Can. J. Math., Vol. XXXV, No. 2, 1983, pp. 339–352

MIXING ON SEQUENCES

NATHANIEL A. FRIEDMAN

1. Introduction. Our aim is to study the mixing sequences of a weak mixing transformation. An ergodic measure preserving transformation is weak mixing if and only if for each pair of sets there exists a sequence of density one on which the transformation mixes the sets [**9**]. An unpublished result of S. Kakutani implies there actually exists a single sequence of density one on which the transformation is mixing for all sets (see Section 3). This result motivated the general definition of a transformation being mixing on a sequence, as well as mixing of higher order on a sequence. Given a weak mixing transformation, there exist sequences along which it is mixing of all degrees. In particular, this is the case for an eventually independent sequence [**7**].

In Section 3 it will be shown that if T is weak mixing but not mixing, then a sequence on which T is two-mixing must have upper density zero. Thus in this case T is mixing on a sequence of density one but T cannot be two-mixing on a sequence of positive density.

In Section 4 we will study the Mean Ergodic Theorem (M.E.T.) for Césaro-averages along a mixing sequence. The Blum-Hansen Theorem [**1**] states that a transformation is mixing if and only if the M.E.T. holds along any sequence. It was proven by L. Jones [**10**] that the M.E.T. holds for any sequence of positive lower density when the transformation is weak mixing. An example will be given of a weak mixing transformation T that is mixing on a certain sequence but the M.E.T. does not hold for T on that sequence. An inspection of the proof in [**1**] shows that the M.E.T. is equivalent to a condition referred to in Section 4 as Césaro uniform mixing. In particular, this implies the M.E.T. holds along each sequence on which the transformation is two-mixing.

In Section 5 we will first verify a uniform version of the Blum-Hansen Theorem which states that if T is mixing, then the Césaro average of any n iterates $T^{k_i}f$, $1 \leqq i \leqq n$, is close to the integral of f if n is large. Here n depends only on f and the closeness is uniform for all choices of k_i, $1 \leqq i \leqq n$. A corollary is that if T is mixing and A is a set of positive measure, then the union of any n iterates of A has measure close to 1 for n sufficiently large. However, if T is mixing on a sequence, then this property can fail for iterates chosen on the sequence. The property holds on a two-mixing sequence.

Received March 16, 1982 and in revised form October 4, 1982. This work was partially supported by N.S.F. Grant No. MCS-8102101.

I would like to thank A. Hajian for organizing the seminar at Northeastern University in May, 1981, where the ideas for this paper began. I would also like to thank S. Kalikow for several conversations that eventually led me to inspect the proof of the Blum-Hansen Theorem and the resulting uniform version in Section 5.

2. Preliminaries. Let $(X, \mathscr{B}, m)$ be a measure space isomorphic to the unit interval with Lebesgue measure. An invertible transformation T defined on X is *weak mixing* if

$$(2.1)\quad \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} |m(T^k A \cap B) - m(A)m(B)| = 0, \quad A, B \in \mathscr{B}.$$

The transformation T is *mixing* if

$$(2.2)\quad \lim_{n\to\infty} m(T^n A \cap B) = m(A)m(B), \quad A, B \in \mathscr{B}.$$

We will only consider transformations that are weak mixing but not mixing.

An increasing sequence of positive integers will be denoted by $s = (s_i)$ or $s = (k : k \in s)$. A limit along s will be denoted by $\lim_{k\in s}$. A transformation T is *mixing on* s if

$$(2.3)\quad \lim_{k\in s} m(T^k A \cap B) = m(A)m(B), \quad A, B \in \mathscr{B}.$$

For each positive integer n, let $n(s)$ be the number of terms in s not exceeding n. Define $D^*(s)$ and $D_*(s)$ as

$$D^*(s) = \lim_{n\to\infty} \sup n(s)/n,$$

$$D_*(s) = \lim_{n\to\infty} \inf n(s)/n.$$

If $D^*(s) = D_*(s) = D$, then s has *density* $D(s) = D$. The following result is proved in [**9**].

(2.4) THEOREM. *A transformation T is weak mixing if and only if for each pair of sets A, $B \in \mathscr{B}$ there exists $s = s(A, B)$ with $D(s) = 1$ and*

$$\lim_{k\in s} m(T^k A \cap B) = m(A)m(B).$$

Since $(\mathscr{B}, m)$ is separable, one can use (2.4) and a diagonalization argument to prove there exists a sequence s on which T is mixing. Moreover, if T is mixing on a sequence, then T must be weak mixing. Thus T is weak mixing if and only if T is mixing on s for some s. In Section 3 it will be proved that s can be chosen to also satisfy $D(s) = 1$.

The sequences on which a transformation is mixing are isomorphism invariants and can be used to distinguish certain weak mixing transformations. For example, given any increasing sequence s, one can construct a weak mixing transformation that is not mixing on s [**7**]. Thus if T_1 is mixing on s, then there exists T_2 not mixing on s. In particular, there does not exist a universal mixing sequence.

The method of independent cutting and stacking [**5**, **6**, **13**] will be used to construct examples. A brief description follows. The construction takes place on the unit interval $[0, 1)$ and all intervals considered will be left-closed and right-open. A *column* C of height h is an ordered set of disjoint intervals I_i, $1 \leqq i \leqq h$, that have the same length. The *base of* C is $\underline{C} = I_1$, the *top of* C is $\bar{C} = I_h$, the *width of* C is $w(C) = m(I_1)$, and the *height of* C is $h(C) = h$. We also let C denote the union of the intervals in C, which we refer to as *levels in* C. A column C can be pictured as the rungs on a ladder with I_i above I_{i-1}, $1 < i \leqq h$.

The corresponding transformation T_C maps I_{i-1} onto I_i by a translation, $1 < i \leqq h$. Thus T_C is defined on $C - \bar{C}$.

A *tower* G is an ordered set of disjoint columns. The *top of* G is the union of the tops of the columns in G, denoted by $\bar{G}$. The *base of* G is the union of the bases of the columns in G, denoted by $\underline{G}$. The *width of* G is $w(G) = m(\underline{G}) = m(\bar{G})$. The transformation T_G consists of T_C acting on C in G. A level in a column in G is simply called a level in G. We also let G denote the union of levels in G. Thus T_G is defined on $G - \bar{G}$.

Let C be a column of height h with base I. Let J be a subinterval of I. We refer to $C_J = (T_C{}^i J : 0 \leqq i < h)$ as a *subcolumn* of C. Let

$$p = m(J)/m(I) \leqq 1.$$

We also refer to C_J as a *p-copy of* C and denote $C_J = pC$.

Given a tower $G = (C_j : 1 \leqq j \leqq k)$, denote a p-copy of G as

$$pG = (pC_j : 1 \leqq j \leqq k).$$

Let $p_j = w(C_j)/w(G)$, $1 \leqq j \leqq k$; hence $(p_1 + \ldots + p_h) = 1$. Cut G into disjoint copies $G_0 = .5G$ and $.5G$. Cut the latter $.5G$ into k disjoint copies $G_j = p_j(.5G)$, $1 \leqq j \leqq k$; hence $w(G_j) = .5w(C_j)$, $1 \leqq j \leqq k$. Thus the width of G_j is the same as the width of the jth column in G_0, $1 \leqq j \leqq k$.

Form the tower SG obtained by placing G_j above the jth column $.5C_j$ in G_0, $1 \leqq j \leqq k$. The tower SG has k columns above each column in G_0; hence SG has k^2 columns. The width of SG is $w(SG) = w(G_0) = w(G)/2$. Note that T_{SG} extends T_G to a set of measure $m(\bar{G}_0) = w(G)/2$, by mapping $\overline{p_j C_j}$ onto $\underline{G}_j$, $1 \leqq j \leqq k$.

We refer to SG as the tower obtained by *independent cutting and stacking of* G. Let $S^k G = S(S^{k-1}G)$ and $T_k = T_{S^k G}$, $k \geqq 1$, As a set, $G_k = G$ so T_k is defined on $G - \bar{G}_k$, where $m(\bar{G}_k) = w(G)/2^k$, $k \geqq 1$.

If x is in a level in G, then $T_k(x)$ will be defined for k sufficiently large. Thus a transformation $T(G)$ can be defined on G as

$$(2.5) \quad T(G)(x) = \lim_{k\to\infty} T_k(x).$$

A tower G is an *M-tower* if two columns in G have heights that are mutually prime. In particular, G is an M-tower if two heights differ by one. If G is an M-tower, then $T(G)$ is mixing [**5**]. Moreover, $T(G)$ is a mixing Markov shift and isomorphic to a Bernoulli shift [**6, 13**]. Mixing implies that given $\epsilon > 0$, there exists a positive integer $N(G, \epsilon)$ such that

$$(2.6) \quad |m(T(G)^n I \cap J) - m(I)m(J)/m(G)| < \epsilon, \quad n \geqq N(G, \epsilon),$$

where I and J are levels in G.

Fix $N \geqq N(G, \epsilon)$. By (2.5) we can choose $k = k(G, \epsilon, N)$ so large that if T extends T_k, then (2.6) implies

$$(2.7) \quad |m(T^n I \cap J) - m(I)m(J)/m(G)| < \epsilon, \quad N(G, \epsilon) \leqq n \leqq N,$$

where I and J are levels in G.

Let G be a tower with columns with rational widths. Using the greatest common divisor, we can cut the columns in G into subcolumns all of the same width w. These sub-columns are now stacked consecutively to form one column of width w that we denote by $C(G)$. Note that if I is a level in G, then I appears as a finite union of levels in $C(G)$.

Let C be a column of height h and width w. Let u be a positive integer. The column C can be cut into u subcolumns of equal width w/u. These u subcolumns are stacked consecutively to form a single column denoted by $S_u C$, with height uh and width w/u.

Let $\epsilon > 0$ and t a positive integer. Choose $u \geqq \epsilon/t$ and let T be any extension of $T_{S_u C}$. If J is a level in C, then the construction of $S_u C$ implies

$$(2.8) \quad m\left(\bigcap_{j=0}^{t} T^{jh} J\right) \geqq (1 - \epsilon)m(J).$$

3. Sequences. A sequence s_1 *eventually contains* a sequence s_2 if all but a finite number of terms in s_2 are in s_1. The union of a countable set of sequences of density zero can have positive density. However, the following unpublished result of S. Kakutani [**11**] states that there exists a sequence of density zero that eventually contains each sequence of density zero in the countable set. A proof is included for completeness.

(3.1) THEOREM. *Let $D(s^n) = 0$, $n \geqq 1$. There exists s with $D(s) = 0$ such that s eventually contains s^n, $n \geqq 1$.*

Proof. Let $s^n = (s_j^n)$ and $\epsilon_n = 1/n^2$, $n \geqq 1$. Given $s = (s_j)$, let

$$(1) \quad d^*(s, u) = \lim_{k\to\infty} \sup V_k/k,$$

where V_k is the number of terms s_j that do not exceed k for $j \geqq u$. For $n \geqq 1$, $d^*(s^n, u)$ decreases to 0 as $u \to \infty$.

Choose u_1 such that

$$(2) \quad d^*(s^1, u_1) < \epsilon_1.$$

Assume $u_1 < u_2 < \ldots < u_r$ have been chosen so that for $1 \leqq v \leqq r$,

(3) $\quad d^*(s^n, u_v) < \epsilon_v, \quad 1 \leqq n \leqq v.$

Choose $u_{r+1} > u_r$ such that

(4) $\quad d^*(s^n, u_{r+1}) < \epsilon_{r+1}, \quad 1 \leqq n \leqq r + 1.$

By induction we obtain an increasing sequence (u_v) satisfying (3) for $v \geqq 1$. Now form s as the union of s_j^n for $j \geqq u_n$, $n \geqq 1$. Since $n\epsilon_n \to 0$, it follows that $D(s) = 0$.

(3.2) COROLLARY. *A transformation is weak mixing if and only if it is mixing on a sequence of density one.*

Proof. Since $(X, \mathscr{B}, m)$ is separable there exists a sequence of pairs (A_k, B_k) that are dense in the sense that for any pair (A, B) we have

(1) $\quad \lim_{k\to\infty} \inf\, (m(A \Delta A_k) + m(B \Delta B_k)) = 0,$

where Δ denotes the symmetric difference. We can choose a sequence s^k with $D(s^k) = 1$ such that Theorem (2.4) holds with $s(A_k, B_k) = s^k$, $k \geqq 1$. Let t^k be the complement of s^k in N. Apply Theorem (3.1) to obtain t with $D(t) = 0$ so that t eventually contains t^k, $k \geqq 1$. Let s be the complement of t. It follows that T is mixing on s and $D(s) = 1$.

We will now consider higher order mixing on a sequence s. A transformation T is *2-mixing on s* if $A, B, C \in \mathscr{B}$ imply

(3.3) $\quad \lim_{k,n\in s} m(T^n A \cap T^k B \cap C) = m(A)m(B)m(C),$

where $k \to \infty$ and $n - k \to \infty$. Since T is measure preserving, $C = X$ in (3.3) implies

(3.4) $\quad \lim_{k,n\in s} m(T^{n-k} A \cap B) = m(A)m(B).$

The reason that T may be mixing on s but not 2-mixing on s is that $k, n \in s$ does not imply $n - k \in s$. In particular, T may be mixing on s, but (3.4) may not hold. If (3.4) holds, then we will say T is *uniform mixing on s*, in the sense that $T^n A$ mixes into $T^k B$ uniformly with respect to $n - k$. Note that (3.4) may not imply (3.3). Otherwise one could prove mixing implies 2-mixing since a mixing transformation is uniform mixing on every sequence.

As in [**8**], a sequence s has *upper density* $U(s) = u$ if u is the largest number for which there exist $a_j \to \infty$, $b_j - a_j \to \infty$, and the number of terms in the sequence between a_j and b_j divided by $b_j - a_j$ converges to u as $j \to \infty$. Note that s may have $D(s) = 0$ but $U(s) = 1$ because s contains long blocks of consecutive integers with even longer gaps of consecutive integers in between. We will now prove that uniform mixing implies $U(s) = 0$.

(3.5) THEOREM. *If T is weak mixing but not mixing and T is uniform mixing on s, then $U(s) = 0$.*

Proof. Consider the set D of positive differences $p = n - k$ where $k, n \in s$. The set D can be written as $D = \{p_i : i \geqq 1\}$, where $p_i < p_{i+1}$, $i \geqq 1$. The *gaps* in D are $p_{i+1} - p_i$, $i \geqq 1$. Suppose the gaps are bounded by a positive integer g.

Since T is assumed weak mixing but not mixing, there exist A, B and $\epsilon > 0$ such that

(1) $\lim_{n\to\infty} \sup m(T^nA \cap B) \geqq m(A)m(B) + \epsilon.$

Thus there exist $r_j \to \infty$ such that

(2) $m(T^{r_j}A \cap B) \geqq m(A)m(B) + \epsilon, \quad j \geqq 1.$

For each r_j there exists t_j, $0 \leqq t_j \leqq g - 1$, such that $r_j + t_j \in D$. Since there are only g possible values for t_j, one value t must repeat infinitely often. Thus $r_j + t \in D$ for infinitely many j. Now

(3) $m(T^{r_j+t}A \cap T^tB) = m(T^{r_j}A \cap B).$

Let $v_j = r_j + t$ and $B_1 = T^tB$; hence (2) and (3) imply

(4) $m(T^{v_j}A \cap B_1) \geqq m(A)m(B_1) + \epsilon.$

Now $v_j \in D$; hence $v_j = n_j - k_j$ so (4) implies

(5) $m(T^{n_j}A \cap T^{k_j}B_1) \geqq m(A)m(B_1) + \epsilon.$

Now (5) contradicts uniform mixing. Thus T cannot be uniform mixing on s if $s - s$ has bounded gaps.

The proof is completed by a remark in [**8**] that states that if $s - s$ does not have bounded gaps, then s has upper density zero. A simple proof of this result, shown to me by B. Weiss, will be included for completeness. It suffices to verify that s has n mutually disjoint translates for $n \geqq 1$. The translate of s by k is the set $i + k$, $i \in s$, which is denoted by $s + k$. Note that s is disjoint from $s + k$ if and only if $k \notin s - s$.

Since $s - s$ has unbounded gaps, there exists a positive integer $k_1 \notin s - s$; hence $s \cap (s + k_1) = \emptyset$. Now choose a gap in $s - s$ starting at k_2 such that the gap exceeds k_1; hence $k_1 + k_2 \notin s - s$. Therefore s, $s + k_1$, and $s + k_1 + k_2$ are mutually disjoint. Note that $k_2 \notin s - s$ implies

$$(s + k_1) \cap (s + k_1 + k_2) = s \cap (s + k_2) = \emptyset.$$

Proceeding inductively, suppose k_1, $1 \leqq i \leqq n$, have been chosen so that

(6) $s, \quad s + \sum_{t=1}^{n} k_t, \quad 1 \leqq r \leqq n,$

are mutually disjoint. Choose a gap starting at k_{n+1} such that the gap size exceeds $\sum_{i=1}^{n} k_i$. It follows that (6) holds with n replaced by $n + 1$. Thus s has n mutually disjoint translates for $n \geqq 1$.

Since 2-mixing on s implies uniform mixing on s, Theorem (3.5) yields the following result.

(3.6) COROLLARY. *If T is weak mixing but not mixing and T is 2-mixing on s, then $U(s) = 0$.*

In Example (4.6) we will consider a case where T is mixing on s and $U(s) = 0$, but T is not 2-mixing on s.

In [8] Furstenberg defined a transformation T to be weak mixing of order r if $A_i \in \mathscr{B}$, $0 \leqq i \leqq r$, imply

$$(3.7)\quad \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left| m\left(\bigcap_{i=0}^{r} T^{ki}A_i\right) - \prod_{i=0}^{r} m(A_i)\right| = 0.$$

Furstenberg proved that weak mixing implied weak mixing of all orders. As in the case of weak mixing, one can use Theorem (3.1) and (3.7) to show there exists a sequence s with $D(s) = 1$ such that

$$(3.8)\quad \lim_{n\in s} m\left(\bigcap_{i=0}^{r} T^{ni}A_i\right) = \prod_{i=0}^{r} m(A_i), \quad A_i \in \mathscr{B}, \quad 0 \leqq i \leqq r.$$

Furthermore, another application of Theorem (3.1) yields a single sequence s with $D(s) = 1$ such that (3.8) holds for all $r \geqq 1$. In particular, for $r = 2$ we can rewrite (3.8) as

$$(3.9)\quad \lim_{n\in s} m(T^{2n}A \cap T^nB \cap C) = m(A)m(B)m(C), \quad A, B, C \in \mathscr{B}.$$

Thus (3.9) holds for $D(s) = 1$, in contrast to Corollary (3.6).

In Section 4 a transformation will be constructed that is mixing on a sequence s but is not uniform mixing on s. We have been unable to construct a transformation that is uniform mixing on a sequence s but is not 2-mixing on s. Another problem is to construct a transformation that is 2-mixing on a sequence s but is not r-mixing on s for some $r > 2$.

4. Mean convergence. We will now consider mean convergence of Césaro averages along a sequence $s = (s_i)$. Let $f \in L^p$, $p \geqq 1$, and denote

$$(4.1)\quad f_n(x) = \frac{1}{n} \sum_{i=1}^{n} f(T^{-s_i}x).$$

The following result [1] relates mixing and the mean convergence of f_n to the integral $m(f)$ of f with respect to m.

(4.2) BLUM-HANSEN THEOREM. *A transformation T is mixing if and only if for each sequence s, f_n converges to $m(f)$ in L^p, $f \in L^p$, $p \geqq 1$.*

Now suppose T is mixing on s. An example will be constructed to show that Césaro-averages along s need not converge in the mean. The idea of the example can be illustrated by a mixing sequence of sets. Let (A_n) be a sequence of sets with $m(A_n) = a$, $n \geqq 1$. The sequence is *mixing* [**12**] if

(4.3) $\lim_{n\to\infty} m(A_n \cap B) = am(B), \quad B \in \mathscr{B}.$

Let $\epsilon_n > 0$ and $\epsilon_n \to 0$. Let (t_n) be an increasing sequence of positive integers that satisfy

(4.4) $\sum_{i=1}^{n-1} t_i/t_n < \epsilon_n, \quad n > 1.$

Let (A_n) satisfy (4.3) with $a = 1/2$. Consider the sequence of sets (B_n) obtained by repeating A_n t_n times, $n \geqq 1$. This sequence will also be mixing. Let $b(n) = (t_1 + \ldots + t_n)$. The characteristic function of a set A will be denoted by $A(x)$. The Césaro-average of the first $b(n)$ characteristic functions of sets in (B_n) is denoted by $g_n(x)$; hence

(4.5) $g_n(x) = \sum_{i=1}^{n} t_i A_i(x)/b(n).$

Since $m(X) = 1$, it follows from (4.4) that

$$\|g_n - 1/2\|_1 \geqq 1/2 - \epsilon_n.$$

Thus g_n does not converge to $1/2$ in the mean.

We will now construct a transformation T, a corresponding mixing sequence s, and a set A of measure close to $1/2$ such that T^iA, $i \in s$, consists of blocks of length t_n that are approximately the same set, $n \geqq 1$.

(4.6) *Example*. The construction is by induction and the nth stage begins with an M-tower G_n with columns with rational widths. If I is a level in G_i, $1 \leqq i < n$, then I appears as a union of levels in G_n. Let L_n be the total number of levels in G_n and let $\epsilon_n < w_n/100L_n^2$. With reference to (2.6) and (2.7), let $N_n = N(G_n, \epsilon_n)$ and $k_n = k(G_n, \epsilon_n, N_n)$. Choose a positive integer

$$r_n \geqq \max \{k_n, N_n/\epsilon_n\}$$

and form $G_{n1} = S^{r_n}G_n$. We let T_{nj} denote $T_{G_{nj}}$ for notational convenience. Since $r_n \geqq k_n$, (2.7) implies that if T extends T_{n1}, then

(1) $|m(T^iI \cap J) - m(I)m(J)/m(G_n)| < \epsilon_n, \quad i = N_n,$

where I and J are levels in G_n.

Now form the column $G_{n2} = C(G_{n1})$. Each set A that is a union of levels in G_n will also appear as a union of levels in G_{n2}. Moreover, the choice of r_n implies $T^{N_n}A$ appears as a union of levels in G_{n2}, except possibly for a set of measure at most ϵ_n. This is because only the top N_n

levels in columns in G_{n1} pass through the top of G_{n1} under T^{N_n}. Thus we have

$$(2) \qquad T^{N_n}A = \bigcup_{i=1}^{v} J_i \cup E,$$

where J_i is a level in G_{n2}, $1 \leqq i \leqq v$, and $m(E) < \epsilon_n$.

We also have positive integers t_j, $1 \leqq j < n$, and choose t_n to satisfy (4.4); hence

$$(3) \qquad b(n-1)/t_n < \epsilon_n.$$

Now choose a positive integer $u_n \geqq t_n/\epsilon_n$. Form the column $G_{n3} = S_{u_n}G_{n2}$. Let h_n be the height of G_{n2}. If J is a level in G_{n2} and T extends T_{n3}, then (2.8) implies

$$(4) \qquad m\left(\bigcap_{j \leqq t_n} T^{jh_n} J\right) \geqq (1-\epsilon_n)m(J).$$

Let $s_j = N_n + (j - b(n-1))h_n$, $b(n-1) \leqq j < b(n)$. From (2) and (4) we obtain

$$(5) \qquad m\left(\bigcap_{j=b(n-1)}^{b(n)-1} T^{s_j}A\right) \geqq (1-2\epsilon_n)m(A).$$

Lastly, let G_{n+1} be the tower obtained by cutting G_{n3} into two equal columns and adding an extra interval above one column. Thus G_{n+1} is an M-tower consisting of two columns with heights differing by one. The levels in G_n appear as unions of levels in G_{n+1} and the columns in G_{n+1} have rational width. This completes the induction step.

We begin with an M-tower G_1 with columns of rational widths. Take $b(0) = 1$ in (3). At each stage we add an interval to form G_{n+1}. It is easy to see that the sum of the measures of these intervals is finite. Let $X = \bigcup_{n=1}^{\infty} G_n$ and assume m is normalized so that $m(X) = 1$. Thus we obtain a transformation T defined by

$$(6) \qquad T(x) = \lim_{n\to\infty} T_{G_n}(x), \quad x \in X.$$

We first verify T is mixing on $s = (s_j)$. Let A and B be sets that are unions of levels in G_1; hence A and B appear as levels in G_n, $n \geqq 1$. If n is large, then $m(G_n)$ is essentially 1 and (1) implies

$$(7) \qquad |m(T^{N_n}A \cap B) - m(A)m(B)| \leqq L_n^2\epsilon_n = w_n.$$

It is easily seen that $w_n \to 0$; hence

$$(8) \qquad \lim_{n\to\infty} m(T^{N_n}A \cap B) = m(A)m(B).$$

The same proof holds if A and B are unions of levels in G_k, $k \geqq 1$. Since these sets generate $\mathscr{B}$, it follows that T is mixing on (N_n). From (5) we conclude T is mixing on s.

To verify that the Mean Ergodic Theorem does not hold on s, choose k large and fix A consisting of a union of levels in G_k such that

(9) $\quad |m(A) - 1/2| < 1/100 \quad \text{and} \quad \epsilon_k < 1/100.$

Now (5) holds for $n > k$. Let $g_n(x)$ be as in (4.5) with $A_i = T^i A$, $i \geqq 1$. From (9) and (5) we obtain

$$\|g_n - m(A)\|_1 \geqq 1/8.$$

Thus the M.E.T. does not hold on s.

The preceding example shows that mixing on s does not imply the M.E.T. on s. An inspection of the proof in [**1**] yields the following mixing condition that is equivalent to the M.E.T. on s.

(4.7) *Definition.* A transformation T is *Césaro uniform mixing on* s if $A, B \in \mathscr{B}$ imply

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i,j=1}^{n} m(T^{s_i}A \cap T^{s_j}B) = m(A)m(B).$$

(4.8) THEOREM. *The Mean Ergodic Theorem holds for T on s if and only if T is Césaro uniform mixing on s.*

Proof. Let $f_n(x)$ be defined as in (4.1) with $f(x) = A(x)$. In L^2 we have

(1) $$\|f_n - m(A)\|_2^2 = \frac{1}{n^2} \sum_{i,j=1}^{n} m(T^{s_i}A \cap T^{s_j}A) - m(A)^2.$$

If T is Césaro uniform mixing, then (1) implies f_n converges to $m(A)$ in L^2, $A \in \mathscr{B}$. The M.E.T. now follows as in [**1**]. Conversely, suppose the M.E.T. holds. Let f_n be defined as above and let g_n replace f_n in (4.1) with $f(x) = B(x)$ for $B \in \mathscr{B}$. Thus f_n and g_n converge in L^2 to $m(A)$ and $m(B)$, respectively. Thus $f_n g_n$ converges to $m(A)m(B)$ in L^2. Hence $f_n g_n$ converges to $m(A)m(B)$ in L^1 and this yields Césaro uniform mixing on s.

The proof in [**1**] can be used to verify uniform mixing implies Césaro uniform mixing. Since 2-mixing on s implies uniform mixing on s, we have the following result.

(4.9) COROLLARY. *If T is 2-mixing on s, then the Mean Ergodic Theorem holds for T on s.*

The theorem of L. Jones [**10**] states that the M.E.T. holds on s for all weak mixing transformations when $D_*(s) > 0$. Thus Theorem (4.8) is useful only when $D_*(s) = 0$. In particular, this is the case in Corollary (4.9).

In Example (4.6), $h_n \to \infty$ implies $U(s) = 0$. Theorem (4.8) implies T is not Césaro uniform mixing on s. In particular, T is not 2-mixing on s. This also follows directly from (5).

5. Uniform sweeping out. Given an increasing sequence $s = (k_i)$, we say T *sweeps out on* s if $m(A) > 0$ implies

$$m\left(\bigcup_{i=1}^{\infty} T^{k_i}A\right) = 1.$$

If T sweeps out on all s, then we simply say T *sweeps out*. If T is mixing, then T sweeps out. In [**2**] *sequence mixing* is the term used for sweeps out. To avoid confusion with mixing on a sequence, we will use the latter term. The following characterization is proved in [**2**].

(5.1) THEOREM. *A transformation T sweeps out if and only if*

$$\liminf_{n\to\infty} m(T^nA \cap B) > 0, \quad m(A)m(B) > 0.$$

If T sweeps out, then T is weak mixing [**3**]. Hence if T sweeps out, then T is mixing on a sequence of density one by Corollary (3.2). However, there exist weak mixing transformations that do not sweep out. There also exist transformations that sweep out that are not mixing [**4**]. We will now consider a uniform type of sweeping out defined as follows.

(5.2) *Definition. T sweeps out uniformly* if given a set A of positive measure and $\epsilon > 0$, there exists $N = N(A, \epsilon)$ such that $n \geqq N$ implies

$$m\left(\bigcup_{i=1}^{n} T^{k_i}A\right) > 1 - \epsilon \quad \text{for all } k_1 < k_2 < \ldots < k_n.$$

It is shown below that mixing implies uniform sweeping out. The following result is motivated by Lemma 1 [**1**].

(5.3) LEMMA. *Let T be mixing, $m(A) > 0$, and $\epsilon > 0$. There exists $N = N(A, \epsilon)$ such that $n \geqq N$ implies*

$$\frac{1}{n^2}\sum_{i,j=1}^{n} |m(T^{k_i}A \cap T^{k_j}A) - m(A)^2| < \epsilon,$$

for all $k_1 < k_2 < \ldots < k_n$.

Proof. Since T is mixing, we have

(1) $\quad \lim_{|u-v|\to\infty} m(T^uA \cap T^vA) = m(A)^2.$

Choose w so large that $|u - v| > w$ implies

(2) $\quad |m(T^uA \cap T^vA) - m(A)^2| < \epsilon/2.$

Choose $N > (4w + 2)/\epsilon$. Now consider k_i, $1 \leqq i \leqq n$, $n \geqq N$. For each i there are at most $2w + 1$ values of j such that $|k_i - k_j| \leqq w$. Since a term on the left of (2) is bounded by 1, we have

(3) $$\frac{1}{n^2}\sum_{i,j=1}^{n} |m(T^{k_i}A \cap T^{k_j}A) - m(A)^2| \leqq \frac{(2w+1)n}{n^2} + \frac{\epsilon}{2} < \epsilon.$$

Lemma (5.3) will now be used to obtain a uniform version of the Blum-Hansen Theorem. We denote

$$f_n(x) = \frac{1}{n}\sum_{i=1}^{n} f(T^{-k_i}x).$$

(5.4) THEOREM. *Let T be mixing, $p \geqq 1$, and $f \in L^p$. Given $\epsilon > 0$, there exists $N = N(f, \epsilon)$ such that $n \geqq N$ implies*

$$\|f_n - m(f)\|_p < \epsilon \quad \textit{for all } k_1 < \ldots < k_n.$$

Proof. If $f(x) = A(x)$, then Lemma (5.3) yields the result for $p = 2$ since (3) above with ϵ replaced by ϵ^2 implies

(1) $\quad \|f_n - m(A)\|_2 \leqq \epsilon.$

If f is a simple function of the form

(2) $\quad f(x) = \sum_{i=1}^{k} a_i A_i(x),$

then we have

(3) $\quad \|f_n - m(f)\|_2 \leqq \sum_{i=1}^{k} |a_i| \, \|f_{n,i} - m(A_i)\|_2.$

Here $f_{n,i}$ corresponds to $f = A_i$, $1 \leqq i \leqq k$. Choose

$$N = \max\{N(A_i, \epsilon/k|a_i|), \quad 1 \leqq i \leqq k\}.$$

Thus $n \geqq N$ implies the right side of (3) is less than ϵ. For $f \in L^2$, we approximate by a simple function g so that $\|f - g\|_2 < \epsilon/3$. Since T is measure preserving, we obtain

(4) $\quad \|f_n - g_n\|_2 < \epsilon/3, \quad n \geqq 1.$

Now choose $N = N(g, \epsilon/3)$; hence $n \geqq N$ implies

(5) $\quad \|f_n - m(f)\|_2 \leqq \|f_n - g_n\|_2 + \|g_n - m(g)\|_2 + |m(g) - m(f)| < \epsilon.$

If $p = 1$, then the result follows from Holders inequality and the result for $p = 2$. If $p > 1$, then as in [**1**], let g be bounded by M; hence

(6) $\quad \|g\|_p^p \leqq (1 + M^p)\|g\|_1.$

The result now follows from p from (6) and the result for $p = 1$ since simple functions are bounded.

(5.5) COROLLARY. *If T is mixing, then T sweeps out uniformly.*

Proof. Let $m(A) > 0$ and choose $N = N(A, \epsilon^2 m(A)^2)$ in Lemma (5.3). Let

$$B = \left(\bigcup_{i=1}^{n} T^{k_i}A\right)^c;$$

https://doi.org/10.4153/CJM-1983-019-2 Published online by Cambridge University Press

hence

$$B \cap T^{k_i}A = \emptyset, \quad 1 \leqq i \leqq n.$$

Let f_n correspond to $f = A$. Therefore Lemma (5.3) implies

$$m(A)m(B) = \left| \int_B (f_n(x) - m(A))dm \right| \leqq \|f_n - m(A)\|_2 < \epsilon m(A).$$

Thus $m(B) < \epsilon$.

Let us now consider the following version of Theorem (5.1) for a sequence s. The proof follows as in [**2**].

(5.6) THEOREM. *T sweeps out on all subsequences of s if and only if*

$$\liminf_{n \in s} m(T^nA \cap B) > 0, \quad m(A)m(B) > 0.$$

In particular, if T is mixing on s, then T sweeps out on all subsequences of s. However, mixing on s does not imply uniform sweeping out on s. For consider Example (4.6) (5). This implies $T^{s_j}A$ is essentially invariant for $b(n-1) \leqq j < b(n)$ and $b(n) - b(n-1) \to \infty$.

If T is uniform mixing on s, then the same proof of Lemma (5.3) yields the conclusion for $k_i \in s$, $1 \leqq i \leqq n$. In this case Theorem (5.4) holds for $k_i \in s$, $1 \leqq i \leqq n$. The analog of Corollary (5.5) also holds, where uniform sweeping out on s corresponds to Definition (5.2) with $k_i \in s$, $1 \leqq i \leqq n$. In particular, if T is 2-mixing on s, then there is uniform mean convergence on s and T sweeps out uniformly on s.

An open problem is whether the converse of Corollary (5.5) holds. There is also the question of whether sweeping out uniformly on s implies uniform mixing on s.

The following corollary of Theorem (5.4) states that given a set A and $\epsilon > 0$, there exists N such that for any set B, not more than N iterates of A can be badly mixed in B (with respect to ϵ). The original formulation of this result (and (5.8) below) is due to S. Kalikow, where $m(B)$ had to be bounded away from zero.

(5.7) COROLLARY. *Let T be mixing, $m(A) > 0$, and $\epsilon > 0$. There exists $N = N(A, \epsilon)$ such that for any set B there are at most N positive integers k such that*

$$|m(T^kA \cap B) - m(A)m(B)| > \epsilon.$$

Proof. Let $f(x) = A(x)$ with $p = 1$ in (5.4) and let $N_1 = N(f, \epsilon)$ in (5.4). Choose $N = 2N_1$ and suppose the conclusion does not hold. Hence there exist B and k_i, $1 \leqq i \leqq N_1$, such that

(1) $\quad m(T^{k_i}A \cap B) - m(A)m(B) > \epsilon \quad (\text{or} < -\epsilon), \quad 1 \leqq i \leqq N_1.$

Now (1) implies

$$(2) \qquad \epsilon < \frac{1}{N_1} \sum_{i=1}^{N_1} m(T^{k_i}A \cap B) - m(A)m(B)$$

$$= \int_B (f_{N_1}(x) - m(A))dm \leqq \|f_{N_1} - m(A)\|_1 < \epsilon.$$

This contradiction implies $N = 2N_1$ and hence the conclusion.

(5.8) COROLLARY. *Let T be mixing, $m(A) > 0$, and $\epsilon > 0$. There exists $N = N(A, \epsilon)$ such that for any sets B and C and j sufficiently large there are at most N positive integers k such that*

$$|m(T^kA \cap T^jB \cap C) - m(A)m(B)m(C)| > \epsilon.$$

Proof. Let $N = N(A, \epsilon/2)$ in (5.7). Choose j sufficiently large so that

$$(1) \qquad |m(T^jB \cap C) - m(B)m(C)| < \epsilon/2.$$

The conclusion follows from (5.7) with B replaced by $T^jB \cap C$.

Note that Corollary (5.8) is in the direction of mixing implying 2-mixing.

REFERENCES

1. J. R. Blum and D. L. Hanson, *On the mean ergodic theorem for subsequences*, Bull. A.M.S. *66* (1960), 308–311.
2. J. R. Blum, S. L. M. Christianson and D. Quiring, *Sequence mixing and α-mixing*, Illinois J. Math. *18* (1974).
3. J. W. England and N. F. G. Martin, *On weak mixing metric automorphisms*, Bull. A.M.S. *74* (1968), 505–507.
4. N. A. Friedman and D. S. Ornstein, *On partially mixing transformations*, Illinois J. Math. *16* (1972), 61–68.
5. N. A. Friedman, *Mixing, entropy, and generators*, J. Appl. Math. Anal. *26* (1969), 512–528.
6. ——— *Bernoulli shifts induce Bernoulli shifts*, Advances in Math. *10* (1973), 39–48.
7. ——— *Eventually independent sequences*, Israel J. Math. *33* (1979), 310–316.
8. H. Furstenberg, *Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic sequences*, J. Analyse Math. *31* (1977), 204–256.
9. P. R. Halmos, *Lectures on ergodic theory*, Publ. Math. Soc. Japan *3* (1965).
10. L. K. Jones, *A mean ergodic theorem for weakly mixing operators*, Advances in Math. *7* (1971), 211–216.
11. S. Kakutani, Ergodic Theory Seminar, Northeastern University (1981).
12. A. Rényi, *On mixing sequences of sets*, Actu. Math. Acad. Sci. Hungary *9* (1958), 215–228.
13. P. Shields, *Cutting and independent stacking of intervals*, Math. Systems Theory *7* (1973), 1–4.

State University of New York,
Albany, New York