WEAKLY WANDERING VECTORS AND
WEAKLY INDEPENDENT PARTITIONS([1])

BY
ULRICH KRENGEL

Abstract. We first characterize continuous spectrum and purely discrete spectrum of an isometry U of a Hilbert space geometrically by the existence of a spanning system, resp. by the absence, of vectors with infinitely many orthogonal images under powers of U. We then characterize weak mixing and discrete spectrum of an invertible measure preserving transformation of a probability space in terms of the null sets of the space. Finally for two-fold weakly mixing transformations the result on isometries is strengthened by proving the density of the set of partitions with infinitely many mutually independent images in the set of all finite partitions.

0. **Introduction.** This paper consists of three parts. Part 1 deals with spectral theory in Hilbert space, part 2 and part 3 deal with spectral properties of measure preserving (m.p.) transformations of a probability space. The three parts are closely related, but can be studied independently—except that some preliminary results in part 1 (up to Proposition 1.2) are used in part 2.

We first consider an isometry U of a complex Hilbert space $\mathfrak{H}$. We say that U has purely discrete spectrum if $\mathfrak{H}$ is spanned by the eigenvectors of U, and that U has continuous spectrum if U has no eigenvectors. We give a geometric characterization as follows: A vector $f \in \mathfrak{H}$ is called weakly wandering if there exists a strictly increasing sequence $0 = k_0 < k_1 < k_2 < \cdots$ of nonnegative integers such that the vectors $U^{k_i} f \, (i \in Z^+)$ are orthogonal to each other. We show that U has continuous spectrum if and only if the weakly wandering vectors span $\mathfrak{H}$ and that U has purely discrete spectrum if and only if there exist no nonzero weakly wandering vectors. In the first case the weakly wandering vectors turn out to be dense in $\mathfrak{H}$. (The concept of weakly wandering vector is analogous to Hajian and Kakutani's concept of a weakly wandering set in a measure space [2]. Halmos [3] has used the term wandering subspace for subspaces $\mathfrak{K} \subseteq \mathfrak{H}$ all images $U^k \mathfrak{K} \, (k \in Z^+)$ of which are orthogonal to each other. It follows from the orthogonality relations of the functions $e^{2\pi i n x} \, (n \in Z, x \in [0, 1])$ that the existence of sufficiently many wandering vectors characterizes absolutely continuous spectrum, i.e. Lebesgue spectrum.)

In part 2 we consider a m.p. transformation T in a probability space $(\Omega, \mathscr{F}, \mu)$. T induces an isometric operator U_T in $\mathfrak{L}_2(\Omega, \mathscr{F}, \mu)$ by $U_T f = f \circ T$. T is called weakly mixing if U_T has continuous spectrum on the orthogonal complement of the space consisting of the equivalence classes of the constant functions. T is said to have discrete spectrum if U_T has purely discrete spectrum. A sequence (B_n) of $\mathscr{F}$-measurable sets is called remotely trivial if the intersection of the σ-algebras $\mathscr{G}_m$ $(m \geqq 0)$ generated by $\{B_{m+k}, k \in Z^+\}$ contains only sets of measure zero or one. Sucheston [18] has shown that T is strongly mixing if and only if for all $A \in \mathscr{F}$ every subsequence of $(T^{-n}A)$ contains a further subsequence which is remotely trivial. We give similar characterizations for weak mixing and discrete spectrum as follows: T is weakly mixing if and only if for all $A \in \mathscr{F}$ the sequence $(T^{-n}A)$ contains a remotely trivial subsequence. If T is ergodic T has discrete spectrum if and only if, for no $A \in \mathscr{F}$ other than null sets and their complements, the sequence $(T^{-n}A)$ contains a remotely trivial subsequence. A sequence (B_n) is remotely trivial if it has sufficiently good asymptotic independence properties. We are therefore able to derive Kušnirenko's characterizations of discrete spectrum in terms of his entropy for subsequences [12] from our result.

The fact that weak mixing can be characterized in terms of the null sets of the measure space suggests the consideration of weak mixing for nonsingular transformations. This approach has been taken in the case of strong mixing by Professor Sucheston and the author [10]. We give a heuristic explanation of the more formal definitions given in [10]. It will then be clear that the formal approach (which was preferred in [10] in order to obtain a unified treatment of mixing and complete mixing) should not be applied to the above characterization of weak mixing.

In part 3 we ask for exact independence along a subsequence rather than for asymptotic independence. Let T be an invertible ergodic m.p. transformation of a probability space $(\Omega, \mathscr{F}, \mu)$. If the entropy $h(T)$ of T is positive, Sinai's theorem [20] implies the existence of *independent* partitions, i.e. of partitions $\xi = \{A_1, A_2, \ldots, A_n\}$ of Ω, for which $\xi, T^{-1}\xi, T^{-2}\xi, \ldots$ are mutually independent. We call a partition ξ *weakly independent*, if there exists a subsequence $0 = k_0 < k_1 < k_2 < \cdots$ of the integers such that $T^{-k_0}\xi, T^{-k_1}\xi, T^{-k_2}\xi, \ldots$ are mutually independent. In the case $h(T) = 0$ Kušnirenko's results on entropy for subsequences suggest the search for weakly independent partitions. If T is weakly mixing this comes down to the question whether approximate independence of partitions $T^{-k_0}\xi, T^{-k_1}\xi, \ldots, T^{-k_t}\xi$ can be changed into exact independence by a modification of the partition. Observe that only n sets can be modified, but for independence to hold $(n-1)n^t$ equations must be satisfied.

We call T 2-sided weakly mixing if for all $A, B, C \in \mathscr{F}$,

$$(0.1) \qquad \lim_{m \to \infty} \frac{1}{m} \sum_{k=0}^{m-1} |\mu(T^{-k}A \cap B \cap T^k C) - \mu(A)\mu(B)\mu(C)| = 0.$$

Weak mixing is the special case where (0.1) holds for all A, B and for $C = \Omega$.

We show that weak mixing is necessary and 2-sided weak mixing is sufficient for the density of the set of weakly independent partitions in the set of all finite partitions. We conjecture that weak mixing is both necessary and sufficient.

We do not know whether weak mixing and 2-sided weak mixing are equivalent. In fact, 2-sided weak mixing is implied by 2-fold weak mixing as defined by Leonov [13] and the question whether 2-fold weak mixing coincides with weak mixing is open. In the class of stationary Gaussian processes weak mixing and r-fold weak mixing are equivalent for each $r \geqq 2$.

If f is a function with zero integral, taking only finitely many values $c_1, \ldots, c_n$ and if the partition $\xi = \{A_1, \ldots, A_n\}$ with $A_i = \{f = c_i\}$ is weakly independent, then f is a weakly wandering vector in $\mathfrak{L}_2$ for the isometry U_T. Our result on weakly independent partitions therefore strengthens the general theorem on isometries with continuous spectrum in the special case of 2-sided weakly mixing transformations. Note that even in the case of positive entropy the present result cannot be replaced by Sinai's theorem, because that theorem makes no density statement and therefore does not yield a system of weakly wandering vectors spanning the Hilbert space.

The results of part 2 and part 3 of this paper suggest the following question: Is it true that nontrivial weakly independent partitions exist if and only if T does not have discrete spectrum? The "only if" part is of course obvious.

We remark that all relations between sets and functions are meant to hold modulo null sets.

Notation. A^c denotes the complement of a set A, $A \triangle B$ the symmetric difference of A and B, and $A \backslash B$ the set $A \cap B^c$. 1_A denotes the indicator function of A, i.e. the function which takes the value 1 in A and the value 0 in A^c. Let

$$Z = \{0, \pm 1, \pm 2, \ldots\}, \quad Z^+ = \{0, 1, 2, \ldots\} \quad \text{and} \quad N = \{1, 2, 3, \ldots\}.$$

If E is a finite set, $|E|$ denotes the number of its elements.

I would like to acknowledge my indebtedness to ideas of Kušnirenko and Sucheston. I would like to thank Professor J. Neveu for sending me a copy of his lecture notes on ergodic theory, Professor W. Parry for informing me about a relevant result of Mrs. Dowker, and Professors Akcoglu, Papangelou and Sucheston for stimulating discussions.

1. **Weakly wandering vectors.** We consider a complex Hilbert space $\mathfrak{H}$ with scalar product $(\cdot, \cdot)$ and norm $\|\cdot\|$. The complex conjugate of a scalar c is denoted by $\bar{c}$.

For the convenience of the reader we start with a preliminary discussion of some spectral theoretic results, which will be needed for the proof of our theorem. This discussion is based on lecture notes of Neveu [14].

A subset M of Z^+ is said to have density zero if $n^{-1} \sum_{k=0}^{n-1} 1_M(k)$ tends to zero. A sequence $(a_n, n \in Z^+)$ of complex numbers is said to converge strongly Cesàro

to a (in symbols $a_n \curvearrowright a$) if $n^{-1} \sum_{k=0}^{n-1} |a_n - a|$ tends to zero. It is an exercise to show that a bounded sequence $(a_n, n \in Z^+)$ converges strongly Cesàro to a if and only if there exists a subset M of Z^+ of density zero such that $\lim_{n \to \infty; n \notin M} a_n = a$.

LEMMA 1.1. *Let v be a bounded measure on the 1-dimensional torus $T_1 = [0, 1[$ and $\hat{v}(n) = \int e^{2\pi i n x} v(dx)$ $(n \in Z)$ its Fourier-transform. The sequence $(\hat{v}(n), n \in Z^+)$ converges strongly Cesàro to zero if and only if v has no atoms (i.e., $v(\{y\}) = 0$ for each $y \in T_1$).*

Proof. As $(\hat{v}(n))$ is bounded the above characterization of strong Cesàro convergence implies that $\hat{v}(n) \curvearrowright 0$ is equivalent to $\hat{v}(n)^2 \curvearrowright 0$. We have

$$n^{-1} \sum_{k=0}^{n-1} |\hat{v}(k)|^2 = n^{-1} \sum_{k=0}^{n-1} \int_{T_1} v(dx) e^{2\pi i x k} \int_{T_1} v(dy) e^{-2\pi i y k}$$

$$= \int \left[n^{-1} \sum_{k=0}^{n-1} e^{2\pi i (x-y)k} \right] v(dx) v(dy).$$

The last term tends to

$$(1.1) \qquad \int_{\{x=y\}} v(dx) v(dy) = \sum_{z \in T_1} v(\{z\})^2$$

as $n \to \infty$, because the integrands [. . .] are everywhere bounded by 1 and converge to $1_{\{x=y\}}$. □

PROPOSITION 1.1. *Let U be an isometry in the Hilbert space $\mathfrak{H}$ and let U_n $(n \in Z)$ be defined by $U_n = U^n$ $(n \geq 0)$ and $U_n = U^{*|n|}$ $(n < 0)$. For every $f \in \mathfrak{H}$ there exists a unique bounded nonnegative measure μ_f on $T_1 = [0, 1[$ such that*

$$(1.2) \qquad (U_n f, f) = \int_{T_1} e^{2\pi i n x} \mu_f(dx) \qquad (n \in Z).$$

μ_f *is called the spectral measure of f.*

Proof. For every finite sequence $\{c_k, |k| \leq n\}$ of complex numbers one has

$$\sum_l \sum_m c_l \bar{c}_m (U_{l-m} f, f) = \sum_l \sum_m c_l \bar{c}_m (U_l f, U_m f) = \left\| \sum_l c_l U_l f \right\|^2 \geq 0.$$

The sequence $(U_k f, f)$ is therefore nonnegative definite and the proposition follows from a general theorem of Herglotz [5]. (The proposition holds even for contractions, but the proof is then more difficult.) □

An element $0 \neq f \in \mathfrak{H}$ is called an eigenvector for the isometry U if there exists a complex number c (the eigenvalue) such that $Uf = cf$. Let $\mathfrak{H}_d$ be the subspace of $\mathfrak{H}$ which is spanned by the eigenvectors of U and let $\mathfrak{H}_c$ be its orthogonal complement.

PROPOSITION 1.2. *The following statements about an element $f \in \mathfrak{H}$ are equivalent:*
(a) $f \in \mathfrak{H}_c$,
(b) μ_f *has no atoms,*
(c) $(U^k f, f) \curvearrowright 0$,
(d) $(U^k f, g) \curvearrowright 0$ *for all $g \in \mathfrak{H}$.*

Proof. (a) $\Rightarrow$ (b). Let $f \in \mathfrak{H}_c$ and assume $\mu_f(\{y\}) > 0$ for some $y \in T_1$. Applying von Neumann's mean ergodic theorem to the contraction $e^{-2\pi iy}U$ we find that the averages $n^{-1} \sum_{k=0}^{n-1} (e^{-2\pi iy}U)^k f$ converge to an element $g \in \mathfrak{H}$ with $e^{-2\pi iy}Ug = g$. We have

$$(g, f) = \lim_{n \to \infty} n^{-1} \sum_{k=0}^{n-1} e^{-2\pi iyk}(U^k f, f)$$

$$= \lim_{n \to \infty} \int_{T_1} n^{-1} \sum_{k=0}^{n-1} e^{2\pi ik(x-y)} \mu_f(dx) = \mu_f(\{y\}) > 0.$$

Hence $g \neq 0$, and g is an eigenvector (with eigenvalue $e^{2\pi iy}$) which is not orthogonal to f. This contradicts our assumption.

(b) $\Rightarrow$ (c). This follows from Lemma 1.1.

(c) $\Rightarrow$ (d). The set of elements $g \in \mathfrak{H}$ such that $(U^k f, g) \rightsquigarrow 0$ is a closed linear subspace of $\mathfrak{H}$. From $(U^k f, f) \rightsquigarrow 0$ and the fact that U is an isometry it follows that every g of the form $U_l f$ belongs to that subspace. For every vector g which is orthogonal to all vectors $U_l f$ we even have $(Uf, g) \equiv 0$. The subspace therefore coincides with $\mathfrak{H}$.

(d) $\Rightarrow$ (a). Let $g \neq 0$ be an eigenvector with eigenvalue c. As U is an isometry we have $|c| = 1$ and $U^*g = \bar{c}g$. Now $(U^k f, g) \rightsquigarrow 0$ and $(U^k f, g) = (f, U^{*k}g) = (f, \bar{c}^k g) = c^k(f, g)$ imply that f is orthogonal to g. As g was an arbitrary eigenvector, $f \in \mathfrak{H}_c$ follows. $\square$

Recall that U is said to have purely discrete spectrum if $\mathfrak{H} = \mathfrak{H}_d$ and continuous spectrum if $\mathfrak{H} = \mathfrak{H}_c$. We are now in a position to prove our main result.

THEOREM 1.1. *An isometry U of a Hilbert space $\mathfrak{H}$ has continuous spectrum if and only if the weakly wandering vectors span $\mathfrak{H}$. It has purely discrete spectrum if and only if there exist no nonzero weakly wandering vectors. If the weakly wandering vectors span $\mathfrak{H}$, they are even dense in $\mathfrak{H}$.*

Proof. Let us first show that the existence of a spanning system of weakly wandering vectors implies for all $f, g \in \mathfrak{H}$ a convergence statement which is stronger than $(U^k f, g) \rightsquigarrow 0$.

LEMMA 1.2. *Let U be an isometry in $\mathfrak{H}$, $\|f\| = 1$, and assume that for some finite sequence $0 = k_0 < k_1 < \cdots < k_{r-1}$ the vectors $U^{k_0}f, \ldots, U^{k_{r-1}}f$ are orthogonal. Then we have for all $h \in \mathfrak{H}$*

$$(1.3) \qquad \limsup_{n \to \infty} \left(\sup_{j \geq 0} n^{-1} \sum_{k=j}^{j+n-1} |(U^k f, h)|^2 \right) \leq r^{-1}\|h\|^2.$$

Proof. We may assume $\|h\| = 1$. Let $f_i = U^{k_i}f$ $(0 \leq i < r)$. For each $l \in N$ the vectors $U^l f_i$ are orthogonal and have length ≤ 1. Bessel's inequality therefore implies for all $l \in N$

$$\sum_{i=0}^{r-1} |(U^l f_i, h)|^2 \leq \|h\|^2 = 1.$$

For any $n \geqq 2k_{r-1}$ and any $j \geqq 0$ we have

$$\sum_{k=j}^{j+n-1} |(U^k f, h)|^2 \leqq k_{r-1} + \sum_{k=j+k_{r-1}}^{j+n-1} |(U^{k-k_i} f_i, h)|^2 \qquad (0 \leqq i < r).$$

Hence

$$\sum_{k=j}^{j+n-1} |(U^k f, h)|^2 \leqq k_{r-1} + r^{-1} \sum_{k=j+k_{r-1}}^{j+n-1} \sum_{i=0}^{r-1} |(U^{k-k_i} f_i, h)|^2$$

$$\leqq k_{r-1} + 2k_{r-1} + r^{-1} n.$$

Dividing by n and passing to the limit we get (1.3). □

The lemma implies for weakly wandering vectors $f \in \mathfrak{H}$ the relation

$$(1.4) \qquad \lim_{n \to \infty} \left(\sup_{j \geqq 0} n^{-1} \sum_{k=j}^{j+n-1} |(U^k f, h)|^2 \right) = 0 \qquad (h \in \mathfrak{H}).$$

(1.4) means that, for every $\varepsilon > 0$, $n^{-1} |\{k : j \leqq k \leqq j+n-1, |(U^k f, h)|^2 > \varepsilon\}|$ tends to zero uniformly in j. Thus (1.4) is equivalent to

$$(1.5) \qquad \lim_{n \to \infty} \left(\sup_{j \geqq 0} n^{-1} \sum_{k=j}^{j+n-1} |(U^k f, h)| \right) = 0 \qquad (h \in \mathfrak{H}).$$

The set of vectors $f \in \mathfrak{H}$, for which (1.5) holds, is a closed subspace of $\mathfrak{H}$. Therefore, if the weakly wandering vectors span $\mathfrak{H}$, (1.5) holds for all $f \in \mathfrak{H}$. Hence $(U^k f, h) \leadsto 0$ $(f, h \in \mathfrak{H})$, therefore $\mathfrak{H}_c = \mathfrak{H}$. (1.5) can be interpreted as the statement that the maximal Banach limit of the sequence $(|(U^k f, h)|)$ is zero for all $f, h \in \mathfrak{H}$, see [19]. The fact that (1.5) holds in the case of continuous spectrum is known [7].

It is now sufficient to prove that in the case $\mathfrak{H} = \mathfrak{H}_c$ the weakly wandering vectors are dense in $\mathfrak{H}$. This is the crucial part of the proof. If U does not have purely discrete spectrum we can apply this result to the restriction of U to the invariant subspace $\mathfrak{H}_c \neq \{0\}$ to obtain a nonzero weakly wandering vector. Conversely, if there exists a weakly wandering vector $f \neq 0$, Lemma 1.2 and the considerations following it imply $(U^k f, f) \leadsto 0$. Proposition 1.2 then implies $\mathfrak{H}_c \neq \{0\}$. We formulate the required result as a lemma:

LEMMA 1.3. *Let U be an isometry of the Hilbert space $\mathfrak{H}$ and let $\mathfrak{H} = \mathfrak{H}_c$. Then for all $f \in \mathfrak{H}$ and all $\varepsilon > 0$ there exists an $f^{(\varepsilon)} \in \mathfrak{H}$ with*

$$(1.6) \qquad \qquad \|f - f^{(\varepsilon)}\| < \varepsilon$$

and a sequence $0 = k_0 < k_1 < k_2 < \cdots$ of integers such that

$$(1.7) \qquad (U^{k_i} f^{(\varepsilon)}, U^{k_j} f^{(\varepsilon)}) = 0 \qquad (0 \leqq i, j < \infty, i \neq j).$$

The proof proceeds by an inductive construction. Following a suggestion of the referee we start with an informal outline designed to familiarize the reader with the ideas. The proof below is independent of it.

Assume, for simplicity, $\|f\| = 1$. Proposition 1.2 implies that $|(U^k f, f)|$ is small for most sufficiently large values of k. We start by modifying $f_0 = f$ into an f_1 close to f_0 so that $(U^{k_1} f_1, f_1) = 0$ for some k_1. We explain below how this is done.

When f_t and $0 = k_0 < k_1 < \cdots < k_t$ have been determined such that the vectors $U^{k_i}f_t$ $(0 \le i \le t)$ are orthogonal, we want to find k_{t+1} and to modify f_t into f_{t+1} such that the vectors $U^{k_i}f_{t+1}$ $(0 \le i \le t+1)$ are orthogonal. It seems to be difficult to obtain the modification $f_t \to f_{t+1}$ in one step. We introduce a mapping χ_t of a neighborhood of f_t into a neighborhood of f_t such that the deviation from orthogonality (measured in terms of $M_t^a(g) = \max \{\|g\|^{-2} |(U^{k_{t+1}}g, U^{k_i}g)|, 0 \le i \le t\}$) is cut in half, i.e. $M_t^a(\chi_t(g)) \le 2^{-1} M_t^a(g)$.

The definition of χ_t is most intuitive for $t = 0$: Take k_1 such that $f_0 = f$, $U^{k_1}f$, and $U^{2k_1}f$ are nearly orthogonal. Change f_0 into $\chi_0(f_0) = f_0 - (f_0, U^{k_1}f_0)U^{k_1}f_0$ and compare $\chi_0(f_0)$ and $U^{k_1}\chi_0(f_0)$. The reader may like to draw a picture to see that the orthogonality of $\chi_0(f_0)$ and $U^{k_1}\chi_0(f_0)$ is much better than that of f_0 and $U^{k_1}f_0$. In fact, if g is close to f_0 and we define $\chi_0(g) = g - \|g\|^{-2}(g, U^{k_1}g)U^{k_1}g$, then the orthogonality of $\chi_0(g)$ and $U^{k_1}\chi_0(g)$ is much better than that of g and $U^{k_1}g$.

For $t = 0$ we can apply χ_t over and over again and the sequence $\chi_t^n(g)$ converges to a vector f_1 with $(f_1, U^{k_1}f_1) = 0$. Unfortunately for $t \ge 1$, χ_t has the property that the vectors $U^{k_i}\chi_t(f_t)$ $(0 \le i \le t)$ need no longer be orthogonal. We fix this by applying the entire orthogonalization procedure defined up to step t to $\chi_t(f_t)$. Let $\varphi_t(g)$ be the vector obtained by applying the orthogonalization procedure of the previous steps to g. We change f_t into $\chi_t(f_t)$, then into $\varphi_t\chi_t(f_t)$, then into $\chi_t\varphi_t\chi_t(f_t)$, then into $\varphi_t\chi_t\varphi_t\chi_t(f_t)$, etc. Under suitable assumptions the limit exists. We call it f_{t+1}. The constructions are possible if we define χ_t in such a way that for $\psi_t = \varphi_t\chi_t$

(*) $$M_t^a(\psi_t(g)) \le 2^{-1} M_t^a(g)$$

and such that in the course of the construction we never leave the neighborhood of f_t, in which χ_t and φ_t are defined and satisfy the required inequalities. It is, in fact, sufficient to have (*) for those g for which $g, U^{k_1}g, \ldots, U^{k_t}g$ are orthogonal. Each $\psi_t(g)$ has this property, therefore we get $M_t^a(\psi_t^n(g)) \le 2^{-(n-1)} M_t^a(\psi_t(g))$.

To avoid that one of the modifications of f_t lies outside the prescribed neighborhoods we need an estimate of $\|\varphi_t(g) - g\|$. Such an estimate is given in the induction hypothesis below.

We hope that these remarks help to clarify the formal proof which follows.

Proof of Lemma 1.3. We may assume

(1.8) $$\|f\| = 1 \quad \text{and} \quad \varepsilon < \tfrac{1}{4}.$$

Step 0 of the construction consists in specifying $f_0 = f$, $k_0 = 0$, $C_0 = 1$, $\delta_0 = 1$, and $\varphi_0(g) = g$ $(g \in \mathfrak{H})$.

At the end of step t $(t \ge 0)$ of the construction we have determined a sequence of integers $0 = k_0 < k_1 < \cdots < k_t$, an element $f_t \in \mathfrak{H}$, constants $C_t > 0$ and $\delta_t > 0$ and a mapping φ_t of a neighborhood of f_t into $\mathfrak{H}$. It is convenient to introduce the notation $M_0(g) = 0$,

$$M_t(g) = \max \{|(U^{k_i}g, U^{k_j}g)|, 0 \le i < j \le t\} \qquad (t > 0).$$

The element f_t satisfies $M_t(f_t)=0$ and

$$(1.9) \qquad \qquad \|f-f_t\| < 2^{-2}\varepsilon(2^{-0}+2^{-1}+\cdots+2^{-t}).$$

C_t, δ_t and φ_t are such that $\varphi_t(g)$ is defined for all $g \in \mathfrak{H}$ with $\|f_t-g\|+C_tM_t(g)<\delta_t$ and such that $\varphi_t(g)$ for these elements g of $\mathfrak{H}$ satisfies $M_t(\varphi_t(g))=0$ and $\|g-\varphi_t(g)\| \leqq C_tM_t(g)$.

Observe that these inequalities imply $\varphi_t(f_t)=f_t$.

Step $t+1$. We proceed to construct k_{t+1}, f_{t+1}, C_{t+1}, δ_{t+1} and φ_{t+1}. In this construction we shall deal with modifications g of f_t and we shall assume without mention that they will have the property

$$(1.10) \qquad \qquad 2^{-1} < \|g\| < 2.$$

This assumption will be justified later. For a modification g of f_t and for an integer $k > k_t$ let

$$a_i(g) = -\|g\|^{-2}(U^{k_i}g, U^kg), \qquad M_t^a(g) = \max\{|a_i(g)|, 0 \leqq i \leqq t\}$$

and

$$\chi_t(g) = g+ \sum_{\tau=0}^{t} a_\tau(g)U^{k-k_\tau}g.$$

The dependence of $a_i(g)$, $M_t^a(g)$, and $\chi_t(g)$ from k is suppressed in the notation, because we will choose a fixed value of k later in the construction. We have to consider the scalar products

$$(U^{k_i}\chi_t(g), U^k\chi_t(g)) = (U^{k_i}g, U^kg)+a_i(g)(U^kg, U^kg)$$

$$(1.11) \qquad \qquad + \sum_{\tau \neq i} a_\tau(g)(U^{k-k_\tau+k_i}g, U^kg)+ \sum_{\sigma=0}^{t} \bar{a}_\sigma(g)(U^{k_i}g, U^{2k-k_\sigma}g)$$

$$+ \sum_{\sigma=0}^{t}\sum_{\tau=0}^{t} \bar{a}_\sigma(g)a_\tau(g)(U^{k-k_\tau+k_i}g, U^{2k-k_\sigma}g),$$

and

$$(U^{k_i}\chi_t(g), U^{k_j}\chi_t(g)) = (U^{k_i}g, U^{k_j}g)+ \sum_{\tau=0}^{t} a_\tau(g)(U^{k-k_\tau+k_i}g, U^{k_j}g)$$

$$(1.12) \qquad \qquad + \sum_{\sigma=0}^{t} \bar{a}_\sigma(g)(U^{k_i}g, U^{k-k_\sigma+k_j}g)$$

$$+ \sum_{\sigma=0}^{t}\sum_{\tau=0}^{t} \bar{a}_\sigma(g)a_\tau(g)(U^{k-k_\tau+k_i}g, U^{k-k_\sigma+k_j}g).$$

Let

$$M_t^* = \max\{|U^{k_i}f_t, U^{2k-k_\sigma}f_t)|, |(U^{k_i}f_t, U^{k-k_\sigma+k_j}f_t)|, 0 \leqq i, j, \sigma \leqq t\}.$$

Observe that the first two terms on the right side of (1.11) cancel by our choice of $a_i(g)$. As U is an isometry the definitions of $a_i(g)$ and $M_t^a(g)$ give

$$(1.13) \qquad \qquad |M_t^a(g)| \leqq 1.$$

The equations (1.10), (1.11), (1.12) and (1.13) now yield the following estimates:

(1.14) $|(U^{k_i}\chi_t(g), U^k\chi_t(g))| \leqq tM_t^a(g)M_t(g)+2(t+1)^2M_t^a(g)(M_t^*+4\|f_t-g\|)$

and for $k_i < k_j$

(1.15)
$$|(U^{k_i}\chi_t(g), U^{k_j}\chi_t(g))|$$
$$\leqq M_t(g)+2(t+1)(M_t^*+4\|f_t-g\|)M_t^a(g)+4(t+1)^2M_t^a(g)^2.$$

It follows that

(1.16) $M_t(\chi_t(g)) \leqq M_t(g)+2(t+1)(M_t^*+4\|f_t-g\|)M_t^a(g)+4(t+1)^2M_t^a(g)^2.$

If g is such that

(1.17) $\|f_t-\chi_t(g)\|+C_tM_t(\chi_t(g)) < \delta_t$

then we can find an element $\varphi_t(\chi_t(g))=\psi_t(g) \in \mathfrak{H}$ with the properties

(1.18) $M_t(\psi_t(g)) = 0$ and $\|\chi_t(g)-\psi_t(g)\| \leqq C_tM_t(\chi_t(g)).$

The equations (1.10), (1.14), (1.16) and (1.18) yield the following estimate for $M_t^a(\psi_t(g))$:

$$M_t^a(\psi_t(g)) \leqq 4 \max \{|(U^{k_i}\psi_t(g), U^k\psi_t(g))|, 0 \leqq i \leqq t\}$$
$$\leqq 4(\max \{|(U^{k_i}\chi_t(g), U^k\chi_t(g))|, 0 \leqq i \leqq t\}+4\|\chi_t(g)-\psi_t(g)\|)$$
$$\leqq 4(tM_t(g)M_t^a(g)+2(t+1)^2(M_t^*+4\|f_t-g\|)M_t^a(g))+16C_tM_t(\chi_t(g))$$
$$\leqq 16C_tM_t(g)+M_t^a(g)\{4tM_t(g)+8(t+1)^2(M_t^*+4\|f_t-g\|)$$
$$+32(t+1)C_t(M_t^*+4\|f_t-g\|)+64(t+1)^2C_tM_t^a(g)\}.$$

Clearly the same estimate holds for $M_t^a(\chi_t(g))$. We will show later that it is possible to choose k such that the expression { } after the last inequality is smaller than 2^{-1}, and such that this remains true if g in { } is replaced by a modification occurring later in the construction (such as $g^{(1)}, g^{(2)}, \dots$ defined below). This k will be the k_{t+1}.

By induction we have $M_t(\psi_t(g))=0$. Applying the above estimates to $g^{(0)}=g$, $g^{(n+1)}=\psi_t(g^{(n)})$ we find $M_t(g^{(n)})=0$ ($n \geqq 1$) and therefore

(1.19) $M_t^a(g^{(0)}) = M_t^a(g),$

(1.20) $M_t^a(g^{(1)}) \leqq 16C_tM_t(g)+2^{-1}M_t^a(g),$

(1.21) $M_t^a(g^{(n+1)}) \leqq 2^{-1}M_t^a(g^{(n)})$ ($n \geqq 1$).

For each $n \geqq 0$ the estimate given for $M_t^a(g^{(n+1)})$ is also an upper estimate for $M_t^a(\chi_t(g^{(n)}))$. We shall show that the sequence $g^{(n)}$ converges and that the construction does not break down at some stage, if $k=k_{t+1}$ is selected appropriately. The vector $\varphi_{t+1}(g)=\lim_{n \to \infty} g^{(n)}=\lim_{n \to \infty} \psi_t^n(g)$ has then the required orthogonality

property $M_{t+1}(\varphi_{t+1}(g)) = 0$. We estimate the distance of the modification $\psi_t(g)$ of g from g:

$$\|\psi_t(g) - g\| \leqq \|g - \chi_t(g)\| + \|\chi_t(g) - \psi_t(g)\|$$

$$(1.22) \qquad \leqq 2(t+1)M_t^a(g) + C_t M_t(\chi_t(g))$$

$$\leqq C_t M_t(g) + M_t^a(g)\{2(t+1) + 2(t+1)C_t(M_t^* + 4\|f_t - g\|) + 4(t+1)^2 C_t M_t^a(g)\}.$$

Similarly, for $n \geqq 1$,

$$(1.23) \qquad \begin{aligned} &\|g^{(n)} - g^{(n+1)}\| \\ &\leqq M_t^a(g^{(n)})\{2(t+1) + 2(t+1)C_t(M_t^* + 4\|f_t - g^{(n)}\|) + 4(t+1)^2 C_t M_t^a(g^{(n)})\}. \end{aligned}$$

For each $n \geqq 0$ the estimate given for $\|g^{(n+1)} - g^{(n)}\|$ is also an upper estimate for $\|\chi_t(g^{(n)}) - g^{(n)}\|$. As long as we take care that $2^{-1} < \|g^{(n)}\| < 2$ holds for all $n \geqq 0$, we may estimate $\{\ \}$ in (1.22) and (1.23) from above by a fixed constant D_t. It follows that $g^{(n)}$ converges (if the construction of all $g^{(n)}$ is possible and our various assumptions do not break down at some stage), and that

$$\|g - \varphi_{t+1}(g)\| \leqq \sum_{n=0}^{\infty} \|g^{(n)} - g^{(n+1)}\| \leqq C_t M_t(g) + D_t \sum_{n=0}^{\infty} M_t^a(g^{(n)}).$$

If we write k_{t+1} for k in the definition of $a_i(g)$ we see that

$$(1.24) \qquad M_{t+1}(g) = \max\{M_t(g), \|g\|^2 M_t^a(g)\}.$$

Let $C_{t+1} = (64C_t + 8)D_t$. Then (1.19)–(1.22) and (1.24) imply $\|g - \psi_t(g)\| \leqq 2^{-1}C_{t+1}M_{t+1}(g)$ and $D_t M_t^a(g^{(n)}) \leqq 2^{-(n+1)}C_{t+1}M_{t+1}(g)$ $(n \geqq 0)$ and therefore

$$(1.25) \qquad \|g - \varphi_{t+1}(g)\| \leqq \sum_{n=0}^{\infty} \|g^{(n)} - g^{(n+1)}\| \leqq C_{t+1}M_{t+1}(g).$$

We now have to show that the construction actually works if g is sufficiently close to f_t and $k = k_{t+1}$ is determined in such a way that M_t^* and $M_t^a(f_t)$ are small enough. We first introduce auxiliary constants: By (1.9) we have $2^{-1} < \|f_t\| < 2$. Determine positive real numbers η_t, ρ_t and α_t with $0 < \eta_t < 1$, $0 < \rho_t < 2^{-(t+1)}\varepsilon$, and $0 < \alpha_t < 2^{-2}\rho_t$ such that in the case $M_t^* < \eta_t$ any $h \in \mathfrak{H}$ which satisfies

$$(1.26) \qquad \|h - f_t\| < \rho_t$$

and

$$(1.27) \qquad M_t^a(h) < \alpha_t$$

also satisfies

$$2^{-1} < \|h\| < 2, \qquad \|f_t - h\| + C_t M_t(h) < \delta_t$$

and

$$\begin{aligned} \{4t M_t(h) + 8(t+1)^2(M_t^* + 4\|f_t - h\|) + 32(t+1)C_t(M_t^* + 4\|f_t - h\|) \\ + 64(t+1)^2 C_t M_t^a(h)\} < 2^{-1}. \end{aligned}$$

For such a choice of η_t, ρ_t and α_t the number k does not have to be known. As U has continuous spectrum, we have $(U^k h_1, h_2) \rightsquigarrow 0$ for all $h_1, h_2 \in \mathfrak{H}$. It follows that for any $\gamma > 0$ and for any finite system of vectors $h_{1,\nu}$ and $h_{2,\mu}$ there are arbitrarily large values of k for which $|(U^k h_{1,\nu}, h_{2,\mu})| < \gamma$ and $|(U^{2k} h_{1,\nu}, h_{2,\mu})| < \gamma$ hold for all values of ν and μ. It is therefore possible to find a number $k > k_t$ such that $M_t^* < \eta_t$ and

$$(1.28) \qquad 8C_{t+1} M_t^a(f_t) < 2^{-(t+2)} \alpha_t.$$

This k is now fixed and is called k_{t+1}. It follows from $M_t(f_t) = 0$ and (1.28) that there exists a number $\beta_t > 0$ with $\beta_t < 2^{-1} \alpha_t$ such that for all $g \in \mathfrak{H}$ with

$$(1.29) \qquad \|f_t - g\| < \beta_t$$

the inequalities

$$(1.30) \qquad 32C_{t+1} M_t(g) < \alpha_t \quad \text{and} \quad 16C_{t+1} M_t^a(g) < 2^{-t}\alpha_t$$

are satisfied. We have then for such a $g \in \mathfrak{H}$ the inequalities

$$M_t^a(g^{(0)}) = M_t^a(g) < \alpha_t,$$

$$(1.31) \qquad M_t^a(g^{(1)}) \leqq 16C_t M_t(g) + 2^{-1} M_t^a(g) < \alpha_t$$

and

$$(1.32) \qquad M_t^a(g^{(n+1)}) \leqq M_t^a(g^{(n)}) < \alpha_t \qquad (n \geqq 1).$$

Using (1.30) and (1.25) we get

$$\sum_{n=0}^{\infty} \|g^{(n)} - g^{(n+1)}\| \leqq C_{t+1} M_{t+1}(g) < 2^{-2}\alpha_t + 2^{-2}\alpha_t = 2^{-1}\alpha_t.$$

Therefore each $g^{(n)}$ obtained in the construction satisfies $\|g^{(n)} - f_t\| < \rho_t$. Considering this, (1.31) and (1.32) we see that each $g^{(n)}$ is an h satisfying (1.26) and (1.27). Remember that estimates for $M_t^a(g^{(n+1)})$ also are estimates for $\chi_t(g^{(n)})$, and estimates for $\|g^{(n)} - g^{(n+1)}\|$ also are estimates for $\|g^{(n)} - \chi_t(g^{(n)})\|$. Therefore also each $\chi_t(g^{(n)})$ is an h satisfying (1.26) and (1.27). We have finally verified that for all $g \in \mathfrak{H}$ with (1.29) the assumptions made in the course of the construction remain satisfied for all $g^{(n)}$ and for all $\chi_t(g^{(n)})$.

We define f_{t+1} by

$$(1.33) \qquad f_{t+1} = \varphi_{t+1}(f_t).$$

It follows from (1.28) that

$$(1.34) \qquad \|f_{t+1} - f_t\| \leqq C_{t+1} M_{t+1}(f_t) = C_{t+1} \|f_t\|^2 M_t^a(f_t) < 2^{-(t+2)}\alpha_t.$$

Let $\delta_{t+1} = 2^{-1}\alpha_t$. If $g \in \mathfrak{H}$ satisfies $\|f_{t+1} - g\| + C_{t+1}M_{t+1}(g) < \delta_{t+1}$ then we can apply the construction and find $g^{(0)} = g$, $g^{(1)}$, $g^{(2)}$, ... because we have

$$\|f_t - g^{(n)}\| \leq \|f_t - f_{t+1}\| + \|f_{t+1} - g\| + \sum_{v=0}^{n-1} \|g^{(v)} - g^{(v+1)}\|$$

$$\leq 2^{-2}\alpha_t + \|f_{t+1} - g\| + C_{t+1}M_{t+1}(g) < \alpha_t < \rho_t$$

and

$$D_t \sum_{v=0}^{n} M_t^a(g^{(v)}) \leq C_{t+1}M_{t+1}(g) < \delta_{t+1} < \alpha_t$$

so that each $g^{(v)}$ is an element h satisfying (1.26) and (1.27). Again the same estimates hold for each $\chi_t(g^{(n)})$. The equations (1.9) and (1.34) imply

$$\|f - f_{t+1}\| < 2^{-2}\varepsilon(2^{-0} + \cdots + 2^{-(t+1)}).$$

We have thereby verified the induction hypothesis for $t+1$. By (1.34) the sequence f_t converges to an element $f^{(\varepsilon)} \in \mathfrak{H}$. As f_t satisfies $M_\tau(f_t) = 0$ for all $\tau \leq t$ the limit $f^{(\varepsilon)}$ satisfies (1.7). By (1.9) it also satisfies (1.6). The lemma and therefore also the theorem are completely proved. $\square$

It seems clear that for a unitary operator U the construction can be modified in such a way that a doubly infinite strictly increasing sequence $(k_i, i \in Z)$ and a vector $f^{(\varepsilon)}$ are obtained for which $(U^{k_i}f^{(\varepsilon)}, U^{k_j}f^{(\varepsilon)}) = 0$ $(-\infty < i, j < \infty; i \neq j)$ and $\|f - f^{(\varepsilon)}\| < \varepsilon$.

2. A characterization of weak mixing and discrete spectrum in terms of null-sets.
In this part of the paper we consider a measure preserving (m.p.) transformation T of a probability space $(\Omega, \mathscr{F}, \mu)$. T is said to have discrete spectrum if the isometry U_T in the complex Hilbert space $\mathfrak{L}_2 = \mathfrak{L}_2(\Omega, \mathscr{F}, \mu)$ given by $U_T f = f \circ T$ has purely discrete spectrum, i.e. if the eigenvectors $f \in \mathfrak{L}_2$ span $\mathfrak{L}_2$. T is said to be weakly mixing if

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

T is weakly mixing if and only if the only eigenvectors f of U_T are the constant functions (see [8, p. 81] or apply Proposition 1.2 of this paper).

A sequence $(B_k, k \in Z^+)$ of $\mathscr{F}$-measurable sets is called *remotely trivial* if the intersection of the σ-algebras $\mathscr{G}_m$ $(m \geq 0)$ generated by $\{B_{m+k}, k \in Z^+\}$ contains only sets of measure zero and their complements. A sequence $(A_n, n \geq 0)$ of measurable sets is called *mixing* if, for all $B \in \mathscr{F}$, $\lim_{n \to \infty} (\mu(A_n \cap B) - \mu(A_n)\mu(B)) = 0$. Sucheston [18] has shown that (A_n) is mixing if and only if every subsequence of (A_n) contains a remotely trivial subsequence. A short proof of Sucheston's result has recently been given by Lee K. Jones [22].

THEOREM 2.1. *A m.p. transformation T of a probability space $(\Omega, \mathscr{F}, \mu)$ is weakly mixing if and only if for all $A \in \mathscr{F}$ the sequence $(T^{-n}A, n \in Z^+)$ contains a remotely trivial subsequence.*

Proof. Let T be weakly mixing and let $A \in \mathscr{F}$ be given. Let $n_0 = 0$. When $n_0 < n_1 < \cdots < n_t$ have been determined find a natural number $n_{t+1} > n_t$ such that for $k = 0, \ldots, t$,

$$|\mu(T^{-n_{t+1}}A \cap T^{-n_k}A) - \mu(A)^2| < 2^{-t}.$$

The sequence $(A_t) = (T^{-n_t}A)$ satisfies

(2.2) $$\lim_{t \to \infty} (\mu(A_t \cap A_k) - \mu(A_t)\mu(A_k)) = 0 \qquad (k \geqq 0).$$

It is known [18] that (2.2) implies that (A_t) is mixing. (The proof is a Hilbert space argument similar to the proof of (c) $\Rightarrow$ (d) in Proposition 1.2.) Thus (A_t) has a remotely trivial subsequence $(A_{t_i}, i \geqq 0)$ by Sucheston's theorem. (A_{t_i}) is the desired remotely trivial subsequence of $(T^{-n}A)$.

Next assume that T is not weakly mixing. If T is not ergodic, there exists a set $I \in \mathscr{F}$ with $T^{-1}I = I$ and $0 < \mu(I) < 1$. Then the sequence $(T^{-n}I)$ has no remotely trivial subsequence. We may therefore assume that T is ergodic. Let $f \not\equiv \text{const.}$ be an eigenvector of norm 1 of U_T and $c \neq 1$ the corresponding eigenvalue. The ergodicity of T and the fact that U_T is an isometry imply $|c| = 1$ and $|f| \equiv 1$. If $\gamma = (2\pi i)^{-1} \log c$ is rational, we write γ in the form $\gamma = rs^{-1}$ with $(r, s) = 1$ and $s \geqq 2$. The sets $A_i = \{\omega \in \Omega : (2\pi i)^{-1} \log f(\omega) \in [(i-1)s^{-1}, is^{-1}[\}$ form a partition of Ω and we have $T^{-1}A_i = A_{i-r \pmod s}$. Therefore the sequence $(T^{-n}A_1)$ has no remotely trivial subsequence. If γ is irrational we define a measure λ on the Borel-sets $B \subseteq [0, 1[$ by $\lambda(B) = \mu(A(B))$ where $A(B) = \{\omega \in \Omega : (2\pi i)^{-1} \log f(\omega) \in B\}$. Let $T_\gamma x = x + \gamma \pmod 1$ $(x \in [0, 1[)$. We obtain $T^{-1}A(B) = A(T_\gamma^{-1}B)$, and this implies that λ is an invariant measure for T_γ. Thus λ must be the Lebesgue-measure. Let $A = A(B)$ for $B = [0, \frac{1}{2}[$. We claim that $(T^{-n}A)$ has no remotely trivial subsequence: Let $n_0 < n_1 < n_2 < \cdots$ be a strictly increasing sequence of nonnegative integers. For some subsequence $t_1 < t_2 < \cdots$ of the integers the sequence $(\alpha_k, k = 1, 2, \ldots)$ given by $\alpha_k = n_{t_k}\gamma \pmod 1$ converges in $[0, 1[$, because the circle $[0, 1[\pmod 1$ is compact. Let α be the limit. $T_\gamma^{-n}B$ is the set $B^n = \{x \in [0, 1[: x + n\gamma \pmod 1 \in B\}$. Let $B^\alpha = \{x \in [0, 1[: x + \alpha \pmod 1 \in B\}$. We have $\lambda(B^{n_{t_k}} \triangle B^\alpha) \to 0$ $(k \to \infty)$ and therefore $\mu(A(B^\alpha) \triangle T^{-n_{t_k}}A) \to 0$ $(k \to \infty)$. Therefore $A(B^\alpha)$ belongs $(\text{mod } \mu)$ to each σ-algebra $\mathscr{G}_m$ $(m \geqq 0)$ generated by $(T^{-n_{m+k}}A, k \geqq 0)$. As $\mu(A(B^\alpha)) = \frac{1}{2}$ the sequence $(T^{-n_t}A, t \geqq 0)$ is not remotely trivial. As the sequence (n_k) was arbitrary, $(T^{-n}A)$ has no remotely trivial subsequence. $\quad\square$

Before proceeding to the more difficult characterization of discrete spectrum, we would like to discuss some ideas about mixing and weak mixing of transformations without finite invariant measure.

For some time it was not known how the property of mixing should be defined for m.p. transformations of an infinite, σ-finite measure space $(\Omega, \mathscr{F}, \mu)$. As late as 1968 Krickeberg [11] began a survey paper on mixing in topological measure spaces with the remark: "It has been known for some time that mixing transformations in an abstract measure space cannot be defined in a reasonable way if the measure

of the entire space is infinite." At approximately the same time Professor Sucheston and the author [10] introduced the following definitions: A sequence $(A_n, n \geq 0)$ of measurable sets is called semiremotely trivial (s.r.t.) if every subsequence contains a further subsequence which is remotely trivial. A m.p. transformation T of a σ-finite measure space $(\Omega, \mathscr{F}, \bar{\mu})$ is called mixing if $(T^{-n}A)$ is s.r.t. for all $A \in \mathscr{F}$ of finite measure. T is called completely mixing if $(T^{-n}A)$ is s.r.t. for all $A \in \mathscr{F}$. Sucheston's theorem implies that both notions coincide with strong mixing when the measure space is finite. The notions proved useful in that they permitted extensions of a number of well-known results on mixing in finite measure spaces. However, we failed to point out explicitly that the notions have an intuitive interpretation. I shall, therefore, now describe a heuristic motivation.

In infinite measure spaces "mixing" is a very weak property, equivalent to the property "zero-type" of Hajian-Kakutani [2]: $\bar{\mu}(T^{-n}A \cap B) \to 0$ for all $A, B \in \mathscr{F}$ with finite measure. This only means that sets of finite measure are mapped in such a way that a given set of finite measure finally contains proportionally the right amount, namely, asymptotically nothing. The interest in this notion derives from the fact that it permits a generalization of the Blum-Hanson theorem.

Complete mixing is a rather strong mixing property, and it corresponds well to what intuition asks for: T induces a transformation $\mathbf{T}$ in the space of finite measures φ on $(\Omega, \mathscr{F})$ by $\mathbf{T}\varphi(A) = \varphi(T^{-1}A)$. We may interpret φ as a mass distribution in Ω and $\mathbf{T}\varphi$ as the distribution of the mass after transportation by T. Let μ_1 and μ_2 be probability measures absolutely continuous with respect to $\bar{\mu}$. We may think of μ_1 and μ_2 as distributions of a unit mass of red, resp. blue, paint in Ω. T is completely mixing if and only if for all such $\mu_1, \mu_2,$

$$(2.3) \qquad \mathbf{T}^n\mu_1(A) - \mathbf{T}^n\mu_2(A) \to 0 \qquad (A \in \mathscr{F}).$$

(2.3) means that after some time every set A contains approximately as much red paint as it contains blue paint, independent of where we placed the paint at the start. Of course for sets A of finite measure, (2.3) will often be true because $\mathbf{T}^n\mu_1(A)$ and $\mathbf{T}^n\mu_2(A)$ tend to zero. But in general $\mathbf{T}^n\mu_1(A)$ does not converge. An oscillation theorem of Mrs. Dowker [1, Theorem 3] even implies that there always exist sets $A \in \mathscr{F}$ for which even the Cesàro averages of $\mathbf{T}^n\mu(A)$ diverge. It was shown in [10] that complete mixing in infinite measure spaces even implies that the convergence (2.3) is uniform in A. This is more than is true for finite measure spaces.

Another intuitive interpretation is as follows: Let ν be a probability measure equivalent to $\bar{\mu}$: $\nu \ll \bar{\mu} \ll \nu$. T is completely mixing if and only if for all $A \in \mathscr{F}$,

$$(2.4) \qquad \lim_{n \to \infty} (\nu(T^{-n}A \cap B) - \nu(T^{-n}A)\nu(B)) = 0 \qquad (B \in \mathscr{F}).$$

We see that complete mixing, just like strong mixing in a finite measure space, is an asymptotic independence property. The remarkable fact that the validity of (2.4) is independent of the particular choice of ν follows from Sucheston's theorem —but it can also be established directly by using the Radon-Nikodym theorem.

Rényi [16] had already observed that the statement $\lim_{n \to \infty} \nu(A_n \cap B) = \alpha \cdot \nu(B)$ ($B \in \mathscr{F}$) remains valid when ν is replaced by an equivalent probability measure. But he was only interested in the case where the limit exists.

We hope that these remarks suffice to show that the definitions in [10] are "reasonable".

It is now tempting to use the characterization of weak mixing given in Theorem 2.1 as a definition of weak mixing in infinite measure spaces. However, we do not believe that such a definition would be reasonable. Our characterization is not valid for arbitrary sequences of sets: Call a sequence $(A_n, n \geq 0)$ of measurable sets in a probability space $(\Omega, \mathscr{F}, \mu)$ weakly mixing if

$$(2.5) \qquad \lim_{n \to \infty} n^{-1} \sum_{k=0}^{n-1} |\mu(A_n \cap B) - \mu(A_n)\mu(B)| = 0 \qquad (B \in \mathscr{F}).$$

It can be seen as in the proof of Theorem 2.1 that a weakly mixing sequence (A_n) contains a remotely trivial subsequence. But the converse is not true. There exist sequences (A_n) which contain a remotely trivial subsequence, but are not weakly mixing. The validity of (2.5) remains unchanged if μ is replaced by an equivalent measure. This follows from the Radon-Nikodym theorem and was also pointed out by Rényi [16] in the case where $\mu(A_n)$ converges. It follows that weak mixing for sequences of sets can be characterized in terms of the null sets of the measure space. A rather inelegant characterization can be given as follows: A strictly increasing sequence $\{n_k, k \in \mathbf{Z}^+\}$ is said to have positive lower density if $k^{-1}|\{n_i : 0 \leq n_i \leq k\}|$ does not converge to zero. A sequence $(A_n, n \in \mathbf{Z}^+)$ is weakly mixing if and only if for every strictly increasing sequence $\{n_k\}$ of positive lower density the sequence $(A_{n_k}, k \in \mathbf{Z}^+)$ contains a remotely trivial subsequence. We leave the proof to the reader. We mention that England and Martin [21] have given a characterization of weak mixing in terms of null sets. Like our characterization in Theorem 2.1 it holds only for transformations and not for arbitrary sequences of sets.

It seems natural to call a m.p. transformation T of a σ-finite measure space $(\Omega, \mathscr{F}, \bar{\mu})$ weakly mixing if for some probability measure ν equivalent to $\bar{\mu}$ (and hence for all such measures) each sequence $(T^{-n}A)$ ($A \in \mathscr{F}$) is weakly mixing in $(\Omega, \mathscr{F}, \nu)$. It is not difficult to show that a sequence (A_n) of sets in a probability space $(\Omega, \mathscr{F}, \mu)$ is weakly mixing if and only if

$$(2.6) \qquad \lim_{n \to \infty} n^{-1} \sum_{k=0}^{n-1} |\nu_1(A_n) - \nu_2(A_n)| = 0$$

for all pairs of probability measures $\nu_1, \nu_2 \ll \mu$. Therefore T in $(\Omega, \mathscr{F}, \bar{\mu})$ is weakly mixing if and only if

$$(2.7) \qquad \lim_{n \to \infty} n^{-1} \sum_{k=0}^{n-1} |\mu_1(T^{-n}A) - \mu_2(T^{-n}A)| = 0$$

for all $A \in \mathscr{F}$ and all probability measures $\mu_1, \mu_2 \ll \bar{\mu}$. This gives an interpretation of weak mixing which is analogous to (2.3). Mrs. Dowker has already used (2.7) (with $T^n A$ instead of $T^{-n} A$) as a definition of weak mixing in some unpublished work.

Using her Theorem 4 in [1] and the fact [9] that there exist ergodic invertible m.p. transformations T of an infinite measure space for which $T \times T$ is ergodic, she proved the existence of invertible weakly mixing transformations preserving an infinite measure. This is quite striking when it is compared with Theorem 3.1 in [10] which implies that completely mixing m.p. transformations in an infinite, σ-finite measure space necessarily are noninvertible. The notions of complete mixing as in (2.3) and (2.7) carry over to contractions in $\mathfrak{L}_1$. In this setting they have recently been studied in the thesis of Dr. M. Lin.

Let us now consider the announced characterization of transformations with discrete spectrum.

If $\mathfrak{G}$ is a set of measurable functions, the smallest σ-algebra with respect to which all $f \in \mathfrak{G}$ are measurable is called the σ-algebra generated by $\mathfrak{G}$ and denoted by $\sigma(\mathfrak{G})$. We shall need the following theorem, which is of independent interest.

THEOREM 2.2. *Let T be a m.p. transformation of a probability space $(\Omega, \mathscr{F}, \mu)$ and let $\mathfrak{E}$ be the set of eigenvectors of the isometry U_T. Then $\mathfrak{L}_2(\Omega, \sigma(\mathfrak{E}), \mu)$ is spanned by $\mathfrak{E}$ and $\sigma(\mathfrak{E})$ is an invariant sub-σ-algebra of $\mathscr{F}$.*

Proof. Let $\mathfrak{E}_1 = \{f \in \mathfrak{E} : |f| \equiv 1\}$. We first show $\sigma(\mathfrak{E}_1) = \sigma(\mathfrak{E})$. As U_T is an isometry each eigenvalue has absolute value 1. It follows that for $f \in \mathfrak{E}$ the sets

$$(2.8) \qquad \{\omega \in \Omega : \alpha_1 \leqq |f| < \alpha_2\} \qquad (0 \leqq \alpha_1 < \alpha_2)$$

are invariant. Invariant sets belong to $\sigma(\mathfrak{E}_1)$, since $1_A - 1_{A^c} \in \mathfrak{E}_1$ if A is invariant. For any $f \in \mathfrak{E}$, $\sigma(\{f\})$ is generated by sets of the form (2.8) together with sets of the form

$$(2.9) \qquad \{\omega \in \Omega : |f(\omega)| \neq 0, \beta_1 \leqq \operatorname{Arg} f(\omega) \leqq \beta_2\} \qquad (-\pi < \beta_1 < \beta_2 \leqq \pi).$$

On the invariant set $\{f \neq 0\}$ the set in (2.9) coincides with

$$\{\omega : |f(\omega)| \neq 0, \beta_1 \leqq \operatorname{Arg}(|f(\omega)|^{-1} f(\omega)) \leqq \beta_2\}.$$

Hence $\sigma(\{f\}) \subseteq \sigma(\mathfrak{E}_1)$ and therefore $\sigma(\mathfrak{E}) = \sigma(\mathfrak{E}_1)$.

Let K be the complex circle $\{z : |z| = 1\}$. Let g be a real-valued continuous function on K. The Stone-Weierstrass Theorem [6, p. 98] implies that for each $\varepsilon > 0$ there exist an $n \in N$ and $2n+1$ complex numbers a_k $(-n \leqq k \leqq n)$ such that $|\sum_{k=-n}^{n} a_k z^k - g(z)| < \varepsilon$ $(z \in K)$. Applying this result to m continuous functions $g_1, \ldots, g_m$ on K we find that for every $\varepsilon > 0$ there exists an n and a matrix $\{a_{ik}, i = 1, \ldots, m; |k| \leqq n\}$ such that

$$(2.10) \qquad |g_1(z_1) \cdot g_2(z_2) \cdots g_m(z_m) - (\sum a_{1k} z_1^k) \cdots (\sum a_{mk} z_m^k)| < \varepsilon$$

for $(z_1, \ldots, z_m) \in K^m$. (2.10) implies for $f_1, \ldots, f_m \in \mathfrak{E}_1$,

$$(2.11) \qquad \left\| g_1(f_1) \cdot g_2(f_2) \cdots g_m(f_m) - \left(\sum_{k_1} a_{1k_1} f_1^{k_1} \right) \cdots \left(\sum_{k_m} a_{mk_m} f_m^{k_m} \right) \right\| < \varepsilon.$$

For each choice of $k_1, \ldots, k_m$ the function $f_1^{k_1} \cdot f_2^{k_2} \cdots f_m^{k_m}$ belongs to $\mathfrak{E}_1$, because the product of two eigenfunctions is an eigenfunction. As $\varepsilon > 0$ was arbitrary (2.11) implies that the Hilbert space spanned by $\mathfrak{E}_1$ contains $\prod_{i=1}^m g_i(f_i)$ for arbitrary continuous g_i $(1 \le i \le m)$ on K. It then also contains $\prod_{i=1}^m 1_{B_i} \circ f_i$ for arbitrary Borel subsets B_i $(1 \le i \le m)$ of K. $\mathfrak{E}_1$ therefore spans $\mathfrak{L}_2(\Omega, \sigma(\mathfrak{E}), \mu)$.

It remains to prove $T^{-1}\sigma(\mathfrak{E}) = \sigma(\mathfrak{E})$ mod μ. Let $f \in \mathfrak{E}$, $U_T f = af$ and let B be a Borel subset of the complex plane. Then we have

$$T^{-1}\{\omega : f(\omega) \in B\} = \{\omega : f(\omega) \in a^{-1} \cdot B\}$$

and $a^{-1}B = \{z : az \in B\}$ is a Borel set of the complex plane. Hence $T^{-1}\sigma(\mathfrak{E}) \subseteq \sigma(\mathfrak{E})$. Assume $T^{-1}\sigma(\mathfrak{E}) \ne \sigma(\mathfrak{E})$ mod μ. Then there is an $h \ne 0$ in the orthogonal complement of $\mathfrak{L}_2(\Omega, T^{-1}\sigma(\mathfrak{E}), \mu)$ in $\mathfrak{L}_2(\Omega, \sigma(\mathfrak{E}), \mu)$. $U_T^k h$ belongs to the orthogonal complement of $\mathfrak{L}_2(\Omega, T^{-k-1}\sigma(\mathfrak{E}), \mu)$ in $\mathfrak{L}_2(\Omega, T^{-k}\sigma(\mathfrak{E}), \mu)$ for $k \ge 0$. Thus the vectors $U_T^k h$ are orthogonal. This is impossible since U_T has purely discrete spectrum in $\mathfrak{L}_2(\Omega, \sigma(\mathfrak{E}), \mu)$. $\quad \square$

LEMMA 2.1. *Let T be an ergodic m.p. transformation of a probability space $(\Omega, \mathscr{F}, \mu)$ and let $\mathscr{F}_0$ be a sub-σ-algebra of $\mathscr{F}$ with $T^{-1}\mathscr{F}_0 = \mathscr{F}_0$ (mod μ) and $\mathscr{F}_0 \ne \mathscr{F}$ (mod μ). Then there exists a set $A \in \mathscr{F}$ with $0 < \mu(A) < 1$ such that $\mu(A \cap B) = \mu(A)\mu(B)$ for all $B \in \mathscr{F}_0$.*

Proof. Let $\mu(A|\mathscr{F}_0)$ denote the conditional probability of A under $\mathscr{F}_0$, i.e. $\mu(A|\mathscr{F}_0)$ is defined mod μ by the requirements that it is $\mathscr{F}_0$-measurable and that

$$\int_B \mu(A|\mathscr{F}_0) \, d\mu = \mu(A \cap B) \qquad (B \in \mathscr{F}_0).$$

For each finite partition $\xi = \{A_1, \ldots, A_n\}$ of Ω into $\mathscr{F}$-measurable sets we define $m(\xi) = \max \{\mu(A_i|\mathscr{F}_0), 1 \le i \le n\}$. If η is a partition which is finer than ξ we have $m(\eta) \le m(\xi)$. Let $m(\mathscr{F}) = \inf \{m(\xi), \xi = \text{finite partition}\}$. The infimum is taken in the space $\mathfrak{L}_1(\Omega, \mathscr{F}_0, \mu)$. It follows from $T^{-1}\mathscr{F}_0 = \mathscr{F}_0$ that $\mu(A|\mathscr{F}_0) \circ T = \mu(T^{-1}A|\mathscr{F}_0)$ for all $A \in \mathscr{F}$. This implies $m(\mathscr{F}) \circ T = m(\mathscr{F})$, and now it follows from the ergodicity of T that $m(\mathscr{F})$ is T-invariant. $\mathscr{F}_0 \ne \mathscr{F}$ implies $m(\mathscr{F}) < 1$. The lemma now follows from Rohlin's theorem on cross-sections [17], if $(\Omega, \mathscr{F}, \mu)$ is a Lebesgue space. It was pointed out by Hanen and Neveu [4] (see also Neveu [15, Theorem 2]) that the relevant construction in §4 of [17] can also be carried out if Ω is not a Lebesgue space. In that case one has to work with conditional measures $\mu(A|\mathscr{F}_0)$ instead of measures μ_C defined on the fibers induced by the partition (in the sense of [17]) $\mathscr{F}_0$.

$(m(\mathscr{F}) \equiv \text{const.} < 1$ means the following in the context of [17]: if $m(\mathscr{F}) = 0$, each measure μ_C is nonatomic and for every α with $0 < \alpha < 1$ there is a set A with $\mu(A) = \alpha$ and $\mu(A \cap B) = \mu(A)\mu(B)$ $(B \in \mathscr{F}_0)$. If $m(\mathscr{F}) = \alpha > 0$ then each measure μ_C has a largest atom of size α. We may pick an atom of size α from each fiber C in such a way that the union A is $\mathscr{F}$-measurable.)

THEOREM 2.3. *Let T be an ergodic m.p. transformation of a probability space $(\Omega, \mathscr{F}, \mu)$. T has discrete spectrum if and only if there exists no $A \in \mathscr{F}$ with $0 < \mu(A) < 1$ for which the sequence $(T^{-n}A, n \in Z^+)$ has a remotely trivial subsequence.*

Proof. First assume that T has discrete spectrum. Then the orbit $(U_T^k f, k \in Z^+)$ of each $f \in \mathfrak{L}_2$ is conditionally norm-compact. (To see this just write f in the form $\sum_{i=1}^{\infty} C_i f_i$, where the f_i are normalized eigenvectors and $\sum_{i=1}^{\infty} |C_i|^2 < \infty$.) Let $A \in \mathscr{F}$ and a subsequence $n_1 < n_2 < \cdots$ of the integers be given. There exists a subsequence (m_i) of (n_i) for which the sequence $U_T^{m_i} 1_A$ converges in $\mathfrak{L}_2$. It necessarily converges to an indicator function 1_X with $\mu(X) = \mu(A)$. $\mu(T^{-m_i}A \triangle X) \to 0$ implies that X belongs to every σ-algebra $\mathscr{G}_m$ $(m \geq 0)$ generated by $(T^{-n_j}A, j \geq m)$. Thus $(T^{-n_i}A)$ can be remotely trivial only if $\mu(A) = 0$ or $\mu(A) = 1$. Note that we have not used the ergodicity of T for this part of the proof.

Next assume that T does not have discrete spectrum. Theorem 2.2 implies that the σ-algebra $\sigma(\mathfrak{E})$ generated by the set $\mathfrak{E}$ of eigenvectors of U_T is not the full σ-algebra $\mathscr{F}$. By Lemma 2.1 there exists a set $A \in \mathscr{F}$ with $0 < \mu(A) < 1$ such that A is independent of all $B \in \sigma(\mathfrak{E})$. The function $f = \mu(A^c)1_A - \mu(A)1_{A^c}$ is orthogonal to all eigenvectors of U_T. Proposition 1.2 implies $(U_T^k f, f) \leadsto 0$, and from this we get

$$\lim_{n \to \infty} n^{-1} \sum_{k=0}^{n-1} |\mu(T^{-k}A \cap A) - \mu(A)^2| = 0.$$

The sequence $(T^{-n}A, n \geq 0)$ is therefore a weakly mixing sequence of sets and contains a remotely trivial subsequence. $\square$

In [12] Kušnirenko has studied the following concept of entropy for subsequences: For a finite partition ξ of Ω and a strictly increasing subsequence $\{t_n\}$ of the integers let

$$h_{\{t_n\}}(T, \xi) = \limsup n^{-1} H\left(\bigvee_{i=1}^{n} T^{-t_i}\xi\right),$$

$$h_{\{t_n\}}(T) = \sup \{h_{\{t_n\}}(\xi, T), \xi = \text{finite partition}\}.$$

Here $H(\cdot)$ is the usual entropy of a partition, and $\bigvee_{i=1}^{n} T^{-t_i}\xi$ is the common refinement of the partitions $T^{-t_i}\xi$ $(1 \leq i \leq n)$. Using the invariant $h_{\{t_n\}}(T)$ it was possible to distinguish between transformations, for which all known isomorphy invariants were the same. Kušnirenko has proved the following theorem:

THEOREM 2.4 [12]. *T has discrete spectrum if and only if $h_{\{t_n\}}(T) = 0$ for each strictly increasing $\{t_n\}$.*

(Kušnirenko deals only with invertible transformations, but this theorem carries over to the noninvertible case without additional difficulties.) The "only if" part of the theorem is an immediate consequence of the compactness of the orbits $U_T^k f$. If T is ergodic the "if" part follows from Theorem 2.3. (If $(T^{-m_i}A)$ is remotely trivial, one can select a further subsequence $(T^{-t_i}A)$ such that for an arbitrarily small given $\varepsilon_{i+1} > 0$ the inequality $|\mu(T^{-t_{i+1}}A \cap B) - \mu(A)\mu(B)| < \varepsilon_{i+1}$ holds for all sets B in the field generated by the sets $T^{-t_j}A$ $(j=1,\dots,i)$.) It is easy to give examples which show that Theorem 2.3 would be false without the assumption of ergodicity. However, it seems likely that in the nonergodic case Theorem 2.3 can be modified in such a way that Theorem 2.4 would follow from it. It seems that T does not have discrete spectrum if and only if there exists a T-invariant set I of positive measure for which there exists an $A \subseteq I$ with $0 < \mu(A) < \mu(I)$ such that $(T^{-n}A)$ has a remotely trivial subsequence in the restriction of $(\Omega, \mathscr{F}, \mu)$ to I. Our aim has not been to give an alternative proof of Kušnirenko's result; therefore we have not further pursued this possible extension.

3. **Weakly independent partitions.** Let T be an invertible m.p. transformation, (i.e. an automorphism), of a probability space $(\Omega, \mathscr{F}, \mu)$. A partition (of Ω) is a finite collection of disjoint $\mathscr{F}$-measurable sets, the union of which is Ω. Finitely many partitions $\xi_1 = \{A_{11}, \dots, A_{1,n_1}\}, \dots, \xi_r = \{A_{r1}, \dots, A_{r,n_r}\}$ are called independent if for all $(i_1, i_2, \dots, i_r)$ with $1 \leqq i_j \leqq n_j$ $(j=1, \dots, r)$ the equation

$$(3.1) \qquad \mu\left(\bigcap_{j=1}^{r} A_{j,i_j}\right) = \prod_{|j=1}^{r} \mu(A_{j,i_j})$$

holds. Infinitely many partitions are called independent if every finite subset of them is independent. A partition $\xi = \{A_1, \dots, A_n\}$ is called weakly independent if there exists a strictly increasing sequence $0 = k_0 < k_1 < k_2 < \cdots$ of nonnegative integers such that the partitions $T^{-k_i}\xi$ $(i=0, 1, \dots)$ are independent. Recall that T was called 2-sided weakly mixing if for all $A, B, C \in \mathscr{F}$,

$$(3.2) \qquad \lim_{n \to \infty} n^{-1} \sum_{k=0}^{n-1} |\mu(T^{-k}A \cap B \cap T^k C) - \mu(A)\mu(B)\mu(C)| = 0,$$

and that T is called weakly mixing if (3.2) holds for all $A, B \in \mathscr{F}$ and for $C = \Omega$.

THEOREM 3.1. (a) *If T is 2-sided weakly mixing there exists for every partition $\xi = \{A_1, \dots, A_n\}$ and every $\varepsilon > 0$ a weakly independent partition $\bar{\xi} = \{\bar{A}_1, \dots, \bar{A}_n\}$ with*

$$(3.3) \qquad \sum_{i=1}^{n} \mu(A_i \triangle \bar{A}_i) < \varepsilon$$

and

$$(3.4) \qquad \mu(A_i) = \mu(\bar{A}_i) \qquad (i = 1, \dots, n).$$

(b) *If for every partition $\xi = \{A, A^c\}$ $(A \in \mathscr{F})$ and every $\varepsilon > 0$ there exists a partition $\bar{\xi} = \{\bar{A}, \bar{A}^c\}$ with $\mu(A \triangle \bar{A}) < \varepsilon$ and three integers k_0, k_1, k_2 such that $T^{k_0}\bar{\xi}$, $T^{k_1}\bar{\xi}$, $T^{k_2}\bar{\xi}$ are independent, then T is weakly mixing.*

Proof. (b) We start with the much simpler proof of part (b): If T is not ergodic, there exists an invariant set $E \in \mathscr{F}$ with $0 < \mu(E) < 1$ and it is clear that there is an $\varepsilon > 0$ such that for no $k_1 > 0$ and $\bar{E}$ with $\mu(E \triangle \bar{E}) < \varepsilon$ the partitions $\eta = \{\bar{E}, \bar{E}^c\}$ and $T^{k_1}\eta$ can be independent. Thus T is ergodic.

If T is not weakly mixing there exists a complex number $c \neq 1$ with $|c| = 1$ and a measurable complex valued function f on Ω with $|f| \equiv 1$ such that $f \circ T = cf$.

If $\gamma = (2\pi i)^{-1} \log c$ is rational, we write γ in the form $\gamma = rs^{-1}$ with $(r, s) = 1$ and $s \geqq 2$. Then the sets

$$(3.5) \qquad A_i = \{\omega \in \Omega : (2\pi i)^{-1} \log f(\omega) \in [(i-1)s^{-1}, is^{-1}[\}$$

form a partition of Ω and we have $T^{-1}A_i = A_{i-r(\text{mod } s)}$. Therefore the sets A_i move cyclically under T and it is clear that there cannot exist sets $\bar{A}$ arbitrarily close to $A = A_1$ having the property that at least for one $k > 0$ the equality $\mu(\bar{A} \cap T^k\bar{A}) = \mu(\bar{A})^2$ is satisfied.

Next we consider the case where γ is irrational. For any Borel set $B \subseteq [0, 1[$ define

$$A_B = \{\omega \in \Omega : (2\pi i)^{-1} \log f(\omega) \in B\} \quad \text{and} \quad \lambda(B) = \mu(A_B).$$

Let $T_\gamma x = x + \gamma \pmod 1$ ($x \in [0, 1[$). We obtain $T^{-1}A_B = A_{T_\gamma^{-1}B}$, and this implies that λ is an invariant measure for T_γ. Thus λ must be the Lebesgue-measure.

Let $J = [0, 1/10[$, $\varepsilon = 1/10000$ and $A = A_J$. If $\bar{A} \in \mathscr{F}$ is such that $\mu(A \triangle \bar{A}) < \varepsilon$ and $0 < k_1 < k_2$ are such that $\bar{A}, T^{k_1}\bar{A}, T^{k_2}\bar{A}$ are independent, then we get

$$1/100 - 2\varepsilon < \mu(\bar{A} \cap T^{k_i}\bar{A}) < 1/100 + 2\varepsilon$$

and hence

$$(3.6) \qquad 1/100 - 4\varepsilon < \mu(A \cap T^{k_i}A) < 1/100 + 4\varepsilon \qquad (i = 1, 2).$$

Similarly we get

$$(3.7) \qquad 1/100 - 4\varepsilon < \mu(T^{k_1}A \cap T^{k_2}A) < 1/100 + 4\varepsilon.$$

The considerations above now imply

$$(3.8) \qquad |\lambda(J \cap T_\gamma^{-k_i}J) - 1/100| < 4\varepsilon \qquad (i = 1, 2)$$

and

$$(3.9) \qquad |\lambda(T_\gamma^{-k_1}J \cap T_\gamma^{-k_2}J) - 1/100| < 4\varepsilon.$$

Since J is an interval and $\lambda(J) = 1/10$ the equations (3.8) and (3.9) cannot be simultaneously satisfied. $\quad \square$

For the proof of part (a) we require the following elementary lemma:

LEMMA 3.1. *Let* $\Gamma = (\gamma_{\rho\sigma})$ $(\rho = 1, \ldots, r; \sigma = 1, \ldots, s)$ *be a matrix with real entries such that*

$$(3.10) \qquad \sum_{\rho=1}^{r} \gamma_{\rho\sigma} = 0 \quad (1 \leqq \sigma \leqq s) \quad \text{and} \quad \sum_{\sigma=1}^{s} \gamma_{\rho\sigma} = 0 \quad (1 \leqq \rho \leqq r).$$

Then there exists a unique matrix $(\eta_{\lambda\nu})$ $(1 \leq \lambda \leq r-1, 1 \leq \nu \leq s-1)$ *such that* $\Gamma = \sum_{\lambda=1}^{r-1} \sum_{\nu=1}^{s-1} \eta_{\lambda\nu} \Delta^{\lambda\nu}$ *where* $\Delta^{\lambda\nu}$ *is the matrix* $(\delta_{ij}^{\lambda\nu})$ $(1 \leq i \leq r, 1 \leq j \leq s)$ *given by*

$$\begin{aligned}
\delta_{ij}^{\lambda\nu} &= 1 && (i = \lambda, j = \nu), (i = \lambda+1, j = \nu+1), \\
&= -1 && (i = \lambda, j = \nu+1), (i = \lambda+1, j = \nu), \\
&= 0 && \textit{otherwise.}
\end{aligned}$$

The coefficients $\eta_{\lambda\nu}$ *in this representation satisfy*

(3.11) $$\max_{\lambda, \nu} |\eta_{\lambda\nu}| \leq \min (2^s r^s, 2^r s^r) \max_{\rho, \sigma} |\gamma_{\rho\sigma}|.$$

Sketch of proof. Let $\eta_{11} = \gamma_{11}$, then $\Gamma' = \Gamma - \eta_{11}\Delta^{11}$ is a matrix $(\gamma'_{\rho,\sigma})$ with the property (3.10) and with $\gamma'_{11} = 0$. Next let $\eta_{12} = \gamma'_{12}$ and $\Gamma'' = \Gamma' - \eta_{12}\Delta^{12}$. Now we have $\gamma''_{11} = 0$ and $\gamma''_{12} = 0$. Successively reduce the elements of the first row to zeros, then start with the second row, etc. This proves the existence of the representation and the same argument yields the uniqueness: just observe that η_{11} has to be γ_{11} since $\delta_{11}^{\lambda\nu} = 0$ for all λ, ν except $\lambda = 1$, $\nu = 1$. Then we see that η_{12} has to be γ'_{12}, etc. To prove (3.11) with $\min (2^s r^s, 2^r s^r)$ replaced by $2^r s^r$ assume that some $\varepsilon > 0$ is a bound for the elements $|\gamma_{\rho\sigma}|$. Then $(s-1)\varepsilon$ is a bound for the elements $|\eta_{1\nu}|$. If ε_j is a bound for the elements $|\eta_{j\nu}|$, then $2(s-1)\varepsilon_j + \varepsilon$ is a bound for the elements $|\eta_{j+1,\nu}|$ by our construction. Induction and the remark that the role of rows and columns can be interchanged yield (3.11). $\square$

(a) We now turn to the proof of part (a). Let T be 2-sided weakly mixing and let $\varepsilon > 0$ and a partition $\xi = \{A_1, \ldots, A_n\}$ of Ω be given. We construct a partition $\bar{\xi} = \{\bar{A}_1, \ldots, \bar{A}_n\}$ with $\mu(A_i) = \mu(\bar{A}_i)$ and $\sum_{i=1}^n \mu(A_i \triangle \bar{A}_i) < \varepsilon$, and a sequence $0 = k_0 < k_1 < \cdots$ of integers such that the partitions $T^{k_0}\bar{\xi}$, $T^{k_1}\bar{\xi}$, $T^{k_2}\bar{\xi}$, ... are independent. (In the definition of weak independence of $\bar{\xi}$ independence of the partitions $T^{-k_i}\bar{\xi}$ is required. This seems natural in view of the relation to weakly wandering vectors. Proving independence of $T^{k_i}\bar{\xi}$, however, saves many minus signs in the proof. As T is invertible the same construction can be applied to T^{-1}, to prove the theorem as stated.) Clearly we may assume that $n \geq 2$ and that the sets A_i have strictly positive measure.

Let $a > 0$ be such that $\mu(A_i) > a$ $(i = 1, \ldots, n)$. Let B_{ij} $(1 \leq i, j \leq n, i \neq j)$ be sets of positive measure such that

(3.12) $$B_{ij} \subseteq A_i, \qquad B_{ij} \cap B_{ik} = \varnothing \ (j \neq k), \qquad \mu\left(A_i \backslash \bigcup_{j \neq i} B_{ij}\right) > a,$$

and

(3.13) $$\mu\left(\bigcup_{i,j} B_{ij}\right) < \varepsilon.$$

The desired weakly independent partition $\bar{\xi}$ will be constructed inductively. We start with $\xi^{(0)} = \{A_1^{(0)}, \ldots, A_n^{(0)}\} = \xi$ and $k_0 = 0$. After the tth step of the construction we have determined k_t and an approximation $\xi^{(t)} = \{A_1^{(t)}, \ldots, A_n^{(t)}\}$ of ξ in such a way that the partitions $T^{k_0}\xi^{(t)}$, $T^{k_1}\xi^{(t)}$, ..., $T^{k_t}\xi^{(t)}$ are independent.

We shall specify disjoint sets $D_{ij}^{(t)} \subseteq B_{ij} \setminus \bigcup_{\tau=0}^{t-1} D_{ij}^{(\tau)}$ and $A_i^{(t+1)}$ will be obtained from $A_i^{(t)}$ by

$$(3.14) \qquad A_i^{(t+1)} = \left(A_i^{(t)} \cup \bigcup_{j \neq i} D_{ji}^{(t+1)} \right) \setminus \bigcup_{j \neq i} D_{ij}^{(t+1)}.$$

For $D_{ij}^{(0)}$ we may take the empty set.

It will be convenient to consider also subpartitions, i.e. collections of finitely many disjoint $\mathscr{F}$-measurable sets. (Their union is not necessarily the whole space Ω.) The most important subpartition is $\xi^* = \{A_1^*, \ldots, A_n^*\}$, where $A_i^* = A_i \setminus \bigcup_{j \neq i} B_{ij}$. The refinement $\theta_1 \vee \theta_2 \vee \cdots \vee \theta_p$ of p subpartitions $\theta_j = \{E_{j1}, \ldots, E_{j,l_j}\}$ $(1 \leq j \leq p)$ is the subpartition consisting of the sets $E_{1,\lambda_1} \cap E_{2,\lambda_2} \cap \cdots \cap E_{p,\lambda_p}$ $(1 \leq \lambda_j \leq l_j,$ $1 \leq j \leq p)$. We shall say that $\theta_1 \vee \cdots \vee \theta_p$ has positive atoms on $F \in \mathscr{F}$ when all sets $F \cap E_{1,\lambda_1} \cap \cdots \cap E_{p,\lambda_p}$ have positive measure. We denote the maximal common lower bound for the measure of all these sets by $l(F \cap (\theta_1 \vee \cdots \vee \theta_p))$.

We shall carry out the construction in such a way that for fixed t the measure of all sets $D_{ij}^{(t)}$ $(i \neq j)$ is the same. Since the sets $D_{ij}^{(\tau)}$ $(\tau = 0, 1, \ldots)$ are disjoint we obtain

$$(3.15) \qquad \mu(A_i^{(t)}) = \mu(A_i) \qquad (i = 1, \ldots, n; t \geq 0)$$

from (3.14). It will now be sufficient to show that we can find k_{t+1} $(t \geq 0)$ and sets $D_{ij}^{(t+1)}$ as above such that $T^{k_0}\xi^{(t+1)}, T^{k_1}\xi^{(t+1)}, \ldots, T^{k_{t+1}}\xi^{(t+1)}$ are independent. The partition ξ can then be defined by $\bar{A}_i = \lim_{t \to \infty} A_i^{(t)}$ where the limit is taken in the metric space of all equivalence classes of measurable sets with the metric $d(F, G) = \mu(F \triangle G)$. The equations (3.13) and (3.15) imply (3.3) and (3.4).

In the course of the construction of the sets $D_{ij}^{(t+1)}$ we shall need auxiliary sets $B_{ij}^{(t)}$ and $C_{ij}^{(t)}$ $(i \neq j)$ with

$$(3.16) \qquad B_{ij}^{(t)} \supseteq C_{ij}^{(t)} \supseteq D_{ij}^{(t)}$$

and partitions χ_t and ψ_t. They are determined in the tth step of the construction. We start with $\chi_0 = \psi_0 = \{\Omega\}$. Let $B_{ij}^{(0)} = C_{ij}^{(0)}$ be subsets of B_{ij}, all having the same positive measure $b_0 > 0$.

The numbers $k_0 < k_1 < \cdots < k_t$ will be determined in such a way that

$$(3.17) \qquad l\left(\Omega \cap \bigvee_{\tau=-t}^{t} T^{k_\tau}\xi^* \right) \geq a^{2t+1},$$

where, for simplicity, we set $k_{-\tau} = -k_\tau$ $(\tau \geq 0)$. For $t \geq 1$ the partitions ψ_t and χ_t are defined by

$$(3.18) \qquad \psi_t^+ = T^{k_t}\left(\bigvee_{\tau=-t+1}^{t-1} T^{k_\tau}\xi^* \right),$$

$$(3.19) \qquad \psi_t^- = T^{-k_t}\left(\bigvee_{\tau=-t+1}^{t-1} T^{k_\tau}\xi^* \right) = T^{-2k_t}\psi_t^+,$$

$$(3.20) \qquad \psi_t = \psi_t^+ \vee \psi_t^-,$$

$$(3.21) \qquad \chi_t = \chi_{t-1} \vee \psi_t.$$

We shall also have

$$(3.22) \qquad 0 < \inf\{l((C_{ij}^{(t)} \backslash D_{ij}^{(t)}) \cap \chi_t) : 1 \leq i, j \leq n, i \neq j\}.$$

The infimum on the right side of (3.22) is denoted by c_t. The sets $C_{ij}^{(t)}$ will be such that the union $C^{(t)} = \bigcup_{i,j} C_{ij}^{(t)}$ has disjoint images $T^{k_\tau} C^{(t)}$ ($\tau = 0, \ldots, t$). The number m_t of atoms of ψ_t^+ is n^{2t-1} and the number n_t of atoms of χ_t is n^{2t^2}. We are now ready to describe step $t+1$ of the construction ($t \geq 0$):

   Step $t+1$. First determine a positive number $b_{t+1} < c_t$ so small that

$$(3.23) \qquad (t+1)8(n^2 n_t b_{t+1})^2 < a^{4t+2} b_{t+1}.$$

By (3.22) the atoms of $(C_{ij}^{(t)} \backslash D_{ij}^{(t)}) \cap \chi_t$ have measure greater than b_{t+1} and we can pick $B_{ij}^{(t+1)} \subseteq C_{ij}^{(t)} \backslash D_{ij}^{(t)}$ such that $B_{ij}^{(t+1)}$ contains from each atom of χ_t a set of measure b_{t+1}. Let $B^{(t+1)} = \bigcup_{i,j} B_{ij}^{(t+1)}$.

   Next determine a number $\alpha_{t+1} > 0$ such that

$$(3.24) \qquad 2^{n+3} n^{(t+1)n+t+2} \alpha_{t+1} < b_{t+1} a^{4t+2}.$$

As the partitions $\xi^{(t)}, \ldots, T^{k_t} \xi^{(t)}$ are independent and as T is 2-sided weakly mixing, there exists an integer $k_{t+1} > k_t$ such that

$$(3.25) \qquad \left| \mu\left(\bigcap_{\tau=0}^{t+1} T^{k_\tau} A_{i_\tau}^{(t)}\right) - \prod_{\tau=0}^{t+1} \mu(A_{i_\tau}) \right| < \alpha_{t+1} \qquad (1 \leq i_\tau \leq n, \tau \leq t+1)$$

and

$$(3.26) \qquad \mu(T^{k_\tau} B^{(t+1)} \cap T^{k_{t+1}} B^{(t+1)}) < 2\mu(B^{(t+1)})^2 \qquad (0 \leq \tau \leq t)$$

and

$$(3.27) \qquad l(B_{ij}^{(t+1)} \cap (\chi_t \vee \psi_{t+1})) \geq 2^{-1} b_{t+1} a^{4t+2} \qquad (i \neq j)$$

and

$$(3.28) \qquad l\left(\Omega \cap \bigvee_{\tau=0}^{t+1} T^{k_\tau} \xi^*\right) \geq \alpha^{4t+2}.$$

   We shall now shrink the sets $B_{ij}^{(t+1)}$ a little in order to obtain disjointness of the images. Let

$$(3.29) \qquad C^{(t+1)} = B^{(t+1)} \backslash T^{-k_{t+1}}\left(\bigcup_{\tau=0}^{t} T^{k_\tau} B^{(t+1)} \cap T^{k_{t+1}} B^{(t+1)}\right).$$

The set $B^{(t+1)}$ is a subset of $C^{(t)}$. Therefore the sets $T^{k_\tau} B^{(t+1)}$ ($\tau = 0, \ldots, t$) are disjoint. Together with (3.29) we find that the sets $T^{k_\tau} C^{(t+1)}$ ($0 \leq \tau \leq t+1$) are disjoint. The sets $C_{ij}^{(t+1)}$ are defined by

$$(3.30) \qquad C_{ij}^{(t+1)} = B_{ij}^{(t+1)} \cap C^{(t+1)}.$$

It follows from (3.26) and (3.29) that

$$(3.31) \qquad \mu(B^{(t+1)} \backslash C^{(t+1)}) < 2(t+1)\mu(B^{(t+1)})^2 = 2(t+1)(n^2 n_t b_{t+1})^2.$$

Thus, (3.23) and (3.27) imply

$$(3.32) \qquad l(C_{ij}^{(t+1)} \cap (\chi_t \vee \psi_{t+1})) \geqq 4^{-1} b_{t+1} a^{4t+2} \qquad (i \neq j).$$

If we define $D_{ij}^{(t+1)}$ in such a way that it is a subset of $C_{ij}^{(t+1)}$ then the sets $T^{k_\tau} D_{ij}^{(t+1)}$ $(0 \leqq \tau \leqq t+1)$ are disjoint, too, and we obtain from (3.14) for arbitrary $(i_0, \ldots, i_{t+1}) \in \{1, 2, \ldots, n\}^{t+2}$ the equation

$$(3.33) \qquad \mu\left(\bigcap_{\tau=0}^{t+1} T^{k_\tau} A_{i_\tau}^{(t+1)}\right) = \mu\left(\bigcap_{\tau=0}^{t+1} T^{k_\tau} A_{i_\tau}^{(t)}\right) + \sum_{\sigma=0}^{t+1} R_\sigma(i_0, \ldots, i_{t+1})$$

where

$$(3.34) \qquad \begin{aligned} &R_\sigma(i_0, \ldots, i_{t+1}) \\ &= \sum_{j \neq i_\sigma} \mu\left(T^{k_\sigma} D_{j,i_\sigma}^{(t+1)} \cap \bigcap_{\tau=0; \tau \neq \sigma}^{t+1} T^{k_\tau} A_{i_\tau}^{(t)}\right) - \sum_{j \neq i_\sigma} \mu\left(T^{k_\sigma} D_{i_\sigma,j}^{(t+1)} \cap \bigcap_{\tau=0; \tau \neq \sigma}^{t+1} T^{k_\tau} A_{i_\tau}^{(t)}\right). \end{aligned}$$

Let

$$(3.35) \qquad \gamma(i_0, \ldots, i_{t+1}) = \mu\left(\bigcap_{\tau=0}^{t+1} T^{k_\tau} A_{i_\tau}^{(t)}\right) - \prod_{\tau=0}^{t+1} \mu(A_{i_\tau}^{(t)}).$$

It will now be sufficient to pick the sets $D_{ij}^{(t+1)} \subseteq C_{ij}^{(t+1)}$ in such a way that

$$(3.36) \qquad \mu(D_{ij}^{(t+1)}) = \mu(D_{ji}^{(t+1)}) \qquad (i \neq j)$$

and

$$(3.37) \qquad \gamma(i_0, \ldots, i_{t+1}) = -\sum_{\sigma=0}^{t+1} R_\sigma(i_0, \ldots, i_{t+1})$$

for all $(i_0, \ldots, i_{t+1})$ and such that enough space is left in the atoms of $\chi_{t+1} = \chi_t \vee \psi_{t+1}$ on each $C_{ij}^{(t+1)} \setminus D_{ij}^{(t+1)}$ for the construction to be continued.

It follows from the independence of the sets $T^{k_\tau} A_{i_\tau}^{(t)}$ $(0 \leqq \tau \leqq t)$ that the numbers $\gamma(i_0, \ldots, i_{t+1})$ satisfy the equations

$$(3.38) \qquad \sum_{i_{t+1}=1}^{n} \gamma(i_0, \ldots, i_t, i_{t+1}) = 0$$

and

$$(3.39) \qquad \sum_{i_0=1}^{n} \sum_{i_1=1}^{n} \cdots \sum_{i_t=1}^{n} \gamma(i_0, \ldots, i_t, i_{t+1}) = 0.$$

Let us assume that the $(t+1)$-tuples $(i_0, \ldots, i_t)$ $(1 \leqq i_\tau \leqq n)$ are written down in a lexicographical order. Let $\pi(\rho)$ denote the $(t+1)$-tuple $(i_0, \ldots, i_t)$ which appears on the ρth place of this list. (Of course the mapping π depends on t and should be denoted by π_t. But since we describe only step $(t+1)$ of the construction, it will not lead to confusion, if the index t is dropped. We shall tacitly use this policy concerning other symbols.)

By (3.38) and (3.39) we can apply Lemma 3.1 to the matrix $\Gamma = (\gamma_{\rho\sigma})$ given by $\gamma_{\rho\sigma} = \gamma(\pi(\rho), \sigma)$ ($\rho = 1, 2, \ldots, r = n^{t+1}$; $\sigma = 1, \ldots, s = n$). We obtain a representation

$$(3.40) \qquad \Gamma = \sum_{\lambda=1}^{r-1} \sum_{\nu=1}^{s-1} \eta_{\lambda\nu} \Delta^{\lambda\nu}.$$

By (3.25) we have $\max_{\rho,\sigma} |\gamma(\pi(\rho), \sigma)| < \alpha_{t+1}$. This and (3.11) imply

$$(3.41) \qquad \max_{\lambda,\nu} |\eta_{\lambda\nu}| < 2^n n^{(t+1)n} \alpha_{t+1}.$$

The matrix $(\gamma_{\rho\sigma})$ represents the "errors", i.e. the deviations from independence, which occur at this stage of the construction. We have represented the error matrix as a sum of error-matrices of a very special form. The components of this sum will now be corrected individually.

For each pair (λ, ν) with $1 \leq \lambda \leq r-1$, $1 \leq \nu \leq n-1$ we shall pick sets $E_{ij}(\lambda, \nu)$ $\subseteq C_{ij}^{(t+1)}$ ($i \neq j$) in a way described below. All these sets will be disjoint. (For example pick $E_{ij}(1, 1)$ first, then pick $E_{ij}(1, 2)$ in $C_{ij}^{(t+1)} \backslash E_{ij}(1, 1)$, etc. There will always be enough space left over for picking the next set $E_{ij}(\lambda, \nu)$.) Having picked all sets $E_{ij}(\lambda, \nu)$ we define $D_{ij}^{(t+1)}$ by

$$(3.42) \qquad D_{ij}^{(t+1)} = \bigcup_{\lambda=1}^{r-1} \bigcup_{\nu=1}^{n-1} E_{ij}(\lambda, \nu).$$

We shall pick the sets $E_{ij}(\lambda, \nu)$ in such a way that for fixed (λ, ν) the sets $E_{ij}(\lambda, \nu)$ have the same measure for all pairs (i, j) with $i \neq j$ and such that for all atoms F of the subpartition $\chi_t \vee \psi_{t+1}^+$ the measure of $E_{ij}(\lambda, \nu) \cap F$ is the same. Moreover $E_{ij}(\lambda, \nu)$ shall be contained in the union of the atoms of $\chi_{t+1} = \chi_t \vee \psi_t$.

Before proceeding with the specification of the sets we show that these properties of the sets $E_{ij}(\lambda, \nu)$ imply (3.36) and

$$(3.43) \qquad R_\sigma(i_0, \ldots, i_{t+1}) = 0 \qquad (0 \leq \sigma \leq t; 1 \leq i_\tau \leq n \ (0 \leq \tau \leq t+1)).$$

The equation (3.36) follows from the disjointness of the sets $E_{ij}(\lambda, \nu)$. Let us check (3.43) for $\sigma = 1$: We have

$$R_1(i_0, \ldots, i_t, i_{t+1})$$

$$(3.44) \qquad = \sum_{j \neq i_1} \mu(D_{j,i_1}^{(t+1)} \cap T^{k_0-k_1} A_{i_0}^{(t)} \cap T^{k_2-k_1} A_{i_2}^{(t)} \cap \cdots \cap T^{k_{t+1}-k_1} A_{i_{t+1}}^{(t)})$$

$$- \sum_{j \neq i_1} \mu(D_{i_1,j}^{(t+1)} \cap T^{k_0-k_1} A_{i_0}^{(t)} \cap T^{k_2-k_1} A_{i_2}^{(t)} \cap \cdots \cap T^{k_{t+1}-k_1} A_{i_{t+1}}^{(t)}).$$

This is equal to the same expression with all $A_i^{(t)}$ replaced by A_i^* because the sets $E_{ij}(\lambda, \nu)$ are contained in the union U_{t+1} of the atoms of χ_{t+1}. The intersection of $E_{ij}(\lambda, \nu)$ with $C_{ij}^{(t+1)} \cap G$ has the same measure for each atom G of

$$(3.45) \qquad T^{k_0-k_1} \xi^* \vee T^{k_2-k_1} \xi^* \vee \cdots \vee T^{k_{t+1}-k_1} \xi^*$$

because these atoms are unions of atoms of $\chi_t \vee \psi_{t+1}^+$ and because for each G the number of atoms of $\chi_t \vee \psi_{t+1}^+$ occurring in the union is the same. It follows that to

each term in the first sum of (3.44) there corresponds an equally large term in the second sum. This proves (3.43) for $\sigma = 1$. For $\sigma = 0, 2, \ldots, t$ the proof of (3.43) is essentially the same.

For short let X_ρ denote the set $T^{k_0}A_{i_0}^{(t)} \cap T^{k_1}A_{i_1}^{(t)} \cap \cdots \cap T^{k_t}A_{i_t}^{(t)}$ for which $\pi(\rho) = (i_0, \ldots, i_t)$. Then the definition (3.34) of R_{t+1} has the form

$$
\begin{aligned}
R_{t+1}&(\pi(\rho), i_{t+1}) \\
(3.46) \qquad &= \sum_{j \neq i_{t+1}} \mu(D_{j, i_{t+1}}^{(t+1)} \cap T^{-k_{t+1}}X_\rho) - \sum_{j \neq i_{t+1}} \mu(D_{i_{t+1}, j}^{(t+1)} \cap T^{-k_{t+1}}X_\rho).
\end{aligned}
$$

Let (λ, ν) with $1 \leq \lambda \leq r-1$, $1 \leq \nu \leq n-1$ be given.

Case 1. $\eta_{\lambda\nu} \geq 0$. For $i = \nu$, $j = \nu + 1$ pick $E_{ij}(\lambda, \nu) \subseteq C_{ij}^{(t+1)} \cap U_{t+1}$ such that

$$(3.47) \qquad \mu(E_{ij}(\lambda, \nu) \cap T^{-k_{t+1}}X_\lambda) = \eta_{\lambda\nu},$$

$$(3.48) \qquad \mu(E_{ij}(\lambda, \nu) \cap T^{-k_{t+1}}X_\rho) = 0 \qquad (\rho \neq \lambda)$$

and such that $\mu(E_{ij}(\lambda, \nu) \cap T^{-k_{t+1}}X_\rho \cap F)$ is the same for all atoms F of $\chi_t \vee \psi_{t+1}^+$.

For $i = \nu + 1$, $j = \nu$ pick $E_{ij}(\lambda, \nu) \subseteq C_{ij}^{(t+1)} \cap U_{t+1}$ such that

$$(3.49) \qquad \mu(E_{ij}(\lambda, \nu) \cap T^{-k_{t+1}}X_{\lambda+1}) = \eta_{\lambda\nu},$$

$$(3.50) \qquad \mu(E_{ij}(\lambda, \nu) \cap T^{-k_{t+1}}X_\rho) = 0 \qquad (\rho \neq \lambda+1)$$

and such that $\mu(E_{ij}(\lambda, \nu) \cap T^{-k_{t+1}}X_{\lambda+1} \cap F)$ is the same for all atoms F of $\chi_t \vee \psi_{t+1}^+$.

For all other pairs (i, j) with $1 \leq i, j \leq n$, $i \neq j$ pick $E_{ij}(\lambda, \nu) \subseteq C_{ij}^{(t+1)} \cap U_{t+1}$ such that $\mu(E_{ij}(\lambda, \nu)) = \eta_{\lambda\nu}$ and such that $\mu(E_{ij}(\lambda, \nu) \cap G)$ is the same for all atoms G of χ_{t+1}.

Case 2. $\eta_{\lambda\nu} < 0$. For $i = \nu$, $j = \nu + 1$ pick $E_{ij}(\lambda, \nu) \subseteq C_{ij}^{(t+1)} \cap U_{t+1}$ such that

$$\mu(E_{ij}(\lambda, \nu) \cap T^{-k_{t+1}}X_{\lambda+1}) = |\eta_{\lambda\nu}|,$$

$$\mu(E_{ij}(\lambda, \nu) \cap T^{-k_{t+1}}X_\rho) = 0 \qquad (\rho \neq \lambda+1)$$

and such that $\mu(E_{ij}(\lambda, \nu) \cap T^{-k_{t+1}}X_{\lambda+1} \cap F)$ is the same for all atoms F of $\chi_t \vee \psi_{t+1}^+$.

For $i = \nu + 1$, $j = \nu$ pick $E_{ij}(\lambda, \nu) \subseteq C_{ij}^{(t+1)} \cap U_{t+1}$ such that

$$\mu(E_{ij}(\lambda, \nu) \cap T^{-k_{t+1}}X_\lambda) = |\eta_{\lambda\nu}|,$$

$$\mu(E_{ij}(\lambda, \nu) \cap T^{-k_{t+1}}X_\rho) = 0 \qquad (\rho \neq \lambda)$$

and such that $\mu(E_{ij}(\lambda, \nu) \cap T^{-k_{t+1}}X_\lambda \cap F)$ is the same for all atoms F of $\chi_t \vee \psi_{t+1}^+$.

For all other pairs (i, j) with $1 \leq i, j \leq n$, $i \neq j$ pick $E_{ij}(\lambda, \nu) \subseteq C_{ij}^{(t+1)} \cap U_{t+1}$ such that $\mu(E_{ij}(\lambda, \nu)) = |\eta_{\lambda\nu}|$ and such that $\mu(E_{ij}(\lambda, \nu) \cap G)$ is the same for all atoms G of χ_{t+1}.

This completes the definition of the sets $E_{ij}(\lambda, \nu)$ and therefore of the sets $D_{ij}^{(t+1)}$. Let us make sure that the sets needed for picking $E_{ij}(\lambda, \nu)$ have been large enough: We have picked $(r-1)(n-1) \leq n^{(t+1)} \cdot n$ sets $E_{ij}(\lambda, \nu)$, each of them having measure $|\eta_{\lambda\nu}|$. From (3.41) and (3.24) we get

$$(3.51) \qquad \mu(D_{ij}^{(t+1)}) \leq n \cdot n^{t+1} \cdot 2^n \cdot n^{(t+1)n} \alpha_{t+1} < 2^{-3} b_{t+1} a^{4t+2}.$$

On the other hand (3.32) says that each atom G of χ_{t+1} intersects $C_{ij}^{(t+1)}$ in a set of measure greater or equal to $4^{-1}b_{t+1}a^{4t+2}$. It follows that the construction is possible and that

$$(3.52) \qquad l((C_{ij}^{(t+1)}\backslash D_{ij}^{(t+1)}) \cap \chi_{t+1}) \geq 2^{-3}b_{t+1}a^{4t+2}.$$

This is condition (3.22) with t replaced by $t+1$.

It remains to verify (3.37), which in view of $R_0 = R_1 = \cdots = R_t = 0$ now is equivalent to

$$(3.53) \qquad -R_{t+1}(i_0, \ldots, i_{t+1}) = \gamma(i_0, \ldots, i_{t+1}).$$

Let us write $i_{t+1} = u$ and $(i_0, \ldots, i_t) = \pi(\theta)$. Then we have

$$-R_{t+1}(i_0, \ldots, i_{t+1}) = \sum_{v \neq u} (\mu(D_{uv}^{(t+1)} \cap T^{-k_{t+1}}X_\theta) - \mu(D_{vu}^{(t+1)} \cap T^{-k_{t+1}}X_\theta))$$

$$= \sum_{\lambda, v} \left\{ \sum_{v \neq u} (\mu(E_{uv}(\lambda, v) \cap T^{-k_{t+1}}X_\theta) - \mu(E_{vu}(\lambda, v) \cap T^{-k_{t+1}}X_\theta)) \right\}.$$

The expression inside $\{\ \}$ is denoted by $\beta_{\theta, u}(\lambda, v)$. This quantity is now determined for fixed (λ, v) with $1 \leq \lambda \leq n-1$, $1 \leq v \leq n-1$:

Case 1. $\eta_{\lambda v} \geq 0$.

(a) $\theta = \lambda$, $u = v$. For $v \neq u+1$ we have

$$\mu(E_{uv}(\lambda, v) \cap T^{-k_{t+1}}X_\theta) = \mu(E_{vu}(\lambda, v) \cap T^{-k_{t+1}}X_\theta)$$

so that all terms of the sum which correspond to $v \neq u+1$ cancel. The term corresponding to $v = u+1 = v+1$ is $\eta_{\lambda v} - 0$ by (3.47) and (3.50). Thus $\beta_{\theta, u}(\lambda, v) = \eta_{\lambda v}$.

(b) $\theta = \lambda+1$, $u = v$. As before the terms for $v \neq u+1$ cancel. The equations (3.48) and (3.49) yield $\beta_{\theta, u}(\lambda, v) = -\eta_{\lambda v}$.

(c) $\theta = \lambda$, $u = v+1$. Only the term with $v = v = u-1$ remains. (3.50) and (3.47) give $\beta_{\theta, u}(\lambda, v) = 0 - \eta_{\lambda v} = -\eta_{\lambda v}$.

(d) $\theta = \lambda+1$, $u = v+1$. Again only the terms with $v = v = u-1$ remain. (3.49) and (3.48) give $\beta_{\theta, u}(\lambda, v) = \eta_{\lambda v} - 0$.

In all other cases all terms cancel. We have therefore proved that the matrix $(\beta_{\theta, u}(\lambda, v))$ has the form $\eta_{\lambda v}\Delta^{\lambda v}$.

Case 2. $\eta_{\lambda v} < 0$. Also in this case $(\beta_{\theta, u}(\lambda, v)) = \eta_{\lambda v}\Delta^{\lambda v}$ can be obtained by systematically checking all cases.

Equation (3.40) now implies

$$\sum_{\lambda, v} \beta_{\theta, u}(\lambda, v) = \gamma(\pi(\theta), u) = \gamma(i_0, \ldots, i_{t+1}).$$

This completes the proof of the independence of $\xi^{(t+1)}, T^{k_1}\xi^{(t+1)}, \ldots, T^{k_{t+1}}\xi^{(t+1)}$ and, therefore, the proof of the theorem. □

It seems clear that essentially the same construction can be used to obtain a partition ξ and a bilateral strictly increasing sequence $(k_t, -\infty < t < \infty)$ of integers such that the partitions $T^{k_t}\xi \,(-\infty < t < \infty)$ are independent. However, the inductive definition gets technically more difficult in that case.

References

1. Y. N. Dowker, *On measurable transformations in finite measure spaces*, Ann. of Math. (2) **62** (1955), 504–516. MR **17**, 353.

2. A. B. Hajian and S. Kakutani, *Weakly wandering sets and invariant measures*, Trans. Amer. Math. Soc. **110** (1964), 136–151. MR **27** #4904.

3. P. R. Halmos, *Shifts on Hilbert spaces*, J. Reine Angew. Math. **208** (1961), 102–112. MR **27** #2868.

4. A. Hanen and J. Neveu, *Atomes conditionnels d'un espace de probabilité*, Acta Math. Acad. Sci. Hungar. **17** (1966), 443–449. MR **34** #5116.

5. G. Herglotz, *Über Potenzreihen mit positivem, reellem Teil im Einheitskreis*, S.-B. Sächs. Akad. Wiss. **63** (1911), 501–511.

6. E. Hewitt and K. Stromberg, *Real and abstract analysis. A modern treatment of the theory of functions of a real variable*, Springer-Verlag, New York, 1965. MR **32** #5826.

7. E. Hopf, *Ergodentheorie*, Ergebnisse der Mathematik, Band 5, no. 2, Springer-Verlag, Berlin, 1937; reprint, Chelsea, New York, 1948. MR **10**, 549.

8. K. Jacobs, *Neuere Methoden und Ergebnisse der Ergodentheorie*, Ergebnisse der Mathematik und ihrer Grenzgebiete, N.F., Heft 29, Springer-Verlag, Berlin, 1960. MR **23** #A1000.

9. S. Kakutani and W. Parry, *Infinite measure preserving transformations with "mixing"*, Bull. Amer. Math. Soc. **69** (1963), 752–756. MR **27** #3776.

10. U. Krengel and L. Sucheston, *On mixing in infinite measure spaces*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **13** (1969), 150–164. MR **40** #7424.

11. K. Krickeberg, *Recent results on mixing in topological measure spaces*, Lecture Notes in Math., no. 89, Springer-Verlag, Berlin, 1969, pp. 178–185.

12. A. G. Kušnirenko, *Metric invariants of entropy type*, Uspehi Mat. Nauk **22** (1967), no. 5 (137), 57–65 = Russian Math. Surveys **22** (1967), no. 5, 53–61. MR **36** #348.

13. V. P. Leonov, *The use of the characteristic functional and semi-invariants in the ergodic theory of stationary processes*, Dokl. Akad. Nauk SSSR **133** (1960), 523–526 = Soviet Math. Dokl. **1** (1961), 878–881. MR **23** #A2937.

14. J. Neveu, *Théorie ergodique*, Lecture Notes, University of Paris, 1966.

15. ———, *Atomes conditionnels d'espaces de probabilité et théorie de l'information*, Sympos. on Probability Methods in Analysis (Loutraki, 1966), Springer, Berlin, 1967, pp. 256–271. MR **36** #3602.

16. A. Rényi, *On mixing sequences of sets*, Acta Math. Acad. Sci. Hungar. **9** (1958), 215–228. MR **20** #4623.

17. V. A. Rohlin, *On the fundamental ideas of measure theory*, Mat. Sb. **25** (67) (1949), 107–150; English transl., Amer. Math. Soc. Transl. (1) **10** (1962), 1–54. MR **11**, 18.

18. L. Sucheston, *On mixing and the zero-one law*, J. Math. Anal. Appl. **6** (1963), 447–456. MR **26** #7010.

19. ———, *On existence of finite invariant measures*, Math. Z. **86** (1964), 327–336. MR **32** #7706.

20. L. Zsido, *Conditional expectations and Sinai's theorem*, Stud. Cerc. Mat. **20** (1968), 1045–1111. (Rumanian).

21. J. W. England and N. F. G. Martin, *On weak mixing metric automorphisms*, Bull. Amer. Math. Soc. **74** (1968), 505–507. MR **36** #5303.

22. L. K. Jones, *A short proof of Sucheston's characterization of mixing* (to appear).

Department of Mathematics, Ohio State University, Columbus, Ohio 43210

Current address: Institut für Mathematische Statistik, 34 Göttingen, Bürgerstrasse 32, Federal Republic of Germany