COMPOSITIO MATHEMATICA

JOHANN CIGLER

Some applications of the individual ergodic theorem to problems in number theory

Compositio Mathematica, tome 16 (1964), p. 35-43
<http://www.numdam.org/item?id=CM_1964__16__35_0>

Some applications of the individual ergodic theorem to problems in number theory *

by

Johann Cigler

In this paper we shall give some applications of the individual ergodic theorem to problems which have their origin in number theory, such as the theory of normal numbers and completely uniformly distributed sequences. It will be seen that the concept of normal number admits a very natural generalization to compact topological groups, and we shall show that most of the known theorems on normal numbers remain true in this general case. First of all we state some general results which will be used in the sequel.

Let X be a compact topological Hausdorff space with a countable base. On X we consider the Banach space $C(X)$ of all real-valued continuous functions f with the usual norm $\|f\| = \sup_{x\in X} |f(x)|$. Let $M(X)$ be the set of all positive normed measures μ on X, that is to say, the set of all continuous linear functionals μ on $C(X)$ which satisfy $\mu(f) \geqq 0$ for every nonnegative function f and $\mu(1) = 1$. In $M(X)$ we introduce the weak star topology. This means that a sequence μ_n converges to μ if and only if for every $f \in C(X)$ the sequence $\mu_n(f)$ converges to $\mu(f)$. Let further $L_\mu(X)$ denote the space of all functions integrable with respect to μ.

Let T be a mapping of X onto itself.

We denote the set of measures invariant with respect to T by I_T. This means $\mu \in I_T$ if and only if $\int_X f(Tx)d\mu(x) = \int_X f(x)d\mu(x)$ for all $f \in C(X)$.

We say that μ-almost all $x \in X$ have a certain property if the set of x which do not possess this property has μ-measure zero. We call T ergodic with respect to μ ($\mu \in E_T$) if for each measurable invariant set E ($T^{-1}E = E$ μ-almost everywhere) we have $\mu(E) = 0$ or 1.

A measure ν is said to be absolutely continuous with respect to

* Nijenrode lecture

a measure $\mu(\nu \ll \mu)$, if $\mu(f) = 0$ for an arbitrary function $f \in L_\mu(X)$ implies $\nu(f) = 0$.

Our main tool will be the individual ergodic theorem, which may be stated in the following form:

If $\mu \in E_T$ and $f \in L_\mu(X)$, then for μ-almost all $x \in X$ the relation

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n \leqq N} f(T^n x) = \mu(f) \tag{1}$$

holds.

Another theorem which we shall frequently use is the following one [1]:

If $\nu \in I_T$, $\mu \in E_T$ and $\nu \ll \mu$, then $\nu = \mu$.

The proof is quite easy and runs as follows:

Let $f_n(x)$ denote for brevity $f(T^n x)$. Let A be the set of $x \in X$ which do not satisfy (1). Then $\mu(A) = 0$ and therefore by absolute continuity also $\nu(A) = 0$. From the invariance of ν it follows that $\nu(f) = \nu(f_n)$, $n = 1, 2, 3, \ldots$.

Therefore

$$\nu(f) = \nu\left(\frac{f_1 + \cdots + f_n}{n}\right) = \lim_{n\to\infty} \nu\left(\frac{f_1 + \cdots + f_n}{n}\right)$$
$$= \nu\left(\lim_{n\to\infty} \frac{f_1 + \cdots + f_n}{n}\right) = \nu(\mu(f)) = \mu(f).$$

An easy consequence of this theorem is the following one: (Cf. [2]). *Let F be a set of nonnegative functions f such that every continuous function may be approximated uniformly by linear combinations of functions in F. If there exists a measure $\mu \in E_T$ and a constant $C \geqq 1$, such that*

$$\overline{\lim_{N\to\infty}} \frac{1}{N} \sum_{n \leqq N} f(T^n x) \leqq C\mu(f), \qquad f \in F \tag{2}$$

then the sequence $\{T^n x\}$ possesses the distribution measure μ, i.e.

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n \leqq N} f(T^n x) = \mu(f)$$

for all $f \in C(X)$.

For the proof we note that the set of measures $\{\lambda_N\}$ given by

$$\lambda_N(f) = \frac{1}{N} \sum_{n \leqq N} f(T^n x) \qquad (f \in C(X))$$

[1] Cf. [1], [15].

is compact. If π denotes an arbitrary limit measure, then it follows immediately that $\pi \in I_T$ and that furthermore $\pi(f) \leqq C\mu(f)$ for all $f \in F$. Therefore π is absolutely continuous with respect to μ, and from our theorem we conclude that $\pi = \mu$.

We now consider the following situation: Let X be a compact Abelian topological group with countable base and let T be a continuous endomorphism of X which is ergodic with respect to Haar measure λ on X. The ergodicity of T may be stated in the following purely algebraic way: [2] T is ergodic if and only if for each nontrivial continuous character χ of X all functions $\chi(T^n x)$ $(n = 0, 1, 2, 3, \ldots)$ are different.

We mention some special cases:

Let X be the one-dimensional torus, i.e. the additive group of real numbers reduced mod 1 in its natural topology and let $Tx = ax - [ax]$, where $a > 1$ is an integer. Then the sequence $\{T^n x\}$ becomes the sequence $\{a^n x\}$ mod 1. It is well-known (cf. e.g. [10]) that the sequence $\{a^n x\}$ is uniformly distributed if and only if x is normal to base a.

As another example let X be the k-dimensional torus and let T denote the transformation $\mathfrak{x} \to A\mathfrak{x}$ mod 1, where A is a $k \times k$-matrix with nonvanishing determinant all of whose entries are integers such that no eigenvalue is a root of unity. In analogy with the one-dimensional case we call $\mathfrak{x}$ normal with respect to A if the sequence $\{A^n \mathfrak{x}\}$ is uniformly distributed mod 1.[3] We now turn to the general case. We call an element $x \in X$ normal with respect to the endomorphism T, if the sequence $\{T^n x\}$ is uniformly distributed with respect to Haar measure λ.[4] It follows immediately from the individual ergodic theorem that λ-almost all $x \in X$ are normal with respect to T. For all one has to show is that for almost all $x \in X$ and all nontrivial characters χ of X

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n \leqq N} \chi(T^n x) = 0.$$

The individual ergodic theorem implies this relation for each fixed character χ. From this our theorem follows because the set of characters is denumerable.

Another consequence of the above theorems is that if x is normal with respect to T, then x is also normal with respect to

[2] This theorem is due to Rochlin [18].

[3] A special case has been considered by Maxfield [9].

[4] Cf. [3].

every power of T. From the algebraic characterization of ergodicity we conclude that $\lambda \in E_{T^k}$ for every k. For every nonnegative function $f \in C(X)$ we have the inequality

$$\frac{1}{N}\sum_{n\leqq N} f(T^{kn}x) \leqq k\frac{1}{kN}\sum_{n\leqq kN} f(T^n x).$$

Here the right hand side converges to $k\lambda(f)$ by assumption, therefore we have

$$\overline{\lim_{N\to\infty}}\frac{1}{N}\sum_{n\leqq N} f(T^{kn}x) \leqq k\lambda(f)$$

and this together with our theorem implies

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leqq N} f(T^{kn}) = \lambda(f) \text{ for all } f \in C(X).$$

The converse assertion is trivial, viz. if x is normal with respect to a power of T, then it is normal with respect to T. For we then have

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leqq N} f(T^{kn}x) = \lambda(f)$$

for every continuous function f.

Consider the functions $f(T^j x)$ for $j = 1, 2, \ldots, k-1$ and add these equations. Then one gets

$$\lim_{N\to\infty}\frac{1}{kN}\sum_{n\leqq N}\left(f(T^{kn}x)+f(T^{kn+1}x)+\ldots+f(T^{kn+k-1}x)\right) = \lambda(f)$$

i.e. x is normal with respect to T.

A special case is e.g. the well-known fact that if x is normal to base a, then it is normal to base a^k $(k = 1, 2, 3, \ldots)$ and conversely. In the k-dimensional case we get the corollary: Let A be a nonsingular $k\times k$-matrix with integral entries, none of whose eigenvalues is a root of unity, then the uniform distribution mod 1 of the sequence $\{A^n\mathfrak{x}\}$ implies the uniform distribution of the sequences $\{A^{kn+p}\mathfrak{x}\}$ mod 1 $(l \geqq 1, 0 \leqq p < l)$. Let A denote a matrix of the above form. Then a sufficient condition for the uniform distribution of the sequence $\{A^n\mathfrak{x}\}$ is that for every k-dimensional cube V which lies in the unit cube the relation

$$\overline{\lim_{N\to\infty}}\frac{1}{N}A(N, V) \leqq C|V|$$

holds. Here $A(N, V)$ denotes the number of vectors $\{A^n\mathfrak{x}\}$,

$1 \leqq n \leqq N$, which belong to V and $|V|$ denotes the volume of V.[5]

Let now S be a continuous measure-preserving endomorphism of X. Let H be the set of all $h \in X$ with $Sh = e$. We assume that H is a finite group of order k. Then the following theorem holds: *Let T be an ergodic endomorphism which commutes with S, $TS = ST$. Then from the uniform distribution of the sequence $\{T^n x\}$ follows the uniform distribution of the sequence $\{T^n y\}$ for each $y \in S^{-1} x$.*

PROOF: We first note that if the sequence $\{x_n\}$ is uniformly distributed in X, then the sequence $\{(S^{-1}x_n)\}^*$ [6] is uniformly distributed in the factor-group X/H.

It is sufficient to show [7] that for each closed set M^* in X/H whose boundary has measure zero the relation

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n \leqq N} \varphi_{M^*}\big((S^{-1}x_n)^*\big) = \lambda^*(M^*) \tag{3}$$

holds. Now let M be a measurable set in X. If $x \in M$, then $(S^{-1}x)^* \in (S^{-1}M)^*$ and conversely. By assumption we have $\lambda(M) = \lambda(S^{-1}M) = \lambda^*\big((S^{-1}M)^*\big)$. The last relation follows from the fact that $S^{-1}M$ contains with each element x the whole residue class xH.

Let now M be an arbitrary closed subset of X whose boundary is a null set. Then also the boundary of $(S^{-1}M)^*$ is a null set and we have

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} \sum_{n \leqq N} \varphi_{(S^{-1}M)^*}\big((S^{-1}x_n)^*\big) &= \lim_{N\to\infty} \frac{1}{N} \sum_{n \leqq N} \varphi_M(x_n) \\ &= \lambda(M) = \lambda(S^{-1}M) = \lambda^*\big((S^{-1}M)^*\big).\end{aligned}$$

Now each set $M^* \subseteqq X/H$ may be represented in the form $M^* = (S^{-1}M_1)^*$ with some $M_1 \subseteqq X$; take for example $M_1 = SM^*$. This proves (3). For $y \in S^{-1}x$, it is obvious that, $T^n y \in M$ implies $(S^{-1}T^n x)^* \in M = S^{-1}SM$. Therefore we have for each closed set M whose boundary is a null set

$$\begin{aligned}\overline{\lim_{N\to\infty}} \frac{1}{N} \sum_{n \leqq N} \varphi_M(T^n y) &\leqq \lim_{N\to\infty} \frac{1}{N} \sum_{n \leqq N} \varphi_{(S^{-1}SM)^*}\big((S^{-1}T^n x)^*\big) \\ &= \lambda^*\big((S^{-1}SM)^*\big) = \lambda(S^{-1}SM) \leqq k\lambda(M).\end{aligned}$$

[5] Special cases are considered in [16], [13], [12].

[6] x^* denotes the image of x under the natural homomorphism from X to X/H.

[7] Cf. [5], § 12. In the following formula λ^* denotes Haar measure on X/H and φ_M the characteristic function of the set M.

The last inequality follows from the fact that $M \subseteqq S^{-1}SM \subseteqq \bigcup_{h\in H} Mh$. Therefore each limit measure of the sequence $\{T^n y\}$ is invariant with respect to T and absolutely continuous with respect to λ. This implies that the sequence $\{T^n y\}$ is uniformly distributed in X, q.e.d.

A particular case is the well-known result (cf. [10]) that if x is normal to base a, then also rx is normal to base a, where $r \neq 0$ denotes an arbitrary rational number. Another consequence is that the uniform distribution of the sequence $\{A^n \mathfrak{x}\}$ [8] implies the uniform distribution of the sequence $\{A^n B^{-1}\mathfrak{x}\}$ for each non-singular integral matrix B which commutes with A.

We now consider another situation:

Let X be a compact Hausdorff space with countable base. Let X_ω denote the product space $X_\omega = \prod_{i=1}^\infty X_i$, where all $X_i = X$. X_ω may be identified with the space of all sequences $\omega = \{x_n\}$ with $x_n \in X$. In the usual product topology X_ω is again a compact Hausdorff space with countable base. A set $C \subseteqq X_\omega$ of the form $C = \{\omega = \{x_n\} : x_{k_1} \in A_1, \ldots, x_{k_l} \in A_l\}$, where $k_1, \ldots, k_l$ are arbitrary positive integers and $A_1, \ldots, A_l$ are compact subsets of X, is called a cylinder set. For a measure μ in X let μ_ω denote the associated product measure in X, i.e. μ_ω is defined on the least sigma-algebra containing all cylinder sets and $\mu_\omega(C) = \mu(A_1) \ldots \mu(A_l)$. In X_ω we define a transformation T, called the shift transformation by $T\omega = T\{x_n\} = \{x_{n+1}\}$. It is easily checked that T is measure preserving and ergodic with respect to each product measure μ_ω. Therefore μ_ω-almost all sequences $\{x_n\}$ satisfy

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n\leqq N} f(T^n\omega) = \mu_\omega(f) \tag{4}$$

for every $f \in C(X_\omega)$.

If a sequence ω fulfills (4) it is called completely distributed with respect to the measure μ.

Of course our general theorems are applicable also in this case. For instance, if (2) holds on a set of nonnegative functions f which is uniformly dense in $C(X)$, then the sequence $\omega = \{x_n\}$ is completely distributed with respect to μ. [9].

The most important special case is $X = [0, 1]$ and $\mu =$ Lebesgue measure. It is easy to see that the sequence $\{x_n\}$ is completely

[8] Here A denotes as above an integral matrix with det $A \neq 0$, which has no root of unity as eigenvalue.

[9] In the special case $X = [0, 1]$ this has been proved by Postnikow [14].

uniformly distributed if and only if for each $k \geqq 1$ the sequences $\{(x_n, x_{n+1}, \ldots, x_{n+k-1})\}$ are uniformly distributed in the k-dimensional unit cube.

Let now $x = \sum x_n/a^n$ be the representation of x to base $a > 1$ and let x be normal to this base. Then it is well known that the sequence $\{x_n\}$ (which obviously belongs to ΠX_i, where each $X_i = X = \{0, 1, \ldots, a-1\}$) is completely distributed with respect to the measure μ on X defined by $\mu(i) = 1/a$ $(i = 0, 1, \ldots, a-1)$ and conversely. This shows the close relation between the concepts of normal number and complete distribution of a sequence. This connection has been studied in greater detail by E. Hlawka [7].

It is perhaps interesting to note that if $a > 1$ is a transcendental number, then for almost all x the sequence $\{a^n x\}$ is completely uniformly distributed mod 1 [10].

The last theorem cannot be obtained with the aid of the individual ergodic theorem, because for nonintegral a the expression $a^n x$ is not representable in the form $T^n x$, where T denotes a mapping of the unit interval into itself. It is of course an important question to consider the transformation $Tx = \vartheta x - [\vartheta x]$, where $\vartheta > 1$ is not an integer and to study the asymptotic behavior of the sequence $\{T^n x\}$. This has first been done by A. Rényi [17], who proved that for almost all x the sequence $T^n x$ has a distribution measure. The explicit form of this distribution measure has been obtained independently by Gelfond [6] and Parry [11]. Another proof is given in [4]. This distribution measure is absolutely continuous with respect to Lebesgue measure in [0, 1] and its density $\sigma(t)$ is given by

$$\sigma(t) = \frac{1}{\tau} \sum_{t < T^k(1)} \frac{1}{\vartheta^k}, \quad \text{where} \quad \tau = \sum_0^\infty \frac{T^k(1)}{\vartheta^k}.$$

I might conclude by stating the famous theorem of Gausz and Kusmin: Let $x = [x_1, x_2, \ldots, x_n, \ldots]$ be the continued fraction expansion of the number $x \in [0, 1]$. Let $Tx = 1/x - [1/x]$, i.e. $Tx = [x_2, x_3, \ldots]$ Then for almost all x the sequence $\{T^n x\}$ has a distribution measure which is absolutely continuous with respect to Lebesgue measure and whose density $\rho(x)$ is given by $1/\log 2 \cdot 1/1+x$.

A proof of this theorem belonging to ergodic theory has been given by C. Ryll-Nardzewski [19].

[10] This follows in an easy way from a general theorem of Koksma [8], cf. also [2].

BIBLIOGRAPHY

J. R. Blum and D. L. Hanson

[1] On invariant probability measures I, Pacific J. Math. 10, 1125–1129 (1960).

J. Cigler

[2] Der individuelle Ergodensatz in der Theorie der Gleichverteilung mod 1, J. reine angew. Math. 205, 91–100 (1960).

J. Cigler

[3] Ein gruppentheoretisches Analogon zum Begriff der normalen Zahl, J. reine angew. Math. 206, 5–8 (1961).

J. Cigler

[4] Ziffernverteilung in ϑ-adischen Brüchen, Math. Zeitschr. 75, 8–13 (1961).

J. Cigler und G. Helmberg

[5] Neuere Entwicklungen der Theorie der Gleichverteilung, Jahresber. DMV 64, 1–50 (1961).

A. O. Gelfond

[6] Über eine allgemeine Eigenschaft von Zahlsystemen, (russ.), Izvest. Akad. Nauk SSSR, Ser. mat. 23, 809–814 (1959).

E. Hlawka

[7] Folgen auf kompakten Räumen II, Math. Nachr. 18, 188–202 (1958).

J. F. Koksma

[8] Ein mengentheoretischer Satz über die Gleichverteilung mod. Eins, Compositio math. 2, 250–258 (1935).

J. E. Maxfield

[9] Normal k-tuples, Pacific J. Math. 3, 189–196 (1953).

I. Niven

[10] Irrational Numbers, New York 1956.

W. Parry

[11] On the β-expansion of real numbers, Acta math.(hung.) 11, 401–416 (1960).

A. M. Polosuev

[12] On the uniformity of the distribution of a system of functions forming the solution of a system of linear finite-difference equations of first order. Dokl. Akad. Nauk SSSR 123, 405–406 (1958).

A. G. Postnikow

[13] A criterion for the uniform distribution of an exponential function in the complex domain, Vestn. Leningr. Univ. 13, 81–88 (1957).

A. G. Postnikow

[14] A criterion for a completely uniformly distributed sequence, Dokl. Akad. Nauk SSSR 120, 973–975 (1958).

A. G. Postnikow and I. I. Pyateckii

[15] Markov-sequence of symbols and continued fractions, Izvest. Akad. Nauk SSSR 21, 729–746 (1957).

I. I. Pyateckii

[16] On the distribution functions of the fractional parts of an exponential function, Izvest. Akad. Nauk SSSR 15, 47–52 (1951).

A. Rényi

[17] Representations for real numbers and their ergodic properties, Acta math. (hung.) 8, 477–493 (1957).

W. A. ROCHLIN
[18] On endomorphisms of compact Abelian groups, Izvest. Akad. Nauk SSSR 13, 329–340 (1949).

C. RYLL-NARDZEWSKI
[19] On the ergodic theorems II, Studia math. 12, 74–79 (1951).

(Oblatum 29-5-63)