in $\left\{x^*:\ x^*=\lim\frac{1}{n}\sum\limits_{k=1}^{n}T^k x\,,\ x\in C_0(S)\right\}$, hence $\left\{s:\ x^*(s)=\int\limits_S x(\bar s)\ \pi_I'(\mathrm{d}\bar s)\,,\ s\in S\,,\ x\in X\right\}=\bigcap\limits_p\left\{s:\ x_p^*(s)=\int\limits_S x_p(\bar s)\ \pi_I'(\mathrm{d}\bar s)\,,\ s\in S\right\}$, $\pi_I=c_I\pi_I'$ for all $I\in\mathcal{I}$.

Generally, 1 is not contained in $C_0(S)$ and therefore we cannot exclude the situation $\delta_s^* = 0$ like in Theorem 2.3. That is why for one $I \in \mathcal{I}$ π_I ought not exist.

The just presented example can be extended to metric spaces (S, ϱ) without the local compactness property using $U_d(S)$ and a flow (T^t) of mappings from $U_d(S)$ into $U_d(S)$ with T contractive, provided $U_d(S)$ is separable.

Let us suppose for a moment that Theorem 2.1 can be generalized to non-separable metric spaces. Then for each metric space (S, ϱ), $X = C(S)$ and $Tx(s) = \int_S x(\bar{s})\ P(s, \mathrm{d}\bar{s})$, where P is a Feller transition probability, $\lim \frac{1}{n} \sum_{k=1}^{n} \int_S x(\bar{s})\ P^k(s, \mathrm{d}\bar{s})$ for each $x \in X$ and $s \in S$. Since $(M_\sigma(S), w)$ is sequentially complete [19, Ch. 2, Thm. 6], there is a probability $\pi \in M_\sigma(S)$ such that $\int P(s, B)\ \pi(\mathrm{d}s) = \pi B$ for all $B \in \mathcal{B}(X)$. It is easy to give an example of S and $P(s, B)$ with Feller property for which π does not exist. This event can be explained by

THEOREM 2.7. *If (T^t) is a flow of mappings from (X, d) into (X, d) with T contractive, then for each $m \in M_\sigma(X)$ the following two conditions are equivalent:*

iii) $\left(\frac{1}{t} \int_0^t mT^{-s}\ \mathrm{d}s\right)$ *is a Cauchy sequence in* $(M_\sigma(X), w)$.

iv) $\left(\frac{1}{n} \sum_{k=1}^{n} mT^{-k}\right)$ *is relatively compact in* $(M_\sigma(X), w)$.

P r o o f. The relation iii) $\Longrightarrow$ iv) is a consequence of Remark 2.5. Let $M = \left(\frac{1}{n} \sum_{k=1}^{n} mT^{-k}\right)$. By the assumption M is relatively compact in $(M_\sigma(X), w)$. The weak topology of the duality pair $(M_\sigma(X), U_d(X))$ (denoted, say, w') is coarser than w, but it is again a Hausdorff topology, hence w and w' agree on M [4, Ch. 5, Thm. 8], and (M, w') is relatively compact in $(M_\sigma(X), U_d(X))$ with w'. Now it suffices to apply Theorem 2.1.

R e m a r k 2.8. Theorem 2.7, iv) $\Longrightarrow$ iii), can be proved without the separability of (X, d) and Theorem 2.1 provided X is complete and $m \in M_t(X)$, where $M_t(X)$ is the space of all tight measures on $\mathcal{B}(X)$ [19, Ch. 1] considered in the w topology induced from $M_\sigma(X)$. We can define on $M_t(X)$ a norm $\|\cdot\|$ by the relation $\|m\| = \sup\{|mf| : f \in L(1, d)\}$. This metric defines in $M_t(X)$ the same convergent sequences like w and the norm dual of $M_t(X)$, $\|\cdot\|$ can be identified with $C(X)$ ($m \in M_t(X)$ has a separable support and therefore it can be written as a w limit of a sequence of measures of the type

$\sum_{i=1}^{n} \alpha_i \delta_{x_i}$; therefore each continuous linear functional L on $M_t(X)$ has the property $L(m) = \int L(\delta_x)\, m(\mathrm{d}x)$).

Now we can apply the result from [20, Ch. 8, §3] to the linear operator that relates to $m \in M_t(X)$ the measure $mT^{-1} \in M_t(X)$.

We conclude this section by the question whether the fixed point theorem that holds for contractive mappings with $c_T < 1$ has any extension to $c_T \geq 1$.

PROPOSITION 2.9. *Let (X,d) be a real normed space with $d(x,y) = \|x-y\|$ for $x,y \in X$ and let T be a contractive map from (X,d) into (X,d). If to given $x \in X$ there exists $x_0 \in X$ such that for some subsequence $(n_i) \subset \mathbb{N}$*

v) $\lim \frac{1}{n_i} \sum_{k=1}^{n_i} \|T^k x - x_0\| = 0,$

then x_0 is a fixed point for T.

P r o o f. If v) holds, then $\lim \frac{1}{n} \sum_{k=1}^{n} f \circ T^k x = f x_0$ for each $f \in L_d(X)$. Consequently $f \circ T x_0 = f x_0$ for each $f \in L_d(X)$, specially for each continuous linear functional on X, and we can conclude that $T x_0 = x_0$.

Under which additional assumptions the weaker condition $\left(\frac{1}{n} \sum_{k=1}^{n} T^k x\right)$ *is relatively compact in X* implies v) remains to be an open problem and its answer can show an interesting relation between the above developed theory and the Schauder fixed point theorem [1, Ch. 3].

3. SLLN in general stochastic systems

The following theorem is based on a similar idea like Example 2.6.

THEOREM 3.1. *Let T be a homeomorphism of X onto X. Then X can be equipped by a metric d such that for each $x \in X$ and $f \in U_d(X)$ $\left(\frac{1}{n} \sum_{k=1}^{n} f \circ T^k x\right)$ is a Cauchy sequence. If T belongs to a flow (T^t), then for each $x \in X$ and $f \in U_d(X)$ $\left(\frac{1}{t} \int_0^t f \circ T^s x \,\mathrm{d}s\right)$ is a Cauchy sequence.*

P r o o f. Let $\overline{X}$, $\overline{X}_0$, $\overline{d}$ and H have the same meaning like in the proof of Theorem 2.1 and let $d(x,y) = \overline{d}(H^{-1}x, H^{-1}y)$ for all $x,y \in X$. The homeomorphism H defines an isometric isomorphism between $C(\overline{X})$ and $U_d(X)$ and the map $T_0 = H^{-1} \circ T \circ H$ defines a homeomorphism of $\overline{X}_0$ onto $\overline{X}_0$. If the homeomorphism $\overline{T}$ arises by the continuous extension of $\overline{T}_0$ from $\overline{X}_0$ to $\overline{X}$, then the relation $\overline{f} \to \overline{f} \circ \overline{T}$ defines a contractive map from $C(\overline{X})$ into $C(\overline{X})$. Since $C(\overline{X})$ is now separable in the supremum norm the $MLLN$ holds and $\left(\frac{1}{n}\sum_{k=1}^{n} \overline{f} \circ \overline{T}^k \overline{x}\right)$ is a Cauchy sequence for each $\overline{x} \in \overline{X}$ and $\overline{f} \in C(\overline{X})$. Because of the homeomorphism between $U_d(X)$ and $C(\overline{X})$ we can consider the first assertion proved. The second one can be verified by the same considerations like in the proof of Theorem 2.1. Hence d is the desired metric.

The assertion of Theorem 3.1 will be called the $SLLN$. This election prefers more the practical and historical reasons then the logical; $MLLN$ asserts more than $SLLN$.

COROLLARY 3.2. *Let X be a real separable Banach space with a biorthogonal Schauder basis $(e_k) \subset X$, $(e'_k) \subset X^*$ and T be a map from X into X that satisfies the $SLLN$.*

If for given $x \in X$ $\sup_n \|T^n\| < \infty$ and $\lim_{p\to\infty} \sup_n \left\| \sum_{k=1+p}^{\infty} \langle T^n x, e_k\rangle e'_k \right\| = 0$, then $\delta_x^* \in M_\sigma(X)$ and $\left(\frac{1}{n}\sum_{k=1}^{n} \delta_x T^{-k}\right)$ is a Cauchy sequence in $\big(M_\sigma(X), w\big)$.

If δ_x^* is concentrated in a point x_0, then x_0 is a fixed point for T.

P r o o f. Under the given assumptions $\left(\frac{1}{n}\sum_{k=1}^{n} \delta_x T^{-k}\right)$ is relatively compact in $\big(M_\sigma(X), w\big)$ [18, Ch. 1, Sec. 9]. On $M_\sigma^+(X)$ w agrees with the weak topology of the pair $\big(M_\sigma(X), U_d(X)\big)$. Integrating in the $SLLN$ we can derive the first assertion. As to the second see the proof of Proposition 2.9.

A homeomorphism T of X onto X is called *uniquely ergodic* if there exists just one measure $m^* \in M^+(X)$ such that $m^*T^{-1} = m^*$ and $m^*X = 1$.

As it is well known, if $\overline{X}$ is compact and if $\overline{T}$ is a uniquely ergodic homeomorphism of $\overline{X}$ onto $\overline{X}$, then [15, Ch. 1, Thm. 1], the $MLLN$ holds, and, identifying $\overline{X}$ and $\overline{T}$ with those in the proof of Theorem 3.1, we can derive

COROLLARY 3.3. *If T is a uniquely ergodic homeomorphism of (X, d) onto (X, d), then the $MLLN$ holds.*

Note that if T is uniquely ergodic, then $\overline{T}$ is uniquely ergodic because each $m \in M(X)$ can be written as $m = \overline{m}H$ for some $\overline{m} \in M_\sigma(\overline{X})$ and the correspondence between m and $\overline{m}$ is one to one.

$SLLN$, like $MLLN$ can be used for defining of an ergodic decomposition $\mathcal{E}$ of X. If $I \in \mathcal{E}$, then T restricted to I is uniquely ergodic, hence, we can use Corollary 3.3 for strengthening of the $SLLN$

COROLLARY 3.4. *Let T be a homeomorphism of (X,d) onto (X,d), let $\mathcal{E}$ be the ergodic decomposition of X and let T_I, $I \in \mathcal{E}$, be the restriction of T to I. Then T_I satisfies the $MLLN$.*

THEOREM 3.3. *Let (T^t) be a flow of continuous mappings from (X,d) into (X,d) that satisfies the $MLLN$. If $m \in M_\sigma^+(X)$, then these two conditions are equivalent:*

vi) $\left(\frac{1}{t}\int_0^t mT^{-s}E\,\mathrm{d}s\right)$ *is a Cauchy sequence for each* $E \in \mathcal{B}(X)$.

vii) $\left(\frac{1}{t}\int_0^t mT^{-s}\,\mathrm{d}s\right)$ *is uniformly σ-smooth.*

P r o o f. We will present only the more complicated part of the proof. The uniform σ-smoothness implies the relative compactness of $\left(\frac{1}{n}\sum_{k=1}^{n} mT^{-k}\right)$ in $\left(M_\sigma(X), w\right)$. By Theorem 2.7, that holds for each T satisfying the $MLLN$, $\left(\frac{1}{n}\sum_{k=1}^{n} mT^{-k}\right)$ is a Cauchy sequence in $\left(M_\sigma(X), w\right)$. Using the uniform σ-smoothness and [12, Thm. 1.8] we can prove that $\left(\frac{1}{n}\sum_{k=1}^{n} mT^{-k}E\right)$ is a Cauchy sequence for each $E \in \mathcal{B}(X)$. If we have to do with a continuous time flow, then it suffices to apply the just proved result to the measure $\overline{m} = \int_0^1 mT^{-s}\,\mathrm{d}s$.

THEOREM 3.4. *Let T be a homeomorphism of (X,d) onto (X,d). Then the condition* vii) *is equivalent to*

viii) *there is $m^* \in M_\sigma^+(X)$ such that*

$$\lim \operatorname{var}\left(\frac{1}{t}\int_0^t mT^{-s}\,\mathrm{d}s - m^*\right) = 0 .$$

P r o o f. If T is a homeomorphism and vii) holds for the given $m \in M_\sigma^+(X)$, then by Theorems 3.1 and 3.3 $\left(\frac{1}{n}\sum_{k=1}^{n} mT^{-k}E\right)$ is a Cauchy sequence for each $E \in \mathcal{B}(X)$ and it defines a measure $m^* \in M_\sigma^+(X)$. Let us suppose for a moment that m and m^* are defined on a σ-algebra, this can be always achieved. We will show that mT^{-n} are absolutely continuous with respect to m^*.

Since T is a homeomorphism, the sets $T^n A$ are measurable provided A is measurable and $A^* = \bigcup_{n=-\infty}^{\infty} T^n A$ is a T invariant set. Therefore $mA^* = m^* A^*$ and if $mA > 0$, then $m^* A > 0$ (otherwise $0 < mA^* \sum_{n=-\infty}^{\infty} m^* T^n A = \sum_{k=-\infty}^{\infty} m^* A = 0$ – a contradiction).

Now we can again employ the one to one property of T and show that mT^{-n} is absolutely continuous with respect to $m^* T^{-n}$ for each n and $\mathrm{d}mT^{-n}/\mathrm{d}m^* T^{-n} = f \circ T^{-n}$, where $f = \mathrm{d}m/\mathrm{d}m^*$ [16, Ch. 6, Thm. 48.7]. From the proved results we can conclude that $\lim \frac{1}{n} \sum_{k=1}^{n} \int_E f \circ T^{-k} x \, m^*(\mathrm{d}x) = \int_E m^*(\mathrm{d}x)$ which is equivalent to the weak convergence of the sequence $\left(\frac{1}{n} \sum_{k=1}^{n} f \circ T^{-k}\right)$ to 1 in $L^1(X, m^*)$. The relation $f \to f \circ T^{-k}$ defines a bounded linear operator from $L^1(X, m^*)$ into $L^1(X, m^*)$, hence by [20, Ch. 8] $\frac{1}{n} \sum_{k=1}^{n} f \circ T^{-k}$ converges to 1 in $L^1(X, m^*)$. If the flow is discrete time, then the proof follows from the relation

$$\sup_{E \in \mathcal{B}(X)} \left| \frac{1}{n} \sum_{k=1}^{n} mT^{-k} E - m^* E \right| \leq \int \left| \frac{1}{n} \sum_{k=1}^{n} f \circ T^{-k} x - 1 \right| m^*(\mathrm{d}x).$$

For continuous time flows viii) is a consequence of the last relation applied to $\overline{m} = \int_0^1 mT^{-s} \, \mathrm{d}s$.

Let S be a separable metric space and $\Omega = \prod_{n \in \mathbf{Z}} S$ be the infinite Cartesian product of S endowed with the product topology. As known, Ω is again a separable metric space and the map T, denoting now the Bernoulli shift, $(\omega_n)_{n \in \mathbf{Z}} \to (\omega_{n+1})_{n \in \mathbf{Z}}$, is a homeomorphism of Ω. Due to Theorem 3.1 there exists a metric d on Ω such that $MLLN$ holds for all $f \in U_d(\Omega)$ and $\omega \in \Omega$. The construction of d allows to state, that the restriction $(\omega_n)_{n \in \mathbf{Z}} \to (\omega_n)_{n \in \mathbf{Z}^+}$ is a uniformly continuous map, provided $\prod_{n \in \mathbf{Z}^+} S$ is again considered in the product topology. This allows to state:

LEMMA 3.5. *If S is a separable metric space, then on $\Omega = \prod_{n \in \mathbf{Z}^+} S$ we can define a metric d such that the Bernoulli shift relating to each $(\omega_n)_{n \in \mathbf{Z}^+}$ the element $(\omega_{n+1})_{n \in \mathbf{Z}^+}$ satisfies the $MLLN$.*

In the rest of this section we are going to assume that S and Ω are the spaces from Lemma 3.5. The projection q defined by the relation $(\omega_n)_{n\in\mathbb{Z}^+} \to \omega_0$ is a uniformly continuous map from Ω onto S. The mappings ξ_n, defined by the equations $\xi_n = q \circ T^n$, where T is the Bernoulli shift and $n \in \mathbb{Z}^+$, is called the *canonical stochastic process.* From this definition and arguments above we obtain:

THEOREM 3.6. *If $(\xi_n)_{n\in\mathbb{Z}^+}$ is the canonical stochastic process with values in S, then S can be equipped with a metric ϱ such that for each $x \in U_\varrho(S)$ and $\omega \in \Omega$ $\left(\frac{1}{n}\sum_{k=1}^{n} x \circ \xi_k(\omega)\right)$ is a Cauchy sequence.*

This non-probabilistic version of the $SLLN$ allows to prove the main ergodic theorems arising in probability theory.

Let $\mathcal{B}_0 \subset \mathcal{B}(\Omega)$ be the smallest algebra to which ξ_0 $(= q)$ is measurable and P be a probability measure on $\mathcal{B}(\Omega)$. Due to [16, Ch. 6] P determines on $S \times \mathcal{B}(\Omega)$ a conditional probability P_s with properties $\mathbb{E}_s(x \circ \xi_n \mid \mathcal{B}_0) = \int_\Omega x \circ \xi_n(\omega)\ P_s(\mathrm{d}_\omega)$ a.s. $P\xi_0^{-1}$ for each $x \in C(S)$, and $\int_B P_s E\, P\xi_0^{-1}(\mathrm{d}s) = PE \cap \xi_0^{-1} B$ for each $B \in \mathcal{B}(S)$, and $E \in \mathcal{B}(\Omega)$. In practice usually only P_s is known and there is a question whether there exists a measure $\pi \in M_\sigma(S)$ such that if we put $P\xi_0^{-1} = \pi$, then $P = \int_S P_s \pi(\mathrm{d}s)$ becomes an ergodic invariant measure with respect to the Bernoulli shift. If we put for $B \in \mathcal{B}(S)$ $P^n(s, B) = \int_\Omega \chi_B \circ \xi_n(\omega)\ P_s(\mathrm{d}\omega)$, then $P^n(s, \cdot) \in M_\sigma(S)$ for each $s \in S$.

THEOREM 3.7. *Let (ξ_n) be a canonical stochastic process defined on $(\Omega, \mathcal{B}(\Omega), P)$ and with values in S. Then S can be equipped with a metric ϱ so that S admits a decomposition into a class $\mathcal{I}$ of disjoint sets with properties:*

ix) *to each $I \in \mathcal{I}$ we can relate a measure $\pi_I \in M(S)$ so that for each $x \in U_\varrho(S)$ ($x \in C(S)$ as long as $\pi_I \in M_\sigma(S)$) and $s \in I$*

$$\lim \frac{1}{n}\sum_{k=1}^{n} \int_S x(\overline{s})\ P^k(s, \mathrm{d}\,\overline{s}) = \int_S x(\overline{s})\ \pi_I(\mathrm{d}\,\overline{s}),$$

xi) *if for given $s \in S$ P_s is ergodic, i.e. $P_s J = 0$ or $P_s J = 1$ for each shift invariant set, then for each $x \in U_\varrho(S)$ ($x \in C(S)$ if $\pi_I \in M_\sigma(S)$)*

$$\lim \frac{1}{n}\sum_{k=1}^{n} x \circ \xi_k = \int_S x(\overline{s})\ \pi_I(\mathrm{d}\,\overline{s}) \qquad a.s.\ \ P_s.$$

Proof. Assertion ix) is an immediate consequence of Theorem 3.6 and of the Fubini theorem. The decomposition $\mathcal{I}$ of S can be defined in the same manner like in Example 2.6.

If P_s is ergodic, then the function $x^*(\omega) = \lim \frac{1}{n} \sum_{k=1}^{n} x \circ \xi_k(\omega)$, $x \in U_\varrho(S)$, is shift invariant, hence constant a.s. P_s and after integrating $\int_S x(s)\ \pi_I(\mathrm{d}s) = \int_\Omega x^*(\omega)\ P_s(\mathrm{d}\omega) = x^*(\omega)$ a.s. P_s.

Even if π_I is the measure computed in x), the measure $P = \int_S P_s\ \pi_I(\mathrm{d}s)$ ought not to be shift invariant without additional assumptions about P_s. But if (ξ_n) has with respect to P the Markov property, then it becomes an ergodic invariant measure. This can be verified using

PROPOSITION 3.8. *If P_s belongs to a Markov process, then it is ergodic with respect to the Bernoulli shift.*

Proof. Let $J \in \mathcal{B}(\Omega)$ be shift invariant, i.e. $TJ = J$. If we put $I = qJ$, then $\chi_I(\xi_1) = \chi_{qTJ}\, q \circ T = \chi_{qJ} \circ q = \chi_I(\xi_0)$ and consequently $P(s, I) = \chi_I(s)$. P_s is uniquely determined by its values on cylindrical sets $E = \prod_{n \in \mathbb{Z}^+} B_n$, where $(B_n) \subset \mathcal{B}(S)$ and $B_n \neq S$ for at most first $n_0 + 1$ sets, n_0 finite. If

$$P_s E = \chi_{B_0}(s) \int_{B_1} P(s, \mathrm{d}s_1) \int_{B_2} P(s_1, \mathrm{d}s_2) \cdots \int_{B_{n_0}} P(s_{n_0-1}, \mathrm{d}s_{n_0}),$$

then using the σ-smoothness of P_s and considering J instead E we can easily compute that $P_s J = \chi_I(s)$ and this completes the proof.

An attempt to prove Theorem 3.7 for Markov chains with a locally compact phase space and a unique invariant measure can be found in [10].

THEOREM 3.9. *Let us consider the objects (S, ϱ), $(\Omega, \mathcal{B}(\Omega), P)$, (ξ_n), $\mathcal{I}$ and $(\pi_I)_{I \in \mathcal{I}}$ from Theorem 3.7. Let us suppose that $F\colon \Omega \to \Omega$ is a measurable map, $\eta_n = \xi_n \circ F$ for $n \in \mathbb{Z}^+$, $\pi_I \in M_\sigma(S)$ for any $I \in \mathcal{I}$, P_s is ergodic for any $s \in I$ and $P_s F^{-1}$ is absolutely continuous with respect to P_s. Then there is $\pi'_s \in M_\sigma(S)$ such that for each $x \in C(S)$*

xii) $\lim \frac{1}{n} \sum_{k=1}^{n} x \circ \eta_k = \int_S x(\bar{s})\ \pi'_s(\mathrm{d}\,\bar{s})$ *a.s.* P_s.

Proof. Due to Theorem 3.7 xi) we can find a measurable set $\Omega_s \subset \Omega$ such that $P_s \Omega_s = 1$ and for each $x \in C(S)$ and $\omega \in \Omega_s$ $\left(\frac{1}{n} \sum_{k=1}^{n} x \circ \xi_k(\omega)\right)$ is a

Cauchy sequence. But then $\left(\frac{1}{n}\sum_{k=1}^{n} x\circ\eta_k(\omega)\right)$ is a Cauchy sequence for each $x\in C(S)$ and $\omega\in F^{-1}\Omega_s$. Since P_sF^{-1} is absolutely continuous with respect to P_s, $\left(\frac{1}{n}\sum_{k=1}^{n} x\circ\eta_k(\omega)\right)$ is a Cauchy sequence for all $x\in C(S)$ and almost all $\omega\in\Omega$ ($= F^{-1}\Omega$ a.s. P_s). The assertion is now a consequence of ergodicity of P_s, of the fact that the limit is defined for all $x\in C(S)$ and of the sequential completness of $\left(M_\sigma(S), w\right)$.

COROLLARY 3.10. *Let (ξ_n) be a canonical Markov process on $\left(\Omega,\mathcal{B}(\Omega),P\right)$ with values in (S,ϱ). Let $G\colon S\times S\to S$ be $\left(\mathcal{B}(S\times S),\mathcal{B}(S)\right)$ measurable and let $\eta=(\eta_n)$, where*

$$\eta_{n+1} = G(\eta_n,\xi_{n+1}), \qquad n>0.$$

If the process (ξ_n) is ergodic with a unique invariant measure $\pi\in M_\sigma^+(S)$ and if $P_s\eta^{-1}$ is absolutely continuous with respect to P_s for each $s\in S$, then the process (η_n) is ergodic with a unique invariant measure $\pi'\in M_\sigma^+(S)$ such that xii) *holds for all $x\in C(S)$.*

Proof. By Proposition 3.8 we can apply Theorem 3.9 to the mapping $F\colon(\omega_0,\omega_1,\omega_2,\dots)\to\left(\eta_0,\ G(\eta_0,\omega_1),\ G\big(G(\eta_0,\omega_1),\omega_2\big),\dots\right)$, where the relation $\omega_0\to\eta_0$ can be defined arbitrarily, but measurable.

The previous method cannot be extended immediately to the continuous time case because of the heavy measurability problems arising as long as $\Omega=\prod_{t\in\langle 0,\infty)} S$ is considered in the product topology. However, if we restrict our attention to $\Omega_c=\left\{\omega:\ \omega\in\Omega \text{ and } \omega \text{ is continuous in } t \text{ on } \langle 0,\infty)\right\}$, then the Bernoulli shift (T^t) maps Ω_c onto Ω_c and the mappings T^t restricted to Ω_c are continuous in the product topology induced on Ω_c from Ω. So (T^t) is a flow of continuous mappings from Ω_c onto Ω_c and T agrees with the shift discussed previously on $\prod_{n\in\mathbb{Z}^+} S$.

Let (ξ_t) be the canonical process defined on Ω by $\xi_t(\omega)=q\circ T^t\omega$, $\omega\in\Omega$, $t\ge 0$. Then, using the same trick like in the proof of Theorem 2.1 we can show that for each $x\in U_\varrho(S)$ and $\omega\in\Omega_c$ $\left(\frac{1}{t}\int_0^t x\circ\xi_s(\omega)\,\mathrm{d}s\right)$ is a Cauchy sequence, because $\left(\frac{1}{n}\sum_{k=1}^{n} x\circ\xi_k(\omega)\right)$ is a Cauchy sequence. Considering a probability P on $\mathcal{B}(\Omega)$ with support contained in Ω_c we can easily derive the continuous time analogy of Theorem 3.7.

Comparison of the above derived results with the known ones leads to the conclusion that for Markov chains we have obtained a new method how to derive some assertions from [13, 17] (see there the historical remarks). Theorems 2.7, 3.3 and 3.4 are analogies to the non-topological approach in [3]. Remarkable remains the generality of the results.

Supposing that the assumption T *is homeomorphic* can be replaced by the weaker hypothesis T *is one to one*, our theory could be extended to automorphisms admitting only σ-finite invariant measures (cf [3]), but this problem we leave open.

REFERENCES

[1] CONWAY, J. B.: *A Course in Functional Analysis*, Springer-Verlag, New York-Berlin-Heidelberg-Tokyo, 1985.

[2] DUDLEY, R. M.: *Convergence of Baire measures*, Studia Math. **27** (1966), 251–268.

[3] FRIEDMAN, N. A.: *Introduction to Ergodic Theory*, Van Nostrand, New York-Cincinnati-Toronto-London-Melbourne, 1970.

[4] KELLEY, J. L.: *General Topology*, D. Van Nostrand, New York, 1955.

[5] KOLMOGOROV, A. N.—FOMIN, S. V.: *Introduction to Theory of Functions and Functional Analysis.* (Russian), Nauka, Moscow, 1981.

[6] KORNFEL'D, I. P.—SINAI, J. G.—FOMIN, S. V.: *Ergodic Theory.* (Russian), Nauka, Moscow, 1980.

[7] KOROLJUK, V. S.—TURBIN, A. F: *Mathematical Elements of Phase Amplification in Complex Systems.* (Russian), Naukova Dumka, Kiev, 1978.

[8] KOVALENKO, I. N.—KUZNECOV, N. JU.—SHURENKOV, V. M.: *Random Processes.* (Russian), Naukova Dumka, Kiev, 1983.

[9] LOTZ, H. P.: *Positive linear operators on L^p and the Doeblin condition.* In: Aspects of Positivity in Functional Analysis. Elsevier Science Publishers B.V., North Holand, 1986, pp. 137–156.

[10] MEYN, S. P.: *Ergodic theorems for discrete time stochastic systems using a stochastic Lyapunov function*, SIAM J. Control Optim. **27** (1989), 1409–1439.

[11] MOHAPL, J.: *The Radon measures as functionals on Lipschitz functions*, Czechoslovak Math. J. **41** (1991), 446–453.

[12] MOHAPL, J.: *On weakly convergent nets in spaces of non-negative measures*, Czechoslovak Math. J. **40** (1990), 408–421.

[13] NUMMELIN, E.: *General Irreducible Markov Chains and Non Negative Operators*, Cambridge University Press, Cambridge, 1984.

[14] PACHL, J. K.: *Measures as functionals on uniformly continuous functions*, Pacific J. Math. **82** (1979), 515–521.

[15] PARRY, W.: *Topics in Ergodic Theory*, Cambridge University Press, Cambridge, 1980.

[16] PARTHASARATHY, K. R.: *Introduction to Probability and Measure*, Springer-Verlag, New York-Heidelberg-Berlin, 1978.

[17] SHURENKOV, V. M.: *Ergodic Markov Processes.* (Russian), Nauka, Moscow, 1989.

[18] VACHANIJA, N. N.—TARIELADZE, V. I.—ČOBANJAN, S. A.: *Probability Distributions in Banach Spaces.* (Russian), Nauka, Moscow, 1985.

[19] VARADARJAN, V. S.: *Measures on topological spaces.* (Russian), Mat. Sb. **55** (1961), 35–100.

[20] YOSIDA, K.: *Functional Analysis*, Springer-Verlag, Berlin-Göttingen-Heidelberg, 1965.

[21] ŽDANOK, T. A.: *Fixed point theorem for measurable field of operators with an application to random differential equation.* In: Fifth Japan-USSR Symposium Proceedengs 1986, Springer-Verlag, New York-Heidelberg-Berlin, 1988.

Received September 27, 1991

Revised August 20, 1992

Kosmákova 24
785 01 Šternberg
Czech Republic