Monatshefte für Mathematik 79, 157—163 (1975)

Remarks on Invariant Measures for Number Theoretic Transformations*

By

Michael S. Waterman, Pocatello, Idaho, USA

(Received 24 September 1973)

Abstract

A sufficient condition for the existence of an invariant measure is given that is useful in number theory. A connection with a problem of Blum is pointed out. Kuzmin's theorem is considered from an operator point of view, and the Chacon-Ornstein theorem is applied to give almost everywhere Cesaro convergence to the density of the invariant measure.

1. Introduction

Much of the modern work in the metric theory of number theoretic expansions depends on an invariant measure for the shift transformation associated with these expansions. One of the first problems is the existence of an invariant measure. For a few cases, such as the one-dimensional continued fraction, q-adic expansions, and β-expansions, the invariant measure is known [6, 8]. However, in the more general situation such as f-expansions and F-expansions, it is not easy to exhibit the invariant measure and an existence theorem is required [2, 6, 11]. Section 2 of this paper gives a proof of the existence of an invariant measure that is distinct from the methods used in the papers cited above. The considerations of that section are somewhat more general than indicated here.

Even after an invariant measure μ is shown to exist we must search for an effective way to compute it. That is, we would like to evaluate $\mu(A)$ for all Borel sets A. Kuzmin's theorem and its generalizations [3, 4, 9, 10] have been very useful in this regard.

* This work was supported in part by National Science Foundation grants GP-28313 and GP-28313§1. The work was partly supported under the auspices of the U. S. Atomic Energy Commission while the author was a faculty participant of the Associated Western Universities at Los Alamos Scientific Laboratory.

The proofs of these theorems are quite difficult, however, and it is not clear how to generalize them to general ergodic theory. Section 3 presents an application of the Chacon-Ornstein theorem to the Kuzmin "operator" and we obtain almost everywhere Cesaro convergence to the density of the invariant measure.

2. Existence of an Invariant Measure

The results of this section follow from the theory of Banach limits. Let S be the set of all bounded sequences of reals $s=\{s_n\}_{n\geqslant 1}$ with

$$\|s\| = \sup_n |s_n| .$$

The set S^* of convergent sequences is a subspace of S. Define $p(s)=\limsup_n (s_n)$ and $f(s)=\liminf_n (s_n)$. These functions coincide with the continuous linear functional $l(s)=\lim_n (s_n)$, $s\in S^*$. By the Hahn—Banach theorem [13, p. 102—104], there exists a linear functional L defined on S which satisfies

$$f(s) \leqslant L(s) \leqslant p(s),\ \ L(s)=l(s), s\in S^*,\ \ L(s)=L(t), \tag{1}$$

where $t_n=s_{n+1}$, $n\geqslant 1$. This functional L will be called a Banach limit.

Theorem 1 resulted from an examination of theorem 6.2 of "The Metrical Theory of Jacobi—Perron Algorithm" by F. SCHWEIGER [10].

Theorem 1. *Let $\lambda, \nu_n\, (n\geqslant 1)$ be probability measures on the measurable space $(\Omega,\mathscr{A})$ and assume there exists a constant $C>0$ such that $\nu_n(A)\leqslant C\,\lambda(A)$ for all $A\in\mathscr{A}$. Then $\mu(A)=L(\{\nu_n(A)\})$ is a probability measure on $(\Omega,\mathscr{A})$ and satisfies $\mu(A)\leqslant C\,\lambda(A)$.*

Proof. By the above remarks $\mu(A)$ is defined, and

$$\mu(\emptyset)=L(\{\nu_n(\emptyset)\})=L(\{0\})=0,$$

and

$$\mu(\Omega)=L(\{\nu_n(\Omega)\})=L(\{1\})=1.$$

Suppose $A\cap B=\emptyset$. Then $\nu_n(A\cup B)=\nu_n(A)+\nu_n(B)$ and $\mu(A\cup B)= = L(\{\nu_n(A)\}+\{\nu_n(B)\})=\mu(A)+\mu(B)$. Thus μ is a finitely additive measure. Since $\nu_n(A)\leqslant C\,\lambda(A)$, $\mu(A)=L(\{\nu_n(A)\})\leqslant C\,\lambda(A)$.

The only thing left to show is that μ is countably additive. But $A_n \downarrow \emptyset$ implies $\lambda(A_n) \downarrow 0$ so that $\mu(A_n) \downarrow 0$. By HALMOS [5, p. 39], we have μ countably additive and theorem 1 is proved.

Theorem 1 is true if we assume λ_n $(n \geqslant 1)$ are uniformly absolutely continuous with respect to λ.

We next apply theorem 1 to the problem of an invariant measure.

Corollary 1. Let T be a measurable transformation defined on $(\Omega, \mathscr{A}, \lambda)$ and suppose $\lambda(T^{-n}A) \leqslant C\lambda(A)$ for $n \geqslant 0$ and all $A \in \mathscr{A}$. Then there exists a probability measure $\mu \ll \lambda$ such that T is a measure preserving transformation for μ.

Proof. Let $\nu_n(A) = \lambda(T^{-n}A)$. Now ν_n is a measure and theorem 1 applies so that μ exists and $\mu(A) \leqslant C\lambda(A)$. Therefore $\mu \ll \lambda$. Also by the property $L(s) = L(t)$ in (1), we have $\mu(A) = \mu(T^{-1}A)$ for all $A \in \mathscr{A}$.

We now present some applications of corollary 1 to number theory. In this paragraph λ will denote Lebesgue measure. For f-expansions, RENYI [8] shows that

$$T_1(x) = f^{-1}(x) - [f^{-1}(x)]$$

satisfies $\lambda(T_1^{-1}A) \leqslant C\lambda(A)$. These expansions include 1-dimensional continued fractions and q-adic expansions where $q \in I$ and $q \geqslant 2$. Renyi also shows that for $\beta > 1$, $\beta \notin I$,

$$T_2(x) = \beta x - [\beta x]$$

satisfies $\lambda(T_2^{-1}A) \leqslant (\beta/(\beta - 1))\lambda(A)$. For n-dimensional F-expansions, WATERMAN [11] shows that

$$T_3(x) = F^{-1}(x) - [F^{-1}(x)]$$

satisfies $\lambda(T_3^{-1}A) \leqslant CL^{-1}\lambda(A)$.

The final result of this section is connected with a problem of BLUM [12, p. 302].

Theorem 2. *Let $\lambda, \nu_n (n \geqslant 1)$ be probability measures on the measurable space $(\Omega, \mathscr{A})$ satisfying ν_n uniformly absolutely continuous with respect to λ for all $A \in \mathscr{A}$. Then, if for all $A \in \mathscr{A}$, $\lim_{n\to\infty} (\nu_n(TA) - \nu_n(A)) = 0$ for an invertible measurable transformation T on $(\Omega, \mathscr{A})$, we have T a measure preserving transformation for μ defined by $\mu(A) = L(\{\nu_n(A)\})$.*

Proof. μ exists by the extension of theorem 1 mentioned above. Also

$$L(\{\nu_n(TA)-\nu_n(A)\}) = L(\{\nu_n(TA)\}) - L(\{\nu_n(A)\}) = \mu(TA) - \mu(A).$$

Then, since $\nu_n(TA)-\nu_n(A)\to 0$ as $n\to\infty$, $\mu(TA)=\mu(A)$.

3. The Kuzmin Operator

Let $(\Omega,\mathscr{A},\lambda)$ be a probability space and T be a measurable transformation of Ω into Ω. We assume T is ergodic and that there exists a measure $\mu\sim\lambda$ such that T is a measure preserving transformation for μ. The following additional assumptions are made:

(i) If $A\in\mathscr{A}$, then $TA\in\mathscr{A}$.

(ii) The numbers $\lambda(A)$, $\lambda(TA)$ and $\lambda(T^{-1}A)$ are all positive or all zero.

(iii) There exists a measurable partition $A_1, A_2, A_3, \ldots$ of Ω such that T is one-to-one on each A_i.

If $(\Omega,\mathscr{A},\lambda)$ is a Lebesgue space and T is essentially countable to one, then, according to Parry [7, p. 106—108], (i), (iii), and $\lambda(A)=0$ implies $\lambda(TA)=0$ all hold for a transformation isometric to T.

Since T is one-to-one on A_i, we have an inverse mapping V_i: $TA_i\to A_i$. We can define the Jacobian [7, p. 108] of this mapping by

$$J_i(x)=\frac{d\lambda V_i}{d\lambda}(x),\ x\in TA_i.$$

For $f\in L_1=L_1(\Omega,\mathscr{A},\lambda)$ we define $\|f\|=\int_\Omega |f(t)|\,d\lambda(t)$. Finally we define the *Kuzmin operator*, Ψ, generated by T for $f\in L_1$:

$$\Psi f(x)=\sum_{x\in TA_i} f(V_i(x))J_i(x).$$

Lemma 1. *If $f\in L_1$ and $f\geqslant 0$, then $\|f\|=\|\Psi f\|$.*

Proof. $\int \Psi f(x)\,d\lambda(x)=\int\sum_{x\in TA_i} f(V_i(x))J_i(x)\,d\lambda(x)=$

$$=\int\sum_i I_{TA_i}(x)f(V_i(x))J_i(x)\,d\lambda(x)=$$

$$=\sum_i\int I_{TA_i}(x)f(V_i(x))J_i(x)\,d\lambda(x)=$$

$$=\sum_i\int_{TA_i} f(V_i(x))J_i(x)\,d\lambda(x)=\sum_i\int_{A_i} f(t)\,d\lambda(t)=\int f\,d\lambda.$$

The interchange of summation and integration is by the monotone convergence theorem, and the next to the last step is by the transformation theorem [5, p. 163].

A modification of the proof of this lemma shows $\|\Psi f\| \leqslant \|f\|$ so that $\Psi: L_1 \to L_1$ and

$$\|\Psi\| = \sup_{f \in L_1} \frac{\|\Psi f\|}{\|f\|} = 1 .$$

Also $f \geqslant 0$ implies $\Psi f \geqslant 0$ so that $\Psi \geqslant 0$.

Theorem 3. *Let* $\varrho = d\mu/d\lambda$. *Then* $\Psi \varrho = \varrho$.

Proof. Now $\varrho = d\mu/d\lambda \in L_1$. By calculations similar to those in lemma 1 we show

$$\int_E \varrho(x)\, d\lambda(x) = \int_E \Big(\sum_{x \in T A_i} \varrho(V_i(x))\, J_i(x) \Big) d\lambda(x)$$

and the result follows from this equation. See FISCHER's Satz 2 [2] for some related work.

Theorem 4. *Let* $g \in L_1$ *and* $g > 0$. *Then*

$$\lim_{n \to \infty} (1/n) \sum_{k=0}^{n-1} \Psi^n g(x) = \varrho(x) \|g\| \text{ a. e.}$$

Proof. We have shown that Ψ is a linear operator on L_1 with $\|\Psi\| = 1$ and $\Psi \geqslant 0$. Then, by the CHACON—ORNSTEIN ergodic theorem [1],

$$\lim_{n \to \infty} \frac{\sum_{k=0}^{n-1} \Psi^k g(x)}{\sum_{k=0}^{n-1} \Psi^k \varrho(x)}$$

exists almost everywhere for $g > 0$, $g \in L_1$. Denote the limit function by $g^*(x)$. Theorem 3 yields

$$\varrho(x)\, g^*(x) = \lim_{n \to \infty} n^{-1} \sum_{k=0}^{n-1} \Psi^k g(x) .$$

Since $\Psi\big(\lim_{n\to\infty} n^{-1} \sum_{k=0}^{n-1} \Psi^k g(x)\big) = \lim_{n\to\infty} n^{-1} \sum_{k=0}^{n-1} \Psi^k g(x)$, we have $\Psi(g^*(x)\,\varrho(x)) = g^*(x)\,\varrho(x)$ or, since the invariant measure is unique,

$$\varrho(x) = \frac{g^*(x)\,\varrho(x)}{\int g^* \varrho\, d\lambda} .$$

Therefore g^* is constant a. e. It is easy to show that $g^* = \|g\|$.

We note that theorem 4 can be modified in case $\mu \ll \lambda$. Also theorem 4 clearly implies L_1 convergence.

Corollary 2. For all $A \in \mathscr{A}$, we have

$$\lim_{n \to \infty} n^{-1} \sum_{k=0}^{n-1} \lambda(T^{-k}(x)) = \mu(A).$$

Proof. Let $g = 1$ in theorem 4 and multiply both sides of that equation by I_A. Considerations similar to lemma 1 of [9] and our lemma 1 yield

$$\int I_A(x) \Psi^n(1)(x) \, d\lambda(x) = \lambda(T^{-n} A),$$

so that integration yields the result.

The meaning of this result is that for many to one transformations the invariant measure can be obtained from the usual Cesaro limit of the sequence $\lambda(T^{-k}A)$.

Some previous uses of Kuzmin's operator Ψ in number theory has resulted in uniform pointwise convergence of $\Psi^n g$ [3, 9, 11]. In fact Gordin [4] has geometric convergence for a subclass of the expansions he considers. He obtains a L_1 version of our theorem 4 with $g = 1$ for a somewhat larger class of transformations than he has geometric convergence for. Unfortunately, he does not include any proofs of his results. We have been unable to obtain results on the geometric convergence of $\Psi^n g$ in the general situation considered above.

References

[1] Chacon, R. V., and D. S. Ornstein: A general ergodic theorem. Illinois J. Math. **4**, 153—160 (1960).

[2] Fischer, R.: Ergodische Theorie von Ziffernentwicklungen in Wahrscheinlichkeitsräumen. Math. Z. **128**, 217—230 (1972).

[3] Fischer, R.: Mischungsgeschwindigkeit für Ziffernentwicklungen nach reellen Matrizen. Acta Arith. **23**, 5—12 (1973).

[4] Gordin, M. I.: Exponentially fast mixing. Dokl. Akad. Nauk SSSR **196** (1971); English translation: Soviet Math. Dokl. **12**, 331—335 (1971).

[5] Halmos, P. R.: Measure Theory. Toronto, New York, London: D. van Nostrand Co. 1950.

[6] Parry, W.: On the β-expansion of real numbers. Acta Math. Acad. Sci. Hungar. **11**, 401—416 (1960).

[7] Parry, W.: Entropy and Generators in Ergodic Theory. New York and Amsterdam: Benjamin, Inc. 1969.

[8] RENYI, A.: Representations for real numbers and their ergodic properties. Acta Math. Acad. Sci. Hungar. 8, 477—493 (1957).

[9] SCHWEIGER, F., and M. WATERMAN: Some remarks on Kuzmin's theorem for F-expansions. J. Number Theory 5, 123—131 (1973).

[10] SCHWEIGER, F.: The Metrical Theory of Jacobi-Perron Algorithm. Lecture Notes in Mathematics 334. Berlin—Heidelberg—New York: Springer. 1973.

[11] WATERMAN, M. S.: Some ergodic properties of multi-dimensional F-expansions. Z. Wahrscheinlichkeitstheorie verw. Geb. 16, 77—103 (1970).

[12] WRIGHT, F. B. (editor): Ergodic Theory. New York, London: Academic Press. 1963.

[13] YOSHIDA, K.: Functional Analysis. Berlin—Heidelberg—New York: Springer. 1966.

Dr. M. S. WATERMAN
Department of Mathematics
Idaho State University
Pocatello, ID 83201, USA

Printed in Austria