ISRAEL JOURNAL OF MATHEMATICS, Vol. 65, No. 1, 1989

ON THE EXISTENCE OF INVARIANT PROBABILITY MEASURES

BY
WOLFGANG ADAMSKI
Mathematisches Institut der Universität München, Theresienstr. 39, D-8000 München 2, FRG

ABSTRACT
Let $(X, \mathscr{A})$ be a measurable space and $T: X \to X$ a measurable mapping. Consider a family $\mathscr{M}$ of probability measures on $\mathscr{A}$ which satisfies certain closure conditions. If $\mathscr{A}_0 \subset \mathscr{A}$ is a convergence class for $\mathscr{M}$ such that, for every $A \in \mathscr{A}_0$, the sequence $((1/n) \sum_{i=0}^{n-1} 1_A \circ T^i)$ converges in distribution (with respect to some probability measure $\nu \in \mathscr{M}$), then there exists a T-invariant element in $\mathscr{M}$. In particular, for the special case of a topological space X and a continuous mapping T, sufficient conditions for the existence of T-invariant Borel probability measures with additional regularity properties are obtained.

Introduction

Let $(X, \mathscr{A})$ be a measurable space and $T: X \to X$ a measurable mapping. If there is a T-invariant probability measure μ on $\mathscr{A}$, then, by Birkhoff's ergodic theorem, the averages $(1/n) \cdot \sum_{i=0}^{n-1} f \circ T^i$ converge μ-a.e. for every $f \in \mathscr{L}_1(X, \mathscr{A}, \mu)$. Our paper is concerned, among others, with the converse of this statement.

Consider a family $\mathscr{M}$ of probability measures on $\mathscr{A}$ which satisfies certain closure conditions. If $\mathscr{A}_0 \subset \mathscr{A}$ is a convergence class for $\mathscr{M}$ (cf. [3]) such that, for every $A \in \mathscr{A}_0$, the sequence $((1/n) \cdot \sum_{i=0}^{n-1} 1_A \circ T^i)$ converges in distribution (with respect to some probability measure $\nu \in \mathscr{M}$), then there exists a T-invariant element in $\mathscr{M}$. In particular, for the special case of a topological space X and a continuous mapping T, we obtain sufficient conditions for the existence of T-invariant Borel probability measures with additional regularity properties.

Received December 22, 1987

If $\mathscr{K} \subset \mathscr{A}$ is a semicompact δ-lattice satisfying some further conditions, then another central theorem of this paper states that a T-invariant $\mathscr{K}$-regular probability measure on $\mathscr{A}$ exists provided that, for some $K_0 \in \mathscr{K}$ and $x_0 \in X$,

$$\limsup_n \frac{1}{n} \cdot \sum_{i=0}^{n-1} 1_{K_0} \circ T^i(x_0) > 0$$

holds. Especially, this result enables us to extend a theorem of Oxtoby and Ulam ([10]) on the existence of invariant Radon probability measures from Polish spaces to arbitrary Hausdorff spaces.

2. Definitions and auxiliary results

$\mathbf{N}$ [$\mathbf{N}_0$] denotes the set of positive [nonnegative] integers. The set $\mathbf{R}$ of real numbers is always assumed to be equipped with the Euclidean topology.

Let X be an arbitrary set and let $\mathscr{P}(X)$ be the power set of X. 1_Q denotes the indicator function of a set $Q \in \mathscr{P}(X)$. For $f \in \mathbf{R}^X$ we put $\| f \| := \sup\{ |f(x)| : x \in X\}$. id_X denotes the identity map on X.

Let $\mathscr{K}$ be a subset of $\mathscr{P}(X)$. Then $\mathscr{K}_\delta$ denotes the family of all countable intersections of members of $\mathscr{K}$, while $\sigma(\mathscr{K})$ stands for the σ-algebra generated by $\mathscr{K}$. Furthermore,

$$\mathscr{F}(\mathscr{K}) := \{F \in \mathscr{P}(X) : F \cap K \in \mathscr{K} \text{ for all } K \in \mathscr{K}\}$$

denotes the collection of all "local $\mathscr{K}$-sets". $\mathscr{K}$ is said to be semicompact if every countable subfamily of $\mathscr{K}$ having the finite intersection property has nonvoid intersection. $\mathscr{K}$ is called a lattice [δ-lattice] if $\phi \in \mathscr{K}$ and $\mathscr{K}$ is closed under finite unions and finite [countable] intersections.

By a measure we always understand a $[0, \infty)$-valued σ-additive function defined on a σ-algebra. If μ is a measure on $\mathscr{A}$ and $\mathscr{K}$ is a subset of $\mathscr{A}$, then μ is said to be $\mathscr{K}$-regular if

$$\mu(A) = \sup\{\mu(K) : K \in \mathscr{K}, K \subset A\} \qquad \text{for all } A \in \mathscr{A}.$$

If μ and ν are measures on $\mathscr{A}$, then we write $\mu \ll \nu$ if, for any $A \in \mathscr{A}$, $\nu(A) = 0$ implies $\mu(A) = 0$. We write $\mu \sim \nu$ if $\mu \ll \nu$ and $\nu \ll \mu$.

Let $(X, \mathscr{A}, \mu)$ be a measure space. If $(Y, \mathscr{B})$ is a measurable space and $f: X \to Y$ is $\mathscr{A}, \mathscr{B}$-measurable, then $f(\mu)$ denotes the image measure of μ under the mapping f, i.e. $f(\mu)(B) := \mu(f^{-1}B)$ for $B \in \mathscr{B}$. If $T: X \to X$ is $\mathscr{A}, \mathscr{A}$-measurable, then μ is said to be *T-invariant* if we have $\mu = T(\mu)$.

If X is a topological space, then we denote by $\mathscr{C}(X)$ [$\mathscr{C}^b(X)$] the family of all continuous [and bounded] real-valued functions on X. A cozero-set in X is a set of the form $\{f \neq 0\}$ with $f \in \mathscr{C}(X)$. We write $\mathscr{F}(X)$, $\mathscr{G}(X)$, $\mathscr{K}(X)$, $\mathscr{U}(X)$ for the collection of all closed, open, compact, cozero-sets in X, respectively. We will consider also the class

$$\mathscr{G}_r(X) := \{G \in \mathscr{G}(X) : G = (\bar{G})^0\} = \{F^0 : F \in \mathscr{F}(X)\}$$

of the so-called regular open sets which is, in general, a proper subclass of $\mathscr{G}(X)$. Here $\bar{A}$ $[A^0]$ denotes the closure [interior] of a subset A of X.

Finally, $\mathscr{B}_0(X) := \sigma(\mathscr{U}(X))$ $[\mathscr{B}(X) := \sigma(\mathscr{G}(X))]$ denotes the Baire [Borel] σ-field in X. A measure defined on $\mathscr{B}_0(X)$ $[\mathscr{B}(X)]$ is called a Baire [Borel] measure on X. A Borel measure μ on X is said to be τ-smooth if $\lim_\alpha \mu(F_\alpha) = \mu(F)$ for every net (F_α) in $\mathscr{F}(X)$ with $F_\alpha \downarrow F$. $\mathscr{F}(X)$-regular Borel measures on X are simply called regular. On the other hand, a $\mathscr{K}(X)$-regular Borel measure on a Hausdorff space X is called a Radon measure.

In the sequel let X be an arbitrary set and $\mathscr{V} \subset \mathbf{R}^X$ a vector lattice (with respect to pointwise operations). We write $\sigma(\mathscr{V})$ for the smallest σ-algebra in X making all functions $f \in \mathscr{V}$ measurable. $\Gamma(\mathscr{V})$ denotes the subfamily of $\mathbf{R}^{\mathscr{V}}$ consisting of all nonnegative linear functionals. For a sequence (f_n) in $\mathscr{V}$ we write $f_n \downarrow f$ if (f_n) is decreasing and converges pointwise to $f \in \mathbf{R}^X$. $\Phi \in \Gamma(\mathscr{V})$ is said to be σ-smooth if $f_n \downarrow 0$ implies $\lim_n \Phi(f_n) = 0$. Define $\Gamma_\sigma(\mathscr{V}) := \{\Phi \in \Gamma(\mathscr{V}) : \Phi \text{ is } \sigma\text{-smooth}\}$. We call $\mathscr{V}$ a *Daniell lattice* ([2]) if $\Gamma(\mathscr{V}) = \Gamma_\sigma(\mathscr{V})$.

If X is a topological space, then a classical theorem of Alexandroff (see [4], Theorem 19.3 or [12], Theorem II.19) states that the limit of a pointwise convergent sequence in $\Gamma_\sigma(\mathscr{C}^b(X))$ is again an element of $\Gamma_\sigma(\mathscr{C}^b(X))$. In the following we need this property for arbitrary vector lattices of real-valued functions:

We say that the vector lattice $\mathscr{V} \subset \mathbf{R}^X$ has the *Alexandroff property* if, for every sequence $(\Phi_n) \subset \Gamma_\sigma(\mathscr{V})$ such that $\Phi(f) := \lim_n \Phi_n(f)$ exists in $\mathbf{R}$ for all $f \in \mathscr{V}$, we have $\Phi \in \Gamma_\sigma(\mathscr{V})$.

It is trivial that every Daniell lattice has the Alexandroff property. A deeper result is given in

2.1. Proposition. *$\mathscr{V}$ has the Alexandroff property provided that $\mathscr{V}$ satisfies the following condition*:

(2.1) *If* $f_1, f_2, \ldots \in \mathscr{V}_+$ *and* $\sum_{k \in \mathbf{N}} f_k \in \mathscr{V}$, *then* $\sum_{k \in A} f_k \in \mathscr{V}$ *for all* $A \in \mathscr{P}(\mathbf{N})$.

Proof. Let $(\phi_n) \subset \Gamma_\sigma(\mathscr{V})$ be such that $\Phi(f) := \lim_n \Phi_n(f)$ exists in $\mathbf{R}$ for all $f \in \mathscr{V}$. Obviously we have $\Phi \in \Gamma(\mathscr{V})$. To prove the σ-smoothness of Φ it suffices to show $\Phi(\Sigma_{k \in \mathbf{N}} f_k) = \Sigma_{k \in \mathbf{N}} \Phi(f_k)$ for every sequence (f_k) in $\mathscr{V}_+$ with $\Sigma_{k \in \mathbf{N}} f_k \in \mathscr{V}$.

Let such a sequence (f_k) be given. In view of (2.1), we can define

$$\mu_n(A) := \Phi_n\left(\sum_{k \in A} f_k\right) \quad \text{and} \quad \mu(A) := \Phi\left(\sum_{k \in A} f_k\right) \qquad \text{for } A \in \mathscr{P}(\mathbf{N}) \text{ and } n \in \mathbf{N}.$$

(μ_n) is a sequence of measures on $\mathscr{P}(\mathbf{N})$ such that $\lim_n \mu_n(A) = \mu(A) \in \mathbf{R}$ for all $A \subset \mathbf{N}$. By Nikodym's theorem (see [8], III. 7.4), μ is also a measure on $\mathscr{P}(\mathbf{N})$. This implies

$$\sum_{k=1}^{n} \Phi(f_k) = \Phi\left(\sum_{k=1}^{n} f_k\right) = \mu(\{1, \ldots, n\}) \to \mu(\mathbf{N}) = \Phi\left(\sum_{k \in \mathbf{N}} f_k\right),$$

i.e. $\Phi(\Sigma_{k \in \mathbf{N}} f_k) = \Sigma_{k \in \mathbf{N}} \Phi(f_k)$. □

If $(X, \mathscr{A}, \mu)$ is a measure space and $p \in [1, \infty)$, then it is an immediate consequence of 2.1 that $\mathscr{V} := \mathscr{L}_p(X, \mathscr{A}, \mu)$ has the Alexandroff property.

We will now give another example of a vector lattice with the Alexandroff property. For this purpose consider a δ-lattice $\mathscr{L}$ of subsets of X with $X \in \mathscr{L}$. A real-valued function f on X is said to be *$\mathscr{L}$-continuous* if $f^{-1}F \in \mathscr{L}$ for all closed subsets F of $\mathbf{R}$. Note that a function $f \in \mathbf{R}^X$ is $\mathscr{L}$-continuous iff the sets $\{f \geqq t\}$ and $\{f \leqq t\}$ belong to $\mathscr{L}$ for all $t \in \mathbf{R}$. Define

$$\mathscr{C}(\mathscr{L}) := \{f \in \mathbf{R}^X : f \text{ is } \mathscr{L}\text{-continuous}\}$$

and

$$\mathscr{C}^b(\mathscr{L}) := \{f \in \mathscr{C}(\mathscr{L}) : f \text{ is bounded}\}.$$

Then $\mathscr{C}(\mathscr{L})$ and $\mathscr{C}^b(\mathscr{L})$ are vector lattices containing the constants.

2.2. Examples. (a) Let $(X, \mathscr{A})$ be a measurable space. Then $\mathscr{C}(\mathscr{A})$ $[\mathscr{C}^b(\mathscr{A})]$ is the family of all [bounded] $\mathscr{A}$-measurable real-valued functions on X.

(b) If X is a topological space, then we have $\mathscr{C}(X) = \mathscr{C}(\mathscr{F}(X))$ and $\mathscr{C}^b(X) = \mathscr{C}^b(\mathscr{F}(X))$.

2.3. Proposition. *The vector lattices $\mathscr{C}(\mathscr{L})$ and $\mathscr{C}^b(\mathscr{L})$ have the Alexandroff property.*

Proof. Let $\mathscr{V} \in \{\mathscr{C}(\mathscr{L}), \mathscr{C}^b(\mathscr{L})\}$. We show that $\mathscr{V}$ satisfies (2.1). Let $(f_k) \subset \mathscr{V}_+$ with $f := \Sigma_{k \in \mathbf{N}} f_k \in \mathscr{V}$ be given. We must prove $\Sigma_{k \in A} f_k \in \mathscr{V}$ for every infinite subset A of $\mathbf{N}$. If $\mathbf{N} - A$ is finite, then $\Sigma_{k \in A} f_k = f - \Sigma_{k \in \mathbf{N} - A} f_k \in$

$\mathscr{V}$. Thus we assume that A and $\mathbf{N}-A$ are infinite. Let $A=\{n_1, n_2, \ldots\}$ with $n_1<n_2<\cdots$ and $\mathbf{N}-A=\{r_1, r_2, \ldots\}$ with $r_1<r_2<\cdots$. For any $t\in\mathbf{R}$, we obtain

$$\left\{\sum_{k\in A} f_k \leqq t\right\} = \bigcap_{m\in\mathbf{N}} \left\{\sum_{k=1}^{m} f_{n_k} \leqq t\right\} \in \mathscr{L}$$

and

$$\left\{\sum_{k\in A} f_k \geqq t\right\} = \left\{\sum_{k\in\mathbf{N}-A} f_k - f \leqq -t\right\} = \bigcap_{m\in\mathbf{N}} \left\{\sum_{k=1}^{m} f_{r_k} - f \leqq -t\right\} \in \mathscr{L}.$$

Hence $\Sigma_{k\in A} f_k$ is $\mathscr{L}$-continuous. □

From 2.2 and 2.3 we deduce

2.4. Corollary. (a) *If $(X, \mathscr{A})$ is a measurable space, then the vector lattice of all [bounded] $\mathscr{A}$-measurable real-valued functions on X has the Alexandroff property.*

(b) *If X is a topological space, then the vector lattice of all [bounded] continuous real-valued functions on X has the Alexandroff property.*

2.5. Proposition. *Let $\mathscr{V}_0$ and $\mathscr{V}$ be vector lattices of bounded real-valued functions on X satisfying the following three conditions:*

(i) $1\in\mathscr{V}_0\subset\mathscr{V}$;

(ii) *$\mathscr{V}_0$ is dense in $\mathscr{V}$ (with respect to the topology of uniform convergence);*

(iii) *$\mathscr{V}$ has the Alexandroff property.*

Then $\mathscr{V}_0$ has the Alexandroff property, too.

Proof. Let $(\Phi_n)\subset\Gamma_\sigma(\mathscr{V}_0)$ be a sequence such that $\lim_n \Phi_n(f)$ exists in $\mathbf{R}$ for all $f\in\mathscr{V}_0$. By the Daniell–Stone theorem (see [5], Satz 39.4), there is, for any $n\in\mathbf{N}$, a measure μ_n on $\sigma(\mathscr{V}_0)$ such that $\Phi_n(f)=\int f d\mu_n$ for $f\in\mathscr{V}_0$.

Next we show that

$$\lim_n \int f d\mu_n \quad \text{exists in } \mathbf{R} \text{ for all } f\in\mathscr{V}. \tag{2.2}$$

Note that $\sigma(\mathscr{V})=\sigma(\mathscr{V}_0)$ by (ii). Let $f\in\mathscr{V}$ and $\varepsilon>0$ be given. Put

$$\varepsilon' := \frac{\varepsilon}{3(a+1)} \qquad \text{where } a := \lim_n \Phi_n(1).$$

Choose an $f_0\in\mathscr{V}_0$ such that $\| f-f_0 \| <\varepsilon'$. Then one can find an index n_0 such that

$$|\Phi_m(f_0) - \Phi_n(f_0)| < \frac{\varepsilon}{3} \quad \text{and} \quad \Phi_m(1) \leqq a + 1 \qquad \text{for all } m, n \geqq n_0.$$

This implies

$$\begin{aligned} \left| \int f d\mu_m - \int f d\mu_n \right| &\leqq \int |f - f_0| \, d\mu_m + |\Phi_m(f_0) - \Phi_n(f_0)| + \int |f_0 - f| \, d\mu_n \\ &\leqq \varepsilon' \Phi_m(1) + \frac{\varepsilon}{3} + \varepsilon' \Phi_n(1) \\ &\leqq \varepsilon \qquad \text{for all } m, n \geqq n_0. \end{aligned}$$

Hence (2.2) holds.

Since $\mathscr{V}$ has the Alexandroff property, we infer from (2.2) that $\Phi(f) := \lim_n \int f d\mu_n$, $f \in \mathscr{V}$, defines an element Φ of $\Gamma_\sigma(\mathscr{V})$. In particular, the restriction of Φ onto $\mathscr{V}_0$, i.e. the mapping $f \in \mathscr{V}_0 \rightarrow \lim_n \Phi_n(f)$, is an element of $\Gamma_\sigma(\mathscr{V}_0)$. □

2.6. Corollary. *If $(X, \mathscr{A})$ is a measurable space, then the vector lattice $\mathscr{V}$ of all $\mathscr{A}$-step functions has the Alexandroff property.*

Proof. $\mathscr{V}$ is dense in $\mathscr{C}^b(\mathscr{A})$. By 2.4(a), $\mathscr{C}^b(\mathscr{A})$ has the Alexandroff property. Thus our claim follows from 2.5. □

Remark. If $\mathscr{A}$ is an algebra, then the vector lattice of all $\mathscr{A}$-step functions does not have the Alexandroff property, in general (see 3.5).

3. Main results

A sequence (f_n) of real-valued random variables defined on a probability space $(X, \mathscr{A}, \nu)$ is said to converge in distribution (with respect to ν) if the corresponding sequence $(f_n(\nu))$ of image measures converges weakly to some probability measure ρ defined on $\mathscr{B}(\mathbf{R})$ (i.e. $\lim_n \int g df_n(\nu) = \int g d\rho$ for all $g \in \mathscr{C}^b(\mathbf{R})$). The following result can be proved by the same method as the first part of Theorem 5.4 in [6].

3.1. Lemma. *Let $(X, \mathscr{A}, \nu)$ be a probability space. If (f_n) is a uniformly integrable sequence of real-valued random variables which converges in distribution, then we have*

$$\lim_n \int f_n d\nu = \int \mathrm{id}_{\mathbf{R}} \, d\rho,$$

where ρ denotes the weak limit of the sequence $(f_n(\nu))$.

In the following we consider an arbitrary nonvoid set X and a fixed mapping $T: X \to X$. Then we define

$$A_n f := \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i \qquad \text{for every } n \in \mathbf{N} \text{ and every bounded } f \in \mathbf{R}^X.$$

As $(A_n f) \circ T = A_n(f \circ T)$, we simply write $A_n f \circ T$ for this function. Finally, for $Q \in \mathscr{P}(X)$, we write $A_n Q$ instead of $A_n 1_Q$.

3.2. LEMMA. *Let $(X, \mathscr{A}, \nu)$ be a probability space and $T: X \to X$ an $\mathscr{A}, \mathscr{A}$-measurable mapping. Furthermore, let $\mathscr{V} \subset \mathbf{R}^X$ be a family of bounded $\mathscr{A}$-measurable functions such that, for every $f \in \mathscr{V}$, the sequence $(A_n f)$ converges in distribution. Then we have*

$$\lim_n \int A_n f d\nu = \lim_n \int A_n f \circ T d\nu = \int \mathrm{id}_{\mathbf{R}}\, d\rho_f \qquad \textit{for every } f \in \mathscr{V}.$$

Here ρ_f denotes the weak limit of the sequence $(A_n f(\nu))$.

PROOF. For any $f \in \mathscr{C}^b(\mathscr{A})$, the sequence $(A_n f)$ is uniformly integrable, since

$$|A_n f| \leqq \|f\| \in \mathscr{L}_1(X, \mathscr{A}, \nu) \qquad \text{for all } n \in \mathbf{N}.$$

Thus 3.1 implies

$$\lim_n \int A_n f d\nu = \int \mathrm{id}_{\mathbf{R}}\, d\rho_f \qquad \text{for } f \in \mathscr{V}.$$

Furthermore, as the difference $A_n f \circ T - A_n f$ converges in probability to 0, $(A_n f \circ T(\nu))$ is also weakly convergent to ρ_f by Slutsky's theorem (see [7], Theorem 8.1.1). Therefore

$$\lim_n \int A_n f \circ T d\nu = \int \mathrm{id}_{\mathbf{R}}\, d\rho_f \qquad \text{for } f \in \mathscr{V}$$

holds again by 3.1. □

By means of 3.2, we can now prove the following basic result.

3.3. THEOREM. *Let $\mathscr{V} \subset \mathbf{R}^X$ be a vector lattice of bounded functions such that $1 \in \mathscr{V}$. Assume that $\mathscr{V}$ has the Alexandroff property. If $T: X \to X$ is a mapping such that $f \circ T \in \mathscr{V}$ for $f \in \mathscr{V}$, then the following three statements are equivalent:*

(1) *There is a probability measure* ν *on* $\sigma(\mathscr{V})$ *such that, for every* $f \in \mathscr{V}$, *the sequence* $(A_n f)$ *converges* ν-*a.e.*
(2) *There is a probability measure* ν *on* $\sigma(\mathscr{V})$ *such that, for every* $f \in \mathscr{V}$, *the sequence* $(A_n f)$ *converges in distribution.*
(3) *There is a* T-*invariant probability measure on* $\sigma(\mathscr{V})$.

PROOF. (1)$\Rightarrow$(2) is obvious.

(2)$\Rightarrow$(3). By assumption, there is a probability measure ν on $\sigma(\mathscr{V})$ such that, for every $f \in \mathscr{V}$, the sequence $(A_n f(\nu))$ converges weakly to a probability measure ρ_f on $\mathscr{B}(\mathbf{R})$. Since T is $\sigma(\mathscr{V})$, $\sigma(\mathscr{V})$-measurable, we infer from 3.2

$$\lim_n \int A_n f d\nu = \lim_n \int A_n f \circ T d\nu = \int \mathrm{id}_{\mathbf{R}}\, d\rho_f \qquad \text{for all } f \in \mathscr{V}.$$

We define

$$\Phi_n(f) := \int A_n f d\nu \quad \text{and} \quad \Phi(f) := \int \mathrm{id}_{\mathbf{R}}\, d\rho_f \qquad \text{for } f \in \mathscr{V} \quad \text{and} \quad n \in \mathbf{N}.$$

Then $(\Phi_n) \subset \Gamma_\sigma(\mathscr{V})$ and

$$\Phi(f) = \lim_n \Phi_n(f) = \lim_n \Phi_n(f \circ T) = \Phi(f \circ T) \qquad \text{for all } f \in \mathscr{V}.$$

Since $\mathscr{V}$ has the Alexandroff property, we conclude $\Phi \in \Gamma_\sigma(\mathscr{V})$. Thus, by the Daniell–Stone theorem ([5], Satz 39.4), there is a measure μ on $\sigma(\mathscr{V})$ such that $\Phi(f) = \int f d\mu$ for $f \in \mathscr{V}$. It follows $\mu(X) = \Phi(1) = \lim_n \Phi_n(1) = 1$.

Next we show that μ is T-invariant. For this purpose, it suffices to prove $\mu(H) = \mu(T^{-1}H)$ for all $H \in \mathscr{H}$ where $\mathscr{H} := \{\{f = 0\} : f \in \mathscr{V}\}$ is a lattice generating $\sigma(\mathscr{V})$. Let $H \in \mathscr{H}$ be given, i.e. $H = \{f = 0\}$ for some $f \in \mathscr{V}$. Put $f_n := 1 - \min(1, n|f|)$ for $n \in \mathbf{N}$. Then $\mathscr{V} \ni f_n \downarrow 1_H$ and hence $f_n \circ T \downarrow 1_{T^{-1}H}$. It follows

$$\mu(T^{-1}H) = \lim_n \int f_n \circ T d\mu = \lim_n \Phi(f_n \circ T) = \lim_n \Phi(f_n) = \lim_n \int f_n d\mu = \mu(H).$$

(3)$\Rightarrow$(1). Let μ be a T-invariant probability measure on $\sigma(\mathscr{V})$. Since $\mathscr{V}$ is a subset of $\mathscr{L}_1(X, \sigma(\mathscr{V}), \mu)$, (1) follows from Birkhoff's ergodic theorem (see [13], Theorem 1.14). □

3.4. REMARKS. (a) If, in addition to the assumptions of 3.3, $\mathscr{V}$ is separable (with respect to the topology of uniform convergence), then a straightforward argument (cf. the proof of Lemma 6.13 in [13]) shows that each of the statements (1)–(3) of 3.3 is equivalent to

(4) There is a probability measure ν on $\sigma(\mathscr{V})$ and a set $X_0 \in \sigma(\mathscr{V})$ with $\nu(X_0) = 1$ such that $\lim_n A_n f(x)$ exists for all $x \in X_0$ and all $f \in \mathscr{V}$.

(b) In view of Birkhoff's ergodic theorem, every T-invariant probability measure ν on $\sigma(\mathscr{V})$ satisfies condition (1) of 3.3. However, the converse is not true, in general: Let $X := [-1, 1]$, $T(x) := -x$ and $\mathscr{V} := \mathscr{C}^b(X)$. As the Lebesgue measure λ is T-invariant, the sequence $(A_n f)$ converges λ-a.e. for every $f \in \mathscr{V}$. Let $g \in \mathscr{V}_+$ be such that $\int g d\lambda = 1$ and $\int_{[0,1]} g d\lambda \neq \frac{1}{2}$. Denote by ν the measure on $\sigma(\mathscr{V})$ having the density g. Then $\nu \ll \lambda$ which implies that the sequence $(A_n f)$ converges ν-a.e. for every $f \in \mathscr{V}$, i.e. ν satisfies condition (1) of 3.3. However, ν is not T-invariant, since we have $\nu([-1, 0]) \neq \nu([0, 1])$.

The assumption that $\mathscr{V}$ has the Alexandroff property is essential for the validity of 3.3 as the following example shows.

3.5. Example. Let X be the set of integers, $\mathscr{A} := \{A \subset X : A$ or $X - A$ is finite$\}$ and $\mathscr{V}$ the vector lattice of all $\mathscr{A}$-step functions. We consider the translation $Tx := x + 1$. Then $f \circ T \in \mathscr{V}$ for $f \in \mathscr{V}$, and it is well known that there is no T-invariant probability measure on $\sigma(\mathscr{V}) = \mathscr{P}(X)$. On the other hand, it is easy to see that $\lim_n A_n f(x)$ exists for all $x \in X$ and all $f \in \mathscr{V}$.

3.6. Corollary. *Let X be a topological space and $T: X \to X$ a continuous map. Then the following statements are equivalent:*

(1) *There is a Baire probability measure ν on X such that, for every $f \in \mathscr{C}^b(X)$, the sequence $(A_n f)$ converges ν-a.e.*
(2) *There is a Baire probability measure ν on X such that, for every $f \in \mathscr{C}^b(X)$, the sequence $(A_n f)$ converges in distribution.*
(3) *There is a T-invariant Baire probability measure on X.*

Proof. For $\mathscr{V} := \mathscr{C}^b(X)$, our claim follows from 3.3. Note that $\mathscr{V}$ has the Alexandroff property by 2.4(b). □

Now we consider a probability space $(X, \mathscr{A}, \nu)$ and an $\mathscr{A}, \mathscr{A}$-measurable mapping $T: X \to X$. If $\mathscr{V}$ denotes the vector lattice of all $\mathscr{A}$-step functions, then it is obvious that the two statements

(3.1) $(A_n E)$ converges ν-a.e. for every $E \in \mathscr{A}$

and

(3.2) $(A_n f)$ converges ν-a.e. for every $f \in \mathscr{V}$

are equivalent. Since $\mathscr{V}$ has the Alexandroff property by 2.6, we thus infer from 3.3

3.7. Corollary. *Let $(X, \mathscr{A})$ be a measurable space and $T: X \to X$ an $\mathscr{A}, \mathscr{A}$-measurable mapping. Then the following two statements are equivalent:*

(1) *There is a probability measure ν on $\mathscr{A}$ such that, for every $E \in \mathscr{A}$, the sequence $(A_n E)$ converges ν-a.e.*
(2) *There is a T-invariant probability measure on $\mathscr{A}$.*

We will now give a sharpening of 3.7 in so far as we will show that even the convergence in distribution of the sequence $(A_n E)$ for all elements E of a certain subclass of $\mathscr{A}$ is sufficient for the existence of a T-invariant probability measure on $\mathscr{A}$. For this purpose we need

3.8. Definition. Let $(X, \mathscr{A})$ be a measurable space, and let $\mathscr{M}$ be a family of probability measures on $\mathscr{A}$.

(a) A subset $\mathscr{A}_0$ of $\mathscr{A}$ is said to be a *convergence class for* $\mathscr{M}$ ([3]) if, for every sequence (μ_n) in $\mathscr{M}$, the existence of $\lim_n \mu_n(E)$ for all $E \in \mathscr{A}_0$ implies the existence of $\lim_n \mu_n(A)$ for all $A \in \mathscr{A}$.
(b) If $T: X \to X$ is $\mathscr{A}, \mathscr{A}$-measurable, then $\mathscr{M}$ is called *T-closed* if we have $T(\mu) \in \mathscr{M}$ for all $\mu \in \mathscr{M}$.
(c) $\mathscr{M}$ is said to be *sequentially closed* (with respect to the topology of setwise convergence on $\mathscr{A}$) if, for every sequence (μ_n) in $\mathscr{M}$ such that $\mu(A) := \lim_n \mu_n(A)$ exists for all $A \in \mathscr{A}$, we have $\mu \in \mathscr{M}$.

Remark. Let $\mathscr{A}_0$ be a convergence class for $\mathscr{M}$, and let $\mu, \nu \in \mathscr{M}$. If $\mu(A) = \nu(A)$ for all $A \in \mathscr{A}_0$, then $\mu = \nu$.

3.9. Theorem. *Let $(X, \mathscr{A})$ be a measurable space and $T: X \to X$ an $\mathscr{A}, \mathscr{A}$-measurable mapping. Furthermore, let $\mathscr{M}$ be a convex family of probability measures on $\mathscr{A}$ which is both T-closed and sequentially closed. If $\mathscr{A}_0 \subset \mathscr{A}$ is a convergence class for $\mathscr{M}$, then the following three statements are equivalent:*

(1) *There is a probability measure $\nu \in \mathscr{M}$ such that, for every $E \in \mathscr{A}_0$, the sequence $(A_n E)$ converges ν-a.e.*
(2) *There is a probability measure $\nu \in \mathscr{M}$ such that, for every $E \in \mathscr{A}_0$, the sequence $(A_n E)$ converges in distribution.*
(3) *There is a T-invariant probability measure $\mu \in \mathscr{M}$.*

Proof. (1)$\Rightarrow$(2) is evident.

(2)$\Rightarrow$(3). By assumption, there is a $\nu \in \mathscr{M}$ such that, for every $E \in \mathscr{A}_0$, the sequence $(A_nE(\nu))$ converges weakly to a probability measure ρ_E on $\mathscr{B}(\mathbf{R})$. From 3.2 we infer

$$\lim_n \int A_nE d\nu = \lim_n \int A_nT^{-1}E d\nu = \int \mathrm{id}_{\mathbf{R}}\, d\rho_E \qquad \text{for } E \in \mathscr{A}_0.$$

Now we define

$$\mu_n(F) := \int A_nF d\nu = \frac{1}{n} \cdot \sum_{k=0}^{n-1} T^k(\nu)(F) \qquad \text{for } F \in \mathscr{A} \quad \text{and} \quad n \in \mathbf{N}.$$

As $\mathscr{M}$ is convex and T-closed, we have $(\mu_n) \subset \mathscr{M}$. In addition, $\lim_n \mu_n(E) = \lim_n \mu_n(T^{-1}E)$ exists for all $E \in \mathscr{A}_0$. Since $\mathscr{A}_0$ is a convergence class for the sequentially closed set $\mathscr{M}$, there exists a probability measure $\mu \in \mathscr{M}$ such that $\mu(A) = \lim_n \mu_n(A)$ holds for all $A \in \mathscr{A}$. Furthermore, for $E \in \mathscr{A}_0$, we have

$$\mu(E) = \lim_n \mu_n(E) = \lim_n \mu_n(T^{-1}E) = \mu(T^{-1}E).$$

Thus μ is T-invariant by the preceding remark.

(3)$\Rightarrow$(1) holds in view of Birkhoff's ergodic theorem. □

An analysis of the proof of the implication (2)$\Rightarrow$(3) of 3.9 reveals that the arguments of Wright ([14]) can be used to prove the following generalization of the main results of [14].

3.10. Theorem. *Let $(X, \mathscr{A}, \nu)$ be a probability space and $T: X \to X$ an $\mathscr{A}, \mathscr{A}$-measurable mapping such that $T(\nu) \ll \nu$. Assume that $\nu \in \mathscr{M}$ where $\mathscr{M}$ is a convex family of probability measures on $\mathscr{A}$ which is both T-closed and sequentially closed. Furthermore, let $\mathscr{A}_0 \subset \mathscr{A}$ be a convergence class for $\mathscr{M}$ such that, for every $E \in \mathscr{A}_0$, the sequence (A_nE) converges in distribution. Then there exists a T-invariant probability measure $\mu \in \mathscr{M}$ such that*

(i) $\mu \ll \nu$ *and* (ii) $\mu(A) = \nu(A)$ *for* $A \in \mathscr{A}$ *with* $\nu(A \triangle T^{-1}A) = 0$ *holds.*

If, in addition, T is incompressible (i.e. $\nu(A - T^{-1}A) = 0$ implies $\nu(T^{-1}A - A) = 0$) or T is an automorphism (i.e. T is bijective and T^{-1} is also $\mathscr{A}, \mathscr{A}$-measurable), then (i) *can be replaced by the stronger property* (i′) $\mu \sim \nu$.

Next we will study several topological special cases of 3.9.

3.11. Corollary. *Let X be a topological space and $T: X \to X$ a $\mathscr{B}_0(X)$, $\mathscr{B}_0(X)$-measurable mapping. Then the following statements are equivalent:*

(1) *There is a Baire probability measure on X such that, for any* $U \in \mathscr{U}(X)$, *the sequence* $(A_n U)$ *converges in distribution.*
(2) *There is a T-invariant Baire probability measure on X.*

PROOF. Let $\mathscr{M}$ be the family of all probability measures on $\mathscr{A} := \mathscr{B}_0(X)$. By [3], 4.2, $\mathscr{A}_0 := \mathscr{U}(X)$ is a convergence class for $\mathscr{M}$. Thus our claim follows from 3.9. □

3.12. COROLLARY. *Let X be a Hausdorff space and* $T: X \to X$ *a continuous map. Then the following statements are equivalent:*
(1) *There is a Radon probability measure on X such that, for any* $G \in \mathscr{G}(X)$, *the sequence* $(A_n G)$ *converges in distribution.*
(2) *There is a T-invariant Radon probability measure on X.*

PROOF. Let $\mathscr{M}$ be the family of all Radon probability measures on X. By [3], 4.3, $\mathscr{A}_0 := \mathscr{G}(X)$ is a convergence class for $\mathscr{M}$. Thus 3.9 completes the proof. □

3.13. COROLLARY. *Let X be a regular topological space and* $T: X \to X$ *a continuous map. Then the following statements are equivalent:*
(1) *There is a* τ*-smooth Borel probability measure on X such that, for any* $G \in \mathscr{G}_r(X)$, *the sequence* $(A_n G)$ *converges in distribution.*
(2) *There is a T-invariant* τ*-smooth Borel probability measure on X.*

PROOF. Let $\mathscr{M}$ be the family of all τ-smooth probability measures on $\mathscr{B}(X)$. By [3], 4.9, $\mathscr{A}_0 := \mathscr{G}_r(X)$ is a convergence class for $\mathscr{M}$. Now our claim follows from 3.9. □

3.14. COROLLARY. *Let X be a completely regular topological space and* $T: X \to X$ *a continuous map. Then the following statements are equivalent:*
(1) *There is a* τ*-smooth Borel probability measure on X such that, for any* $U \in \mathscr{U}(X)$, *the sequence* $(A_n U)$ *converges in distribution.*
(2) *There is a T-invariant* τ*-smooth Borel probability measure on X.*

PROOF. Let $\mathscr{M}$ be the family of all τ-smooth probability measures on $\mathscr{B}(X)$. By [3], 4.5, $\mathscr{A}_0 := \mathscr{U}(X)$ is a convergence class for $\mathscr{M}$. Now 3.9 completes the proof. □

3.15. COROLLARY. *Let X be a normal topological space and* $T: X \to X$ *a closed continuous map. Then the following statements are equivalent:*
(1) *There is a regular Borel probability measure on X such that, for any* $U \in \mathscr{U}(X)$, *the sequence* $(A_n U)$ *converges in distribution.*

(2) *There is a regular Borel probability measure on X such that, for any* $G \in \mathscr{G}_r(X)$, *the sequence* $(A_n G)$ *converges in distribution.*
(3) *There is a T-invariant regular Borel probability measure on X.*

PROOF. Let $\mathscr{M}$ be the family of all regular probability measures on $\mathscr{B}(X)$. Since T is closed and continuous, $\mathscr{M}$ is T-closed. By [3], 4.6 and 4.10, both $\mathscr{U}(X)$ and $\mathscr{G}_r(X)$ are convergence classes for $\mathscr{M}$. Now our claim follows from 3.9. □

REMARK. In view of 3.9, one can replace, in the statements 3.10–3.15, the convergence in distribution by the convergence a.e.

Next we will present a substantial generalization of a theorem due to Oxtoby and Ulam [10] (cf. 3.21). As for the proof, we have taken over from [10] the basic idea of using Banach limits. However, whereas Oxtoby and Ulam apply Carathéodory's outer measure method, our procedure is based on an inner measure approach.

Recall that a *Banach limit* is a real-valued linear functional L defined on the vector space l_∞ of all bounded sequences of real numbers such that

(i) $L(x_0, x_1, x_2, \dots) \geqq 0$ if $x_k \geqq 0$ for all $k \in \mathbb{N}_0$;
(ii) $L(x_0, x_1, x_2, \dots) = L(x_1, x_2, x_3, \dots)$;
(iii) $L(1, 1, 1, \dots) = 1$.

We need the following functional analytic result which is a simple consequence of the Hahn–Banach theorem (see [9], pp. 64–65).

3.16. LEMMA. *Let* $y = (y_0, y_1, y_2, \dots)$ *be a fixed element of* l_∞. *Then there exists a Banach limit L such that*

$$L(y) = \limsup_{n \to \infty} \frac{1}{n} \cdot \sum_{i=0}^{n-1} y_i.$$

3.17. THEOREM. *Let* $\mathscr{K} \subset \mathscr{P}(X)$ *be a semicompact lattice,* $\mathscr{A}$ *a σ-algebra in X with* $\mathscr{K} \subset \mathscr{A} \subset \sigma(\mathscr{F}(\mathscr{K}_\delta))$ *and* $T: X \to X$ *an* $\mathscr{A}, \mathscr{A}$*-measurable mapping. Assume that the family* $\mathscr{M}$ *of all* $\mathscr{K}_\delta$*-regular probability measures on* $\mathscr{A}$ *is T-closed. Furthermore, let* $\mathscr{G} \subset \mathscr{P}(X)$ *be a lattice satisfying the following conditions*:

(α) *For every* $K \in \mathscr{K}$, *there is a set* $G \in \mathscr{G}$ *such that* $K \subset G$.
(β) $K \in \mathscr{K}$ *and* $G \in \mathscr{G}$ *imply* $K - G \in \mathscr{K}$.
(γ) $\mathscr{G}$ *separates* $\mathscr{K}$ (*i.e. for any disjoint sets* $K_1, K_2 \in \mathscr{K}$, *there are disjoint sets* $G_1, G_2 \in \mathscr{G}$ *such that* $K_i \subset G_i$ *for* $i = 1, 2$).

(δ) $T^{-1}\mathscr{G} \subset \mathscr{G}$.

Then the following three statements are equivalent:

(1) *There exist a set* $K_0 \in \mathscr{K}$ *and a point* $x_0 \in X$ *such that* $\lim_n A_n K_0(x_0) > 0$.

(2) *There exist a set* $K_0 \in \mathscr{K}$ *and a point* $x_0 \in X$ *such that*

$$\limsup_{n \to \infty} A_n K_0(x_0) > 0.$$

(3) *There exists a T-invariant element of* $\mathscr{M}$.

Proof. We only prove (2)$\Rightarrow$(3), since the implication (3)$\Rightarrow$(1) can be proved in the same way as the corresponding part of Theorem 1 in [10]. For any $Q \subset X$ write $s_Q := (1_Q(T^k(x_0)))_{k \in \mathbf{N}_0}$. Let L be a Banach limit with the additional property

$$L(s_{K_0}) = \limsup_{n \to \infty} A_n K_0(x_0)$$

(see 3.16).

Define a set function $\gamma : \mathscr{G} \to [0, 1]$ by $\gamma(G) := L(s_G)$. Obviously γ is isotone, additive and subadditive. For any $K \in \mathscr{K}$, set $\lambda(K) := \inf\{\gamma(G) : K \subset G \in \mathscr{G}\}$. By [11], Lemma 2, λ is a tight content. Thus, by [1], 2.7, λ can be extended to a $\mathscr{K}_\delta$-regular measure μ on $\mathscr{A}$. Observe that $\lambda(K) \leqq 1$ for all $K \in \mathscr{K}$ and $\lambda(K_0) \geqq L(s_{K_0}) > 0$ by (2). Therefore, w.l.o.g., we can and do assume $\mu(X) = 1$, i.e. $\mu \in \mathscr{M}$. Since $\mathscr{M}$ is T-closed, we have $\nu := T(\mu) \in \mathscr{M}$. To prove the T-invariance of μ it suffices to show $\nu(K) \leqq \mu(K)$ for $K \in \mathscr{K}$. This implies $\nu \leqq \mu$ which together with $\nu(X) = \mu(X)$ yields $\nu = \mu$.

Suppose that we have $\nu(K) > \mu(K)$ for some $K \in \mathscr{K}$. Choose a set $G \in \mathscr{G}$ such that $K \subset G$ and $\nu(K) > \gamma(G)$. As L is a Banach limit, we have $\gamma(G) = \gamma(T^{-1}G)$ and so $\mu(T^{-1}K) > \gamma(T^{-1}G)$. On the other hand, $T^{-1}K \subset T^{-1}G$ implies $\mu(T^{-1}K) \leqq \gamma(T^{-1}G)$. This contradiction proves our claim. □

3.18. Corollary. *Let* $\mathscr{V} \subset \mathbf{R}^X$ *be a Daniell lattice of bounded functions with* $1 \in \mathscr{V}$. *If* $T : X \to X$ *is a mapping such that* $f \circ T \in \mathscr{V}$ *for* $f \in \mathscr{V}$, *then there exists a T-invariant probability measure on* $\sigma(\mathscr{V})$.

Proof. Put $\mathscr{K} := \{\{f \geqq 1\} : f \in \mathscr{V}\}$ and $\mathscr{G} := \{\{f > 1\} : f \in \mathscr{V}\}$. It is easy to see that $\mathscr{K}$ and $\mathscr{G}$ are lattices of subsets of X which satisfy the conditions (α)–(δ) of 3.17. In addition, $\mathscr{K}$ is semicompact by [2], Corollary 1. Since $1 \in \mathscr{V}$ and $\mathscr{A} := \sigma(\mathscr{V})$ is generated by $\mathscr{K}$, every probability measure on $\mathscr{A}$ is $\mathscr{K}_\delta$-regular. As condition (1) of 3.17 is satisfied for $K_0 = X$, the proof is complete. □

3.19. Corollary. *Let* $\mathscr{L} \subset \mathscr{P}(X)$ *be a* δ*-lattice with* $X \in \mathscr{L}$ *such that* $\mathscr{C}^b(\mathscr{L}) = \mathscr{C}(\mathscr{L})$ *holds. Then, for any mapping* $T: X \to X$ *with* $T^{-1}\mathscr{L} \subset \mathscr{L}$, *there exists a T-invariant probability measure on* $\sigma(\mathscr{C}^b(\mathscr{L}))$.

Proof. By [2], Theorem 3, $\mathscr{V} := \mathscr{C}^b(\mathscr{L})$ is a Daniell lattice with $1 \in \mathscr{V}$. If $T: X \to X$ is a mapping satisfying $T^{-1}\mathscr{L} \subset \mathscr{L}$, then $f \circ T \in \mathscr{V}$ for $f \in \mathscr{V}$. Thus our claim follows from 3.18. □

3.20. Corollary. *Let X be a pseudocompact topological space. Then, for any continuous mapping* $T: X \to X$, *there exists a T-invariant Baire probability measure on X.*

Proof. Apply 3.19 with $\mathscr{L} := \mathscr{F}(X)$. □

Remarks. (a) The assumption of pseudocompactness is essential for the validity of 3.20: Consider the translation $T(x) = x + 1$ on the real line $\mathbf{R}$. T is a homeomorphism, but there is no T-invariant probability measure on $\mathscr{B}(\mathbf{R})$.

(b) It seems to be unknown whether the converse of 3.20 is also true, i.e. is a topological space X pseudocompact if for any continuous map $T: X \to X$ a T-invariant Baire probability measure on X exists?

The following result extends Theorem 1 of [10] from Polish spaces to arbitrary Hausdorff spaces.

3.21. Corollary. *Let X be a Hausdorff space and* $T: X \to X$ *a continuous map. Then the following three statements are equivalent:*

(1) *There exist a compact set* $K_0 \subset X$ *and a point* $x_0 \in X$ *such that* $\lim_n A_n K_0(x_0) > 0$.

(2) *There exist a compact set* $K_0 \subset X$ *and a point* $x_0 \in X$ *such that*

$$\limsup_{n \to \infty} A_n K_0(x_0) > 0.$$

(3) *There exists a T-invariant Radon probability measure on X.*

Proof. Immediate consequence of 3.17 with $\mathscr{K} := \mathscr{K}(X)$, $\mathscr{G} := \mathscr{G}(X)$ and $\mathscr{A} := \mathscr{B}(X)$. □

3.22. Corollary. *Let X be an uncountable set and put*

$$\mathscr{K} := \{K \subset X : K \text{ or } X - K \text{ is finite}\}.$$

If $T: X \to X$ *is a mapping such that* $T^{-1}\mathscr{K} \subset \mathscr{K}$ *holds, then there is a*

T-invariant probability measure on $\mathscr{A} := \sigma(\mathscr{K}) = \{A \subset X : A$ *or* $X - A$ *is countable*$\}$.

PROOF. $\mathscr{K}$ is a semicompact algebra. Thus our claim follows from 3.17 with $\mathscr{G} := \mathscr{K}$. □

REMARKS. (a) The following simple example shows that the condition $T^{-1}\mathscr{K} \subset \mathscr{K}$ is essential for the validity of 3.22:

Let $X := [0, \infty)$ and $A_0 := X - \bigcup_{n \in \mathbf{N}} A_n$ where A_n denotes the set of all rationals in the interval $[n-1, n)$, $n \in \mathbf{N}$. Then $A_n \in \sigma(\mathscr{K})$ for $n \in \mathbf{N}_0$. Fix some element $x_n \in A_n$ and define $T(x) := x_{n+1}$ if $x \in A_n$, $n \in \mathbf{N}_0$. Obviously T is $\sigma(\mathscr{K})$, $\sigma(\mathscr{K})$-measurable. Assume that there is a T-invariant probability measure μ on $\sigma(\mathscr{K})$. Setting $p_n := \mu(A_n)$ for $n \in \mathbf{N}_0$, we have

$$0 \leqq p_n = \mu(T^{-1}x_{n+1}) = \mu(\{x_{n+1}\}) \leqq \mu(A_{n+1}) = p_{n+1}.$$

Thus (p_n) is increasing which, however, contradicts $\Sigma_{n=0}^{\infty} p_n = 1$.

(b) Let X be a countable set. Then the algebra $\mathscr{K}$ is, in general, not semicompact. Equipping X with the discrete topology, we infer from 3.21 that, for an arbitrary map $T: X \to X$, a T-invariant probability measure on $\sigma(\mathscr{K})$ exists if and only if there are two points $x, y \in X$ such that

$$\limsup_{n \to \infty} A_n\{y\}(x) > 0$$

holds. That the latter condition is not implied by the inclusion $T^{-1}\mathscr{K} \subset \mathscr{K}$ can be seen e.g. for the translation $T(x) := x + 1$ on the integers.

REFERENCES

1. W. Adamski, *Tight set functions and essential measure*, Lecture Notes in Math., Vol. 945, Springer-Verlag, Berlin, Heidelberg and New York, 1982, pp. 1–14.

2. W. Adamski, *Some characterizations of Daniell lattices*, Arch. Math. **40** (1983), 339–345.

3. W. Adamski, P. Gänßler and S. Kaiser, *On compactness and convergence in spaces of measures*, Math. Ann. **220** (1976), 193–210.

4. A. D. Alexandroff, *Additive set-functions in abstract spaces*. Mat. Sbornik (N.S.) **13** (1943), 169–238.

5. H. Bauer, *Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie*, De Gruyter, Berlin and New York, 1978.

6. P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

7. Y. S. Chow and H. Teicher, *Probability Theory*, Springer-Verlag, New York, Heidelberg and Berlin, 1978.

8. N. Dunford and J. T. Schwartz, *Linear Operators*, Part I, Interscience Publishers, New York, 1958.

9. C. Goffman and G. Pedrick, *First Course in Functional Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1965.

10. J. C. Oxtoby and S. M. Ulam, *On the existence of a measure invariant under a transformation*, Ann. of Math. **40** (1939), 560–566.

11. F. Topsoe, *Compactness in spaces of measures*, Studia Math. **36** (1970), 195–212.

12. V. S. Varadarajan, *Measures on topological spaces*, Am. Math. Soc. Transl. (2) **48** (1965), 161–228.

13. P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, New York, Heidelberg and Berlin, 1982.

14. F. B. Wright, *The converse of the individual ergodic theorem*, Proc. Am. Math. Soc. **11** (1960), 415–420.