A NON HOMOGENEOUS ERGODIC THEOREM

BY

PAUL R. HALMOS[1]

1. **Introduction.** Let X be a measure space with measure m such that $m(X)=1$ and let T be a one-to-one measure preserving transformation of X onto itself. My purpose in this note is to make a remark on the convergence of series such as

$$(*) \qquad \sum_{n=1}^{\infty} \frac{1}{n} f(T^n x).$$

Such series are of interest for several reasons. (1) In a paper with the same title as that of this one Izumi[2] has explicitly raised the problem of the convergence of (*) and offered a proof that, under certain rather restrictive conditions on f and T, (*) converges almost everywhere. (2) Very elegant necessary and sufficient conditions for the convergence of series similar to (*) are known in the theory of probability[3]—an extension of those results would be of significance in the study of asymptotic properties of a more general class of transformations. (3) Since it is an elementary fact[4] that the convergence of the numerical series $\sum_{n=1}^{\infty}(1/n)x_n$ implies that $\lim_{n\to\infty}(1/n)\sum_{i=1}^{n} x_i=0$, the almost everywhere convergence of (*) would be an analytic strengthening of Birkhoff's ergodic theorem.

The principal result of this paper (stated precisely in §3) is that in general (*) does not converge in the mean (of order two).

2. **Izumi's theorem.** Izumi assumes that the transformation T is uniformly mixing in the sense that

$$(2.1) \qquad m(E \cap T^n F) = m(E)m(F) + O(\log^{-3} |n|)$$

uniformly for all measurable sets E and F for which $E \subset F$. (In Izumi's paper "o" appears in place of "O" but only the latter, less restrictive condition is used in his proof.) If the quantitative aspect of (2.1) is ignored, it asserts that

$$(2.2) \qquad m(E \cap T^n F) \to m(E)m(F) \qquad \text{uniformly for } E \subset F,$$

Presented to the Society, September 10, 1948; received by the editors April 5, 1948.

[1] John Simon Guggenheim Memorial Fellow.

[2] Shin-Ichi Izumi, *A non homogeneous ergodic theorem*, Proc. Imp. Acad. Tokyo vol. 15 (1939) pp. 189–192.

[3] See for instance E. R. van Kampen, *Infinite product measures and infinite convolutions*, Amer. J. Math. vol. 62 (1940) pp. 417–448.

[4] Ingeniously exploited by Mark Kac in his proof of the strong law of large numbers, *Sur les fonctions independantes* (I), Studia Math. vol. 6 (1936) p. 54. Cf. also van Kampen, op. cit. p. 426.

or, equivalently,

(2.3) $$m((E \cap F) \cap T^n F) \to m(E \cap F)m(F) \qquad \text{uniformly in } E \text{ and } F.$$

If (2.2) held without the restriction $E \subset F$, then for every $\delta > 0$ there would exist a positive integer $n_0(\delta)$ such that if $n \geqq n_0$ is a positive integer and E and F are arbitrary measurable sets, then

$$|m(E \cap T^n F) - m(E)m(F)| < \delta.$$

Choosing $n = n_0$, $F = T^{-n_0}E$, this implies that

$$|m(E) - (m(E))^2| < \delta.$$

Since δ is arbitrary, it follows that *the value of $m(E)$ is either 0 or 1 for every measurable set E—in other words that X is essentially isomorphic, in the sense of measure theory, with a space containing exactly one point.*

It is easy to show that if (2.2) holds, then so does the strengthened condition obtained from it by removing the restriction $E \subset F$. In fact, to every $\delta > 0$ there corresponds a positive integer $n_0(\delta)$ such that if $n \geqq n_0$ is a positive integer and E and F are arbitrary measurable sets, then

(2.4) $$|m((E \cap F) \cap T^n F) - m(E \cap F)m(F)| < \delta/2.$$

Replacing F by F' (= the complement of F) yields

(2.5) $$|m((E \cap F') \cap T^n F') - m(E \cap F')m(F')| < \delta/2.$$

Since, however,

$$E \cap F' = [(E \cap F') \cap T^n F] \cup [(E \cap F') \cap T^n F'],$$

and the terms of the last written union are disjoint, it follows that

$$m((E \cap F') \cap T^n F') = m(E \cap F') - m((E \cap F') \cap T^n F).$$

Hence (2.5) asserts that

(2.6) $$|m(E \cap F') - m((E \cap F') \cap T^n F) - m(E \cap F')m(F')| < \delta/2.$$

But $m(E \cap F') - m(E \cap F')m(F') = m(E \cap F')m(F)$, so that (2.6) becomes

(2.7) $$|m((E \cap F') \cap T^n F) - m(E \cap F')m(F)| < \delta/2.$$

(In other words (2.2) is valid for disjoint sets E and F.) Adding (2.4) and (2.7) yields

(2.8) $$|m(E \cap T^n F) - m(E)m(F)| < \delta,$$

and this implies (2.2) with the restriction $E \subset F$ removed[5].

---

[5] The fact here communicated, concerning the overly restrictive nature of Izumi's hypothesis, was first proved during a discussion among Ambrose, Kakutani, and me, shortly after the appearance of Izumi's paper.

3. **Convergence in the mean.** In this section I assume that $f\in L_2(X)$ and, instead of studying the almost everywhere convergence of (*), I shall discuss the possibility of its convergence in the mean (of order two).

Let U be the unitary operator induced by T; in other words U is defined, for every $f\in L_2(X)$, by $(Uf)(x)=f(Tx)$. Write $\mathfrak{M}$ for the closed linear manifold of functions $g\in L_2(X)$ for which $g=Ug$, and $\mathfrak{N}$ for the orthogonal complement of $\mathfrak{M}$; then every $f\in L_2(X)$ may be written in the form $f=g+h$ with $g\in\mathfrak{M}$ and $h\in\mathfrak{N}$. Since, by the mean ergodic theorem, $(1/n)\sum_{i=1}^{n} U^if$ converges to g in the mean, and since the convergence of

$$\sum_{n=1}^{\infty}\frac{1}{n}U^nf \tag{3.1}$$

implies that $(1/n)\sum_{i=1}^{n} U^if$ approaches 0, it follows that (3.1) can converge only if $f\in\mathfrak{N}$.

There are in general many f's in $\mathfrak{N}$ for which (3.1) does converge. It converges, for instance, if f is a proper function of proper value $\lambda\neq 1$ (that is, $Uf=\lambda f$), or if the sequence of functions $\{U^nf\}$ is an orthogonal sequence, or if f belongs to the range of the operator $I-U$ (that is, f has the form $h-Uh$ for some $h\in\mathfrak{N}$). The proofs of these statements are straightforward (and in most cases yield also almost everywhere convergence). It is clear also that the set of f's for which (3.1) converges is a linear set invariant under U. Since, moreover, the range of $I-U$ is known to be everywhere dense in $\mathfrak{N}$(6), it is necessary to show only that the convergence set of (3.1) is closed in order to obtain the best possible result along these lines. In all interesting cases, however, just the opposite is true. I shall in fact prove that *if T is metrically transitive*(7) *and X is non atomic, then it is impossible that* (3.1) *converge for every $f\in\mathfrak{N}$.*

The operator U may be represented in the form $U=e^{2\pi iH}$, where H is a Hermitian operator with spectral family $\{E(\lambda)\}$, that is, for every $f, g\in\mathfrak{N}$

$$(Uf, g)=\int_0^1 e^{2\pi i\lambda}d(E(\lambda)f, g).$$

For $0<\lambda<1$, write

$$s_N(\lambda)=\sum_{n=1}^{N}\frac{1}{n}e^{2\pi in\lambda}\quad\text{and}\quad s(\lambda)=\lim_{N\to\infty}s_N(\lambda).$$

If (3.1) converges for every $f\in\mathfrak{N}$ then

$$\left\|\sum_{n=M+1}^{N}\frac{1}{n}U^nf\right\|^2=\int_0^1 |s_M(\lambda)-s_N(\lambda)|^2d\|E(\lambda)f\|^2$$

(6) See E. Hopf, *Ergodentheorie*, Ergebnisse der Mathematik, vol. 5, no. 2, Berlin, 1937, p. 23.

(7) See Hopf, op. cit. p. 29.

approaches 0, that is, the sequence of functions $\{s_N\}$ converges (in the mean of order two with respect to the distribution $\|E(\lambda)f\|^2$) to a function s_f. Since a subsequence must then converge to s_f almost everywhere, it follows that $s_f(\lambda)=s(\lambda)$ almost everywhere (with respect to $\|E(\lambda)f\|^2$). Hence in this case $s(H)f$ is defined for every $f\in\mathfrak{N}$ and therefore[8] there exists a positive constant $c<\infty$ such that $|s(\lambda)|\leqq c$ almost everywhere with respect to the spectral measure of H, that is, the set of those values of λ for which $|s(\lambda)|>c$ has measure zero with respect to every distribution of the form $\|E(\lambda)f\|^2$, $f\in\mathfrak{N}$. Observe that since $s(\lambda)$ is equal to a branch of log $(1-e^{2\pi i\lambda})$, the set

$$A = \{\lambda: |s(\lambda)| > c\}$$

consists of two intervals, symmetrically located at the two extremes of the unit interval. A contradiction to the assumption of universal convergence will be obtained by showing that corresponding to any such set A an $f\in\mathfrak{N}$ can be found so that $\mu_f(A)$ (that is, the measure of A with respect to the distribution $\|E(\lambda)f\|^2$) is positive. For this purpose it is more convenient to think of a two-interval set of the kind described as represented in the form $A=\{\lambda: |1-e^{2\pi i\lambda}|<2\delta\}$.

Let δ be any positive number. I have proved earlier[9] that there exists on X a (pointwise) periodic measure preserving transformation S such that if the unitary operator induced by S on $\mathfrak{N}$ is denoted by V, then $\|Vf-Uf\|<\delta\|f\|$ for all $f\in\mathfrak{N}$. The periodicity of S implies that there exists a measurable function f which takes the values $+1$ and -1 each on a measurable set of measure $1/2$ and is invariant under S. (To construct such an f in case S has everywhere the same period p, let E be a measurable set of measure $1/p$ which is disjoint from its first $p-1$ images under S; let f on E take the values $+1$ and -1 each on a set of measure $m(E)/2$, and extend the domain of definition of f to X by the requirement of S-invariance. The nonatomicity of X is necessary in this argument to ensure the existence of sets of measure $m(E)/2$. If S does not have a constant period on X, then X may be split into possibly countably many sets on each of which the period of S is constant and the argument applies.) It is clear that $\int_X f(x)dm(x)=0$ (so that $f\in\mathfrak{N}$) and $\int_X |f(x)|^2dm(x)=\|f\|^2=1$. The V-invariance of f implies that

$$\|f - Uf\| \leqq \|f - Vf\| + \|Vf - Uf\| < \delta.$$

If $A=\{\lambda: |1-e^{2\pi i\lambda}|<2\delta\}$ and A' is the complement of A in the unit interval,

$$\delta^2 > \|f - Uf\|^2 = \int_0^1 |1 - e^{2\pi i\lambda}|^2 d\|E(\lambda)f\|^2 \geqq \mu_f(A')\cdot 4\delta^2,$$

(8) See M. H. Stone, *Linear transformations in Hilbert space*, Amer. Math. Soc. Colloquium Publications, vol. 15, 1922, Theorem 6.5, p. 220.

(9) *Approximation theories for measure preserving transformations*, Trans. Amer. Math. Soc. vol. 55 (1944) p. 15, Theorem 8. Observe that, in view of Theorem 10, the metric and the uniform topologies for measure preserving transformations are identical.

and therefore

$$\mu_f(A) = \int_0^1 d\|E(\lambda)f\|^2 - \mu_f(A') \geqq \|f\|^2 - \frac{1}{4} = \frac{3}{4}\,.$$

This concludes the proof of the (italicized) principal assertion of this section. It is possibly of interest to remark in connection with the last part of the proof that it was shown that *even after the removal of the trivial proper value* 1 (*that is, even after the restriction of* f *to* $\mathfrak{N}$), *the hypotheses on* T *imply that* 1 *belongs to the spectrum of* U(10).

It should be pointed out also that these methods do not directly yield any answer to the problem of almost everywhere convergence (in place of mean convergence). While it is unlikely that there should exist a transformation T satisfying the conditions described above and such that (*) converges almost everywhere for every f for which $\int_X f(x)dm(x) = 0$, a proof of this fact would be of interest. A more delicate investigation would be needed to discover the precise conditions on f under which the series converges in some preassigned sense(11).

University of Chicago
Chicago, Ill.

(10) For a related result see J. von Neumann, *Zur Operatorenmethode in der klassischen Mechanik*, Ann. of Math. vol. 33 (1932) p. 635.

(11) An interesting special case was recently discussed by M. Kac, R. Salem, and A. Zygmund, *A gap theorem*, Trans. Amer. Math. Soc. vol. 63 (1948) pp. 235–243.