COLLOQUIUM MATHEMATICUM

VOL. 108 2007 NO. 2

WEAKLY MIXING TRANSFORMATIONS AND THE CARATHÉODORY DEFINITION OF MEASURABLE SETS

BY

AMOS KOELLER (Berlin), RODNEY NILLSEN (Wollongong) and GRAHAM WILLIAMS (Wollongong)

Abstract. Let $\mathbb{T}$ denote the set of complex numbers of modulus 1. Let $v \in \mathbb{T}$, v not a root of unity, and let $T : \mathbb{T} \to \mathbb{T}$ be the transformation on $\mathbb{T}$ given by $T(z) = vz$. It is known that the problem of calculating the outer measure of a T-invariant set leads to a condition which formally has a close resemblance to Carathéodory's definition of a measurable set. In ergodic theory terms, T is not weakly mixing. Now there is an example, due to Kakutani, of a transformation $\widetilde{\psi}$ which is weakly mixing but not strongly mixing. The results here show that the problem of calculating the outer measure of a $\widetilde{\psi}$-invariant set leads to a condition formally resembling the Carathéodory definition, as in the case of the transformation T. The methods used bring out some of the more detailed behaviour of the Kakutani transformation. The above mentioned results for T and the Kakutani transformation do not apply for the strongly mixing transformation $z \mapsto z^2$ on $\mathbb{T}$.

1. Introduction and notations. In 1914 Carathéodory [3] gave a definition of measurable sets. This definition is very important in the theory of measure and integration and is the key step in going from an outer measure, defined on all subsets of some given set, to a measure defined on the measurable sets ([1, pp. 101–103] or [4, pp. 44–46]). Carathéodory's definition has received specific comments from various mathematicians (see the discussion in [8, pp. 106–108]). It was shown in [8] that the problem of calculating the Lebesgue outer measure of a subset of $\mathbb{T}$ which is invariant under an irrational rotation leads to Carathéodory's definition. By way of contrast, however, it was mentioned in [8] (and a proof is in [9]) that if S denotes the transformation on $\mathbb{T}$ given by $z \mapsto z^2$, then the corresponding problem of calculating the outer measure of an S-invariant set does not lead to the Carathéodory definition. This contrasting situation seems to be related to the facts that an irrational rotation on the unit circle is not weakly mixing [10, p. 40], whereas the transformation $z \mapsto z^2$ on the unit circle is strongly mixing [10, p. 50].

2000 *Mathematics Subject Classification*: Primary 37A05; Secondary 28A05.

Key words and phrases: ergodic transformation, mixing, outer measure, Kakutani transformation, measurable set, Carathéodory definition.

In this paper it is shown that if the example of a weakly mixing but not strongly mixing transformation, given in 1973 by S. Kakutani in [6], is subject to an analysis similar to that in [8] then Kakutani's transformation behaves in a like manner to an irrational rotation, despite its stronger mixing properties.

Let X be a set and let $T : X \to X$ be a transformation on X. Let T^0 denote the identity transformation, and if $n \in \mathbb{N}$, the set of natural numbers, we write T^n for the composition of T with itself n times. If A is a subset of X and $n = 0$ or $n \in \mathbb{N}$, we set

$$T^n(A) = \{T^n(x) : x \in A\} \quad \text{and} \quad T^{-n}(A) = \{y : y \in X \text{ and } T^n(y) \in A\}.$$

A subset A of X is T-*invariant*, or simply *invariant*, if $T^{-1}(A) = A$.

Furthermore, let $\mathcal{B}$ be a σ-algebra of subsets of X, let $\mu : \mathcal{B} \to [0, \infty]$ be a measure, and assume that $T : X \to X$ is such that $T^{-1}(A) \in \mathcal{B}$ and $\mu(T^{-1}(A)) = \mu(A)$ for all $A \in \mathcal{B}$. In this case, T is called a μ-*measure preserving transformation* on X, or simply a *measure preserving transformation*. A measure preserving transformation T is called *ergodic* if, whenever $A \in \mathcal{B}$ and A is T-invariant, then $\mu(A) = 0$ or $\mu(A^{\mathrm{c}}) = 0$. Following [2, p. 14] and [10, pp. 39–40] we shall say that a measure preserving transformation $T : X \to X$ is *weakly mixing* if for all $A, B \in \mathcal{B}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B)| = 0. \tag{1.1}$$

Also, T is said to be *strongly mixing* if for all $A, B \in \mathcal{B}$,

$$\lim_{n\to\infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B).$$

A weakly mixing transformation is ergodic. Suitable general references on ergodic theory are [2] and [10].

An example of a weakly mixing transformation which is not strongly mixing is due to Kakutani [6]. To describe it, let $\mathcal{B}$ be a σ-algebra of subsets of X, let μ be a measure on $\mathcal{B}$ and let T be a measure preserving transformation on X. Let $A \in \mathcal{B}$ be a subset of X such that $\mu(A) > 0$ and $\mu(A^{\mathrm{c}}) > 0$. Let A' be a set disjoint from X and such that there is a one-to-one mapping τ from A onto A'. Define

$$\begin{aligned}
\widetilde{X} &= X \cup A', \\
\widetilde{\mathcal{B}} &= \{B : B \subseteq \widetilde{X},\ B \cap X \in \mathcal{B} \text{ and } \tau^{-1}(B \cap A') \in \mathcal{B}\}, \\
\widetilde{\mu}(B) &= \mu(B \cap X) + \mu(\tau^{-1}(B \cap A')) \quad \text{ for all } B \in \widetilde{\mathcal{B}}.
\end{aligned}$$

Then $\widetilde{\mathcal{B}}$ is a σ-algebra of subsets of $\widetilde{X}$ and $\widetilde{\mu}$ is a measure on $\widetilde{\mathcal{B}}$. If we define the transformation $\widetilde{T}$ on $\widetilde{X}$ by

$$\widetilde{T}(x) = \begin{cases} \tau(x) & \text{if } x \in A, \\ T(x) & \text{if } x \in A^{\mathrm{c}} \cap X, \\ T(\tau^{-1}(x)) & \text{if } x \in A', \end{cases}$$

then $\widetilde{T}$ is a measure preserving transformation on $\widetilde{X}$. This construction is a special case of an induced measure preserving transformation [5]: the transformation T on X induces the transformation $\widetilde{T}$ on $\widetilde{X}$. If T is one-to-one then $\widetilde{T}$ is one-to-one; the range of $\widetilde{T}$ is the union of the range of T and A'; and if T is ergodic then $\widetilde{T}$ is ergodic.

Kakutani's example is a special case of the construction above. Henceforth, take X to be the unit interval $[0,1)$, $\mathcal{B}$ the family of Borel subsets of X, and μ the Borel measure on $\mathcal{B}$. For $n = 0, 1, 2, \ldots$, put

$$I_n = \left[\, 1 - \frac{1}{2^n}, 1 - \frac{1}{2^{n+1}} \right) \quad \text{and} \quad A = \bigcup_{n=0}^{\infty} I_{2n}.$$

Clearly, $X = [0,1) = \bigcup_{n=0}^{\infty} I_n$. Take A' to be any set disjoint from X such that there is a one-to-one function τ mapping A onto A'. Finally, in place of T we will have Kakutani's transformation ψ on X which is given by

$$\psi(x) = x - 1 + \frac{1}{2^n} + \frac{1}{2^{n+1}} \quad \text{for } x \in I_n. \tag{1.2}$$

Kakutani's transformation is illustrated in Figure 1.

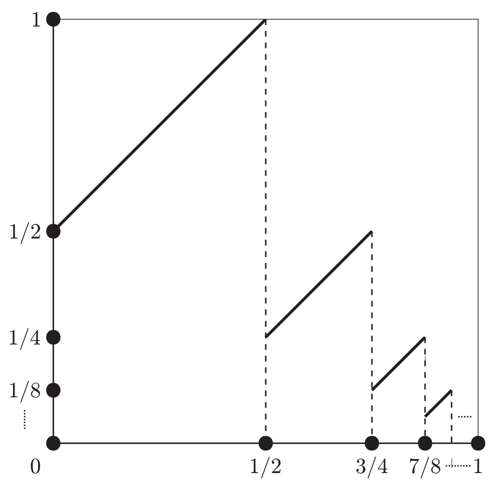

Fig. 1. Graph of the Kakutani transformation

The transformation ψ on X is measure preserving. Then, as above, we construct $(\widetilde{X}, \widetilde{\mathcal{B}}, \widetilde{\mu}, \widetilde{\psi})$ from $(X, \mathcal{B}, \mu, \psi)$. Note that ψ maps I_n linearly onto $[1/2^{n+1}, 1/2^n)$, and $(0,1) = \bigcup_{n=0}^{\infty} [1/2^{n+1}, 1/2^n)$, where the union is disjoint.

It follows that ψ maps $[0,1)$ onto $(0,1)$ one-to-one. It is mentioned in [6] that ψ is ergodic on X. However, it is not weakly mixing, for $\psi(0,1/2)=(1/2,1)$ and $\psi(1/2,1)=(0,1/2)$. Consequently, if we take $A=B=(0,1/2)$ in (1.1), the left hand side of (1.1) equals $1/4$ while the right hand side equals 0. Also, $\widetilde{\psi}$ is one-to-one, ergodic, and its range is $(0,1)\cup A'$. The function $\widetilde{\psi}$ was defined by Kakutani [6], who showed that $\widetilde{\psi}$ is measure preserving and weakly mixing on $\widetilde{X}$, but not strongly mixing. This example of Kakutani's is also discussed in [2, pp. 29–30].

2. Dyadic intervals and orbits. One of the techniques we use in proving our main result is to repeatedly apply $\widetilde{\psi}$ to particular subsets of $\widetilde{X}$. We need some definitions.

DEFINITIONS. Let Y be a set and let $T:Y\to Y$ be a transformation. For any subset D of Y the *orbit of D with respect to T* is the sequence $O_T(D)=D,T(D),T^2(D),T^3(D),\ldots$ of subsets of Y. If D has a single element, $D=\{x\}$ say, $O_T(D)$ is denoted by $O_T(x)$ and called the *orbit of x with respect to T*. A point in $[0,1)$ of the form $k/2^n$ for some $n\in\mathbb{N}$ and $k\in\{0,1,\ldots,2^n-1\}$ is called a *dyadic rational*. An open subinterval J of $[0,1)$ is called a *dyadic interval* or a *dyadic subinterval* if there are $n\in\mathbb{N}$ and $q\in\{0,1,\ldots,2^n-1\}$ such that $J=(q/2^n,(q+1)/2^n)$; the n is uniquely determined and called the *order* of J.

Note that a dyadic interval has order n if and only if its length is 2^{-n}, and that there are 2^n dyadic intervals of order n. Given such a dyadic interval J, and a subset B of J, define

$$B_l=B\cap\left(\frac{q}{2^n},\frac{q+1/2}{2^n}\right)\quad\text{and}\quad B_r=B\cap\left(\frac{q+1/2}{2^n},\frac{q+1}{2^n}\right).$$

If $B=J$, then J_l is called the *left half* of J, and J_r is the *right half*. Each of J_l,J_r is a dyadic interval of order $n+1$. If $A_1,\ldots,A_n$ and $B_1,\ldots,B_n$ are finite sequences of subsets of $[0,1)$, we say $A_1,\ldots,A_n$ *essentially equals* $B_1,\ldots,B_n$ if, for each $i=1,\ldots,n$, the symmetric difference of A_i and B_i consists of a (possibly empty) set of dyadic rationals. In this case we write $A_1\equiv B_1,\ \ldots,\ A_n\equiv B_n$. It is easy to check that if $x\in[0,1)$, then x is a dyadic rational if and only if $\psi(x)$ is a dyadic rational. This means that $A\equiv B$ if and only if $\psi(A)\equiv\psi(B)$.

LEMMA 2.1. *Let $J=(q/2^n,(q+1)/2^n)$ be a dyadic subinterval of $[0,1)$ with $n\in\mathbb{N}$ and $q\in\{0,1,\ldots,2^n-1\}$. Then the following hold*:

(i) *If $q\in\{0,1,\ldots,2^n-2\}$, then there is $k\in\{0,1,\ldots,2^n-1\}$, with $k\neq q$, such that*
$$\psi(J)=\left(\frac{k}{2^n},\frac{k+1}{2^n}\right).$$

If $q = 2^n - 1$, *then*

$$\psi(J) = \left(0, \frac{1}{2^{n+1}}\right) \cup \left(\frac{1}{2^{n+1}}, \frac{1}{2^n}\right),$$

and so $\psi(J) \equiv (0, 1/2^n)$.

(ii) *If* $q \in \{0, 1, \ldots, 2^n - 2\}$, *then*

$$\psi(J_l) = \psi(J)_l \quad \textit{and} \quad \psi(J_r) = \psi(J)_r.$$

If $q = 2^n - 1$, *then*

$$\psi(J_l) = \psi(J)_r \quad \textit{and} \quad \psi(J_r) \equiv \psi(J)_l.$$

(iii) *The* ψ*-orbit of* J *essentially equals some sequential arrangement of the set of all* 2^n *dyadic intervals of order* n.

(iv) *If* $J = (0, 1/2^n)$, *the first* 2^n *elements of the* ψ*-orbit of* J, *not necessarily in this order*, *are*

$$\left(0, \frac{1}{2^n}\right), \left(\frac{1}{2^n}, \frac{2}{2^n}\right), \ldots, \left(\frac{2^n - 1}{2^n}, 1\right),$$

and if $p \in \{0, 1, \ldots, 2^n - 1\}$,

$$\psi^p(J) = \left(\frac{2^n - 1}{2^n}, 1\right) \iff p = 2^n - 1.$$

Proof. (i) As $q/2^n \in [0, 1)$, there is $n_0 \in \{0\} \cup \mathbb{N}$ such that

$$\frac{q}{2^n} \in I_{n_0} = \left[1 - \frac{1}{2^{n_0}}, 1 - \frac{1}{2^{n_0+1}}\right).$$

If $n < n_0$, then $1/2^n > 1/2^{n_0}$ and so

$$1 \geq \frac{q+1}{2^n} \geq 1 - \frac{1}{2^{n_0}} + \frac{1}{2^n} > 1,$$

a contradiction. Hence, $n \geq n_0$, a fact which is used in the ensuing argument. We now consider the two cases: $q = 2^n - 1$ and $q \in \{0, 1, \ldots, 2^n - 2\}$.

Case I. Assume that $q = 2^n - 1$. Then

$$J = \left(\frac{q}{2^n}, \frac{q+1}{2^n}\right) = \left(1 - \frac{1}{2^n}, 1\right) = \left(1 - \frac{1}{2^n}, 1 - \frac{1}{2^{n+1}}\right) \cup \bigcup_{k=n+1}^{\infty} I_k.$$

We now have, using (1.2),

$$\begin{aligned} \psi(J) &= \psi\left(1 - \frac{1}{2^n}, 1 - \frac{1}{2^{n+1}}\right) \cup \bigcup_{k=n+1}^{\infty} \psi(I_k) \\ &= \left(\frac{1}{2^{n+1}}, \frac{1}{2^n}\right) \cup \bigcup_{k=n+1}^{\infty} \left[\frac{1}{2^{k+1}}, \frac{1}{2^k}\right) = \left(0, \frac{1}{2^{n+1}}\right) \cup \left(\frac{1}{2^{n+1}}, \frac{1}{2^n}\right). \end{aligned}$$

This shows that the conclusion of (i) holds in this case.

Case II. Now let $q \in \{0, 1, \ldots, 2^n - 2\}$. Recall that $n_0 \in \{0\} \cup \mathbb{N}$ was chosen so that $q/2^n \in I_{n_0} = [1 - 1/2^{n_0}, 1 - 1/2^{n_0+1})$, we saw that $n \geq n_0$, and although it was not explicitly stated, if $n = n_0$ then $q = 2^n - 1$. In Case II, $n > n_0$. Hence,

$$1 - \frac{1}{2^{n_0}} \leq \frac{q}{2^n} < 1 - \frac{1}{2^{n_0+1}} = \frac{2^n - 2^{n-n_0-1}}{2^n},$$

so that $q < 2^n - 2^{n-n_0-1}$ and

$$1 - \frac{1}{2^{n_0}} \leq \frac{q}{2^n} < \frac{q+1}{2^n} \leq \frac{2^n - 2^{n-n_0-1}}{2^n} = 1 - \frac{1}{2^{n_0+1}}.$$

Thus

$$J = \left(\frac{q}{2^n}, \frac{q+1}{2^n}\right) \subseteq \left[1 - \frac{1}{2^{n_0}}, 1 - \frac{1}{2^{n_0+1}}\right) = I_{n_0}.$$

It now follows from the definition of ψ in (1.2) that

$$\psi\left(\frac{q}{2^n}\right) = \frac{q}{2^n} - 1 + \frac{1}{2^{n_0}} + \frac{1}{2^{n_0+1}} = \frac{q - 2^n + 2^{n-n_0} + 2^{n-n_0-1}}{2^n},$$

and if in the next line we take the limit from the left, we obtain

$$\lim_{x \to (q+1)/2^n} \psi(x) = \frac{q+1}{2^n} - 1 + \frac{1}{2^{n_0}} + \frac{1}{2^{n_0+1}} = \frac{q - 2^n + 2^{n-n_0} + 2^{n-n_0-1} + 1}{2^n}.$$

Now, ψ is linear and increasing on I_{n_0}. Thus, if we put $k = q - 2^n + 2^{n-n_0} + 2^{n-n_0-1}$, we will have $k \in \{0, 1, \ldots, 2^n - 1\}$, $k \neq q$ and also

$$\psi(J) = \psi\left(\frac{q}{2^n}, \frac{q+1}{2^n}\right) = \left(\frac{k}{2^n}, \frac{k+1}{2^n}\right).$$

Thus (i) also holds in this case, which completes the proof of (i).

(ii) The proof is similar to and uses (i), with the same two cases being considered.

(iii) We use induction. For $n = 1$, let $J = (0, 1/2)$. Then $\psi(J) = (1/2, 1)$, and $\psi^2(J) = \psi(1/2, 1) = (0, 1/4) \cup (1/4, 1/2)$. Thus the orbit $O_\psi(J)$ of J essentially equals the two dyadic intervals of order 1, namely $(0, 1/2)$ and $(1/2, 1)$. Hence the claim is true for $n = 1$.

Now assume it is true for $n = k$. That is, if $q \in \{0, 1, \ldots, 2^k - 1\}$ and $J = (q/2^k, (q+1)/2^k)$, then the orbit $O_\psi(J)$ consists essentially of

$$\left(0, \frac{1}{2^k},\right), \left(\frac{1}{2^k}, \frac{2}{2^k}\right), \ldots, \left(\frac{2^k - 1}{2^k}, 1\right).$$

Let $J = ((2^k - 1)/2^k, 1)$, and let

$$L = J_r = \left(\frac{2^k - 1}{2^k}, 1\right)_r = \left(\frac{2^{k+1} - 1}{2^{k+1}}, 1\right).$$

By the inductive asumption, the orbit $J, \psi(J), \dots, \psi^{2^k-1}(J)$ essentially equals the set of all dyadic intervals of order k, and $\psi^{2^k}(J) \equiv J$. It follows by part (ii) that

$$\begin{aligned}
\psi(L) &= \psi(J_r) \equiv \psi(J)_l, \\
\psi^2(L) &= \psi(\psi(L)) \equiv \psi(\psi(J)_l) \equiv \psi(\psi(J))_l \equiv \psi^2(J)_l, \\
\psi^j(L) &= \psi^{j-1}(\psi(L)) \equiv \psi^{j-1}(\psi(J)_l) \equiv \dots \equiv (\psi^j(J))_l,
\end{aligned}$$

for all $j = 1, \dots, 2^k - 1$. Thus, the first 2^k intervals in the ψ-orbit of L are essentially equal to L (which equals J_r), $\psi(J)_l, \psi^2(J)_l \dots, \psi^{2^k-1}(J)_l$.

Now, by (ii) again,

$$\psi^{2^k}(L) = \psi(\psi^{2^k-1}(L)) \equiv \psi((\psi^{2^k-1}(J))_l) \equiv (\psi(\psi^{2^k-1}(J)))_l \equiv (\psi^{2^k}(J))_l \equiv J_l.$$

But by (ii), $\psi(J_l) = \psi(J)_r$, and the previous argument may now be repeated with $\psi(J)_r$ in place of $\psi(J)_l$. We deduce that the next 2^k elements of the orbit of L are essentially equal to $J_l, \psi(J)_r, \dots, \psi^{2^k-1}(J)_r$. In summary, the intervals in the ψ-orbit of L, in order, are

$$J_r, \psi(J)_l, \psi^2(J)_l, \dots, \psi^{2^k-1}(J)_l, J_l, \psi(J)_r, \psi^2(J)_r, \dots, \psi^{2^k-1}(J)_r.$$

However, by the inductive assumption, these intervals comprise essentially all the left and right parts of all dyadic intervals of order k. Since the totality of those left and right parts gives all the dyadic intervals of order $k+1$, we see that the ψ-orbit of L comprises essentially all the dyadic intervals of order $k+1$. But this implies that the ψ-orbit of *any* dyadic interval of order $k+1$ will comprise essentially all the dyadic intervals of order $k+1$. So, if the result is true for dyadic intervals of order k, it is true for dyadic intervals of order $k+1$. This proves (iii).

(iv) Use the inductive argument as in (iii). ∎

The remaining results in this section reveal what happens when $\widetilde{\psi}$ is repeatedly applied to a dyadic interval.

LEMMA 2.2. *Let $J = (q/2^n, (q+1)/2^n)$ be a dyadic subinterval of $[0,1)$ with $n \in \mathbb{N}$ and $q \in \{0, 1, \dots, 2^n - 1\}$. Then the following hold:*

(i) *If $q = 2^n - 1$ then $\widetilde{\psi}(J)$ has non-void intersection with both $\psi(J)$ and A'.*
(ii) *If $q \in \{0, 1, \dots, 2^n - 2\}$ then either $J \cap A = \emptyset$ in which case $\widetilde{\psi}(J) = \psi(J)$, or $J \subseteq A$ in which case $\widetilde{\psi}(J) = \tau(J) \subseteq A'$.*

Proof. (i) Let $q = 2^n - 1$. Then

$$J = \left(1 - \frac{1}{2^n}, 1\right) = \left(1 - \frac{1}{2^n}, 1 - \frac{1}{2^{n+1}}\right) \cup \bigcup_{k=n+1}^{\infty} I_k,$$

so that J contains both $(1-1/2^n, 1-1/2^{n+1})$ and I_{n+1}. If n is even, then

$$\left(1-\frac{1}{2^n}, 1-\frac{1}{2^{n+1}}\right) \subseteq I_n \subseteq A \quad \text{and} \quad I_{n+1} \cap A = \emptyset.$$

The definition of $\widetilde{\psi}$ now gives

$$\widetilde{\psi}\left(1-\frac{1}{2^n}, 1-\frac{1}{2^{n+1}}\right) = \tau\left(1-\frac{1}{2^n}, 1-\frac{1}{2^{n+1}}\right) \subseteq A'$$

and

$$\widetilde{\psi}(I_{n+1}) = \psi(I_{n+1}) = \left[\frac{1}{2^{n+2}}, \frac{1}{2^{n+1}}\right) \subseteq \left(0, \frac{1}{2^n}\right) = \psi(J).$$

Thus,

$$\widetilde{\psi}(J) \cap A' \supseteq \widetilde{\psi}(I_n) \cap A' = \tau(I_n) \neq \emptyset$$

and

$$\widetilde{\psi}(J) \cap \psi(J) \supseteq \widetilde{\psi}(I_{n+1}) \cap \psi(J) = \left(\frac{1}{2^{n+2}}, \frac{1}{2^{n+1}}\right) \neq \emptyset.$$

If instead n is odd and $n+1$ is even, a similar argument works.

(ii) Now let $q \in \{0, 1, \ldots, 2^n - 2\}$. There is a unique $m \in \{0, 1, 2, \ldots\}$ such that $q/2^n \in I_m$. As in the proof of Lemma 2.1(i), it is easy to check that $n \geq m$. Now, recalling that

$$\begin{aligned} I_m &= \left[\frac{2^m - 1}{2^m}, \frac{2^{m+1}-1}{2^{m+1}}\right) = \left[\frac{2^n - 2^{n-m}}{2^n}, \frac{2^n - 2^{n-m-1}}{2^n}\right) \\ &= \bigcup_{j=2^n-2^{n-m}}^{2^n-2^{n-m-1}-1} \left[\frac{j}{2^n}, \frac{j+1}{2^n}\right), \end{aligned}$$

it follows that $q \in \{2^n - 2^{n-m}, \ldots, 2^n - 2^{n-m-1} - 1\}$. However, as $J = [q/2^n, (q+1)/2^n)$, this makes it clear that $J \subseteq I_m$. So by the definition of $\widetilde{\psi}$ we now have

$$\tilde{\psi}(J) = \begin{cases} \tau(J) & \text{if } m \text{ is even}, \\ \psi(J) & \text{if } m \text{ is odd}. \end{cases} \blacksquare$$

LEMMA 2.3. *Let $n \in \mathbb{N}$ and $J = (0, 1/2^n)$. Then for all $j \in \{0, 1, 2, \ldots, 2^n - 1\}$, $\psi^j(J)$ is essentially equal to a set in $O_{\widetilde{\psi}}(J)$.*

Proof. We keep n fixed, and use induction on j. For $j = 0$, we have $\psi^0(J) = J = \widetilde{\psi}^0(J) \in O_{\widetilde{\psi}}(J)$, and the claim is true.

Now, assume it is true for some $j \in \{0, 1, \ldots, 2^n - 2\}$. Then there is $q \in \{0, 1 \ldots\}$ such that $\psi^j(J) = \widetilde{\psi}^q(J)$, and we know that $\psi^j(J)$ is a dyadic interval of order n. Further because $j < 2^n - 1$, Lemma 2.1(iv) shows that $\psi^j(J)$ is not $(1 - 1/2^n, 1)$. Thus, Lemma 2.2(ii) applies to $\psi^j(J)$, and we have

$$\widetilde{\psi}(\psi^j(J)) = \psi(\psi^j(J)) \quad \text{or} \quad \widetilde{\psi}(\psi^j(J)) = \tau(\psi^j(J)).$$

In the former case we have

$$\widetilde{\psi}^{q+1}(J) = \widetilde{\psi}(\widetilde{\psi}^{q}(J)) = \widetilde{\psi}(\psi^{j}(J)) = \psi^{j+1}(J),$$

so that $\psi^{j+1}(J) \in O_{\widetilde{\psi}}(J)$. In the latter case we have

$$\begin{aligned}\widetilde{\psi}^{q+2}(J) &= \widetilde{\psi}^{2}(\widetilde{\psi}^{q}(J)) = \widetilde{\psi}^{2}(\psi^{j}(J)) = \widetilde{\psi}(\widetilde{\psi}(\psi^{j}(J))) = \widetilde{\psi}(\tau(\psi^{j}(J)))\\ &= \psi(\psi^{j}(J)) = \psi^{j+1}(J),\end{aligned}$$

so that again $\psi^{j+1}(J) \in O_{\widetilde{\psi}}(J)$, and the proof is complete. ∎

3. Weak mixing and the Carathéodory definition of measurable sets. Recall that X is the interval $[0,1)$, $\mathcal{B}$ is the σ-algebra of Borel subsets of X, μ is Lebesgue measure on $\mathcal{B}$ and ψ is Kakutani's function. Also, $\widetilde{X}$, $\widetilde{\mathcal{B}}$, $\widetilde{\mu}$, $\widetilde{\psi}$ are constructed from $X, \mathcal{B}, \mu, \psi$ as decribed in Section 1. Note that $\widetilde{\psi}$ is one-to-one and its range is $\widetilde{X} \setminus \{0\}$. Now, let μ_* denote the Lebesgue outer measure on X. The measure $\widetilde{\mu}$ on $\widetilde{\mathcal{B}}$ can be extended, in the usual way, to an outer measure $\widetilde{\mu}_*$ on the whole of $\widetilde{X}$. Then it is readily checked that for all $B \subseteq \widetilde{X}$,

$$\widetilde{\mu}_*(B) = \mu_*(B \cap X) + \mu_*(\tau^{-1}(B \cap A')).$$

Note that if $B \subseteq X$, then $\mu_*(B) = \widetilde{\mu}_*(B)$.

LEMMA 3.1. *Let $B \subseteq \widetilde{X}$, let B be $\widetilde{\psi}$-invariant, let $\widetilde{\mu}_*(B) > 0$ and let $\varepsilon > 0$. Then $\mu_*(B \cap X) > 0$ and there is a dyadic interval J such that*

$$\frac{\mu_*(B \cap J)}{\mu(J)} > 1 - \varepsilon. \tag{3.1}$$

In fact, there is a number $n_0 \in \mathbb{N}$ such that for any $n \in \mathbb{N}$ with $n > n_0$, there is a dyadic interval J of order n such that (3.1) *holds.*

Proof. To show that $\mu_*(B \cap X) > 0$, observe that since

$$\widetilde{\mu}_*(B) = \mu_*(B \cap X) + \mu_*(\tau^{-1}(B \cap A'))$$

and $\widetilde{\mu}_*(B) > 0$, we have either $\mu_*(B \cap X) > 0$ or $\mu_*(\tau^{-1}(B \cap A')) > 0$. In the former case, we are done, while in the latter, after observing that $\widetilde{\psi}^{-1}(B \cap A') = \tau^{-1}(B \cap A')$ and $\widetilde{\psi}^{-1}(A') = A$, we have

$$\begin{aligned}0 &< \mu_*(\tau^{-1}(B \cap A')) = \widetilde{\mu}_*(\widetilde{\psi}^{-1}(B \cap A')) = \widetilde{\mu}_*(\widetilde{\psi}^{-1}(B) \cap \widetilde{\psi}^{-1}(A'))\\ &= \widetilde{\mu}_*(B \cap A)) \le \mu_*(B \cap X).\end{aligned}$$

So, in either case, $\mu_*(B \cap X) > 0$.

As we now know that $\mu_*(B \cap X) > 0$, we can apply an argument in Lemma 4.2 of [8] to deduce that for all sufficiently small $\delta > 0$, there is a subinterval J of X such that

$$\mu(J) = \delta \quad \text{and} \quad \frac{\mu_*(B \cap J)}{\mu(J)} > 1 - \varepsilon.$$

In fact, the argument in Lemma 4.2 of [8], modified slightly so as to apply specifically to dyadic intervals, shows that if $n_0 \in \mathbb{N}$ is such that $1/n_0 < \delta$, then for all $n \in \mathbb{N}$ with $n \geq n_0$, there is a dyadic interval J of order n such that $\mu_*(B \cap J)/\mu(J) > 1 - \varepsilon$, as required. ∎

LEMMA 3.2. *If* $B \subseteq \widetilde{X}$, *then* $\widetilde{\mu}_*(\widetilde{\psi}^{-1}(B)) = \widetilde{\mu}_*(B)$.

Proof. If J is a subinterval of X, it is clear from the definition of ψ as a one-to-one piecewise linear function, where each linear piece has slope 1, that $\mu(\psi(B)) = \mu(B)$ for all Borel subsets B of X. A routine calculation, using the fact that the range of ψ is $X \setminus \{0\}$, shows that

$$\mu_*(\psi^{-1}(B)) = \mu_*(B) \quad \text{for all subsets } B \text{ of } X.$$

Now, if $B \subseteq \widetilde{X}$, then

$$\widetilde{\psi}^{-1}(B) = [A^{\mathrm{c}} \cap X \cap \psi^{-1}(B \cap X)] \cup [A \cap \tau^{-1}(B \cap A')] \cup [A' \cap \tau(\psi^{-1}(B \cap X))].$$

Also, if F, G are subsets of X such that $F \subseteq C$ and $G \subseteq C^{\mathrm{c}}$ for some Borel set $C \subseteq X$, then $\mu_*(F \cup G) = \mu_*(F) + \mu_*(G)$. It follows that

$$\begin{aligned}
\widetilde{\mu}_*(\widetilde{\psi}^{-1}(B)) &= \mu_*([A^{\mathrm{c}} \cap X \cap \psi^{-1}(B \cap X)] \cup [A \cap \tau^{-1}(B \cap A')]) \\
&\quad + \mu_*(\tau^{-1}[A' \cap \tau(\psi^{-1}(B \cap X))]) \\
&= \mu_*(A^{\mathrm{c}} \cap X \cap \psi^{-1}(B \cap X)) + \mu_*(A \cap \tau^{-1}(B \cap A')) \\
&\quad + \mu_*(A \cap \psi^{-1}(B \cap X)) \\
&= \mu_*(\psi^{-1}(B \cap X)) + \mu_*(\tau^{-1}(B \cap A')) \\
&= \mu_*(B \cap X) + \mu_*(\tau^{-1}(B \cap A')) = \widetilde{\mu}_*(B). \ ∎
\end{aligned}$$

THEOREM 3.3. *Let* $B \subseteq \widetilde{X}$ *be* $\widetilde{\psi}$*-invariant. Suppose that there is* $\theta \in [0, 2)$ *such that for all dyadic subintervals* J *of* $[0, 1)$,

$$\widetilde{\mu}_*(B \cap J) + \widetilde{\mu}_*(B^{\mathrm{c}} \cap J) \leq \theta \mu(J). \tag{3.2}$$

Then $\theta \in [1, 2)$ *and either* $\widetilde{\mu}_*(B) = 0$ *or* $\widetilde{\mu}_*(B^{\mathrm{c}}) = 0$.

Proof. The fact that $\theta \in [1, 2)$ is immediate from the observation that

$$\mu(J) = \widetilde{\mu}(J) \leq \widetilde{\mu}_*(B \cap J) + \widetilde{\mu}_*(B^{\mathrm{c}} \cap J) \leq \theta \mu(J).$$

Now, suppose that $\widetilde{\mu}_*(B) > 0$ and $\widetilde{\mu}_*(B^{\mathrm{c}}) > 0$. Set $\varepsilon = (2 - \theta)/3 > 0$. Then, as $\widetilde{\mu}_*(B) > 0$ and $\widetilde{\mu}_*(B^{\mathrm{c}}) > 0$, and as B^{c} is also $\widetilde{\psi}$-invariant, by Lemma 3.1 there exist dyadic intervals J_1, J_2 of the same order such that

$$\widetilde{\mu}_*(B \cap J_1) > (1 - \varepsilon)\mu(J_1) \quad \text{and} \quad \widetilde{\mu}_*(B^{\mathrm{c}} \cap J_2) > (1 - \varepsilon)\mu(J_2).$$

Since J_1, J_2 are dyadic intervals of the same order n, say, Lemma 2.3 shows that both J_1 and J_2 are essentially equal to sets belonging to $O_{\widetilde{\psi}}(0, 1/2^n)$. Thus, there are $r, s \in \{0, 1, 2, \ldots\}$ such that

$$J_1 \equiv \widetilde{\psi}^r((0, 1/2^n)) \quad \text{and} \quad J_2 \equiv \widetilde{\psi}^s((0, 1/2^n)).$$

Without loss of generality we may suppose that $r \le s$, in which case

$$\widetilde{\psi}^{s-r}(J_1) = \widetilde{\psi}^{s-r}\left(\widetilde{\psi}^{r}\left(0, \frac{1}{2^n}\right)\right) = \widetilde{\psi}^{s}\left(\left(0, \frac{1}{2^n}\right)\right) = J_2.$$

Now, $\mu(J_1) = \mu(J_2)$, $\widetilde{\mu}_*$ is $\widetilde{\psi}$-invariant as an outer measure, as expressed by Lemma 3.2, and B^{c} is a $\widetilde{\psi}$-invariant set. Hence,

$$\begin{aligned}(1-\varepsilon)\mu(J_1) = {} & (1-\varepsilon)\mu(J_2)\\ & < \widetilde{\mu}_*(B^{\mathrm{c}} \cap J_2) = \widetilde{\mu}_*(\widetilde{\psi}^{-(s-r)}(B^{\mathrm{c}} \cap J_2))\\ & = \widetilde{\mu}_*(\widetilde{\psi}^{-(s-r)}(B^{\mathrm{c}}) \cap \widetilde{\psi}^{-(s-r)}(J_2)) = \mu_*(B^{\mathrm{c}} \cap J_1).\end{aligned}$$

Thus,

$$\widetilde{\mu}_*(B \cap J_1) + \widetilde{\mu}_*(B^{\mathrm{c}} \cap J_1) > 2(1-\varepsilon)\mu(J_1) = \frac{2}{3}\,(1+\theta)\mu(J_1) \ge \theta\mu(J_1),$$

as $\theta < 2$.

Comparing this with (3.2), with J_1 in place of J, gives an immediate contradiction, and the conclusion follows. ■

The simplest way in which (3.2) can be satisfied is when $\theta = 1$, in which case (3.2) has a formal resemblance to the Carathéodory definition, but perhaps it is preferable to think of (3.2) as a "Carathéodory-type" condition which enables us to deduce that a $\widetilde{\psi}$-invariant set is measurable because sets of measure zero are measurable. Note that it is shown in [7] that there are subsets B of $\widetilde{X}$ which are $\widetilde{\psi}$-invariant but which do not satisfy condition (3.2) and so are necessarily non-measurable. Now $\widetilde{\psi}$ is an ergodic transformation on $\widetilde{X}$, and Kakutani [6] (see also [2, pp. 29–30]) has shown that $\widetilde{\psi}$ is weakly mixing. Thus, Theorem 3.3 shows that Carathéodory's definition of a measurable set has a relationship to the invariant sets of the weakly mixing transformation $\widetilde{\psi}$, which is essentially the same as the relationship between Carathéodory's definition and the invariant sets of an irrational rotation on the unit circle.

REFERENCES

[1] R. Bartle, *The Elements of Integration*, Wiley, New York, 1966.

[2] J. R. Brown, *Ergodic Theory and Topological Dynamics*, Academic Press, New York, 1976.

[3] C. Carathéodory, *Über das lineare Mass von Punktmengen: eine Verallgemeinerung des Längenbegriffs*, Nachr. Gesell. Wiss. Göttingen 1914, 404–426.

[4] P. Halmos, *Measure Theory*, van Nostrand, 1965.

[5] S. Kakutani, *Induced measure preserving transformations*, Proc. Acad. Tokyo 16 (1943), 635–641.

[6] S. Kakutani, *Examples of ergodic measure preserving transformations which are weakly mixing but not strongly mixing*, in: Recent Advances in Topological Dynamics, A. Beck (ed.), Springer, New York, 1973, 143–149.
[7] A. Koeller, P. Nickolas, R. Nillsen and G. Williams, *Non-measurable invariant sets*, in preparation.
[8] R. Nillsen, *Irrational rotations motivate measurable sets*, Elemente Math. 56 (2001) 1–17.
[9] —, *Statistical Behaviour in Dynamical Systems*, lecture notes, School of Mathematics and Applied Statistics, Univ. of Wollongong, 1998.
[10] P. Walters, *An Introduction to Ergodic Theory*, Springer, 1982.

Institut für Mathematik
Fachbereich Mathematik und Informatik
Freie Universität Berlin
Arnimallee 2-6
14195 Berlin, Germany
E-mail: akoeller@mi.fu-berlin.de

School of Mathematics and Applied Statistics
University of Wollongong
Wollongong, NSW 2522 Australia
E-mail: nillsen@uow.edu.au
ghw@uow.edu.au

Received 3 January 2006;
revised 1 December 2006 (4709)