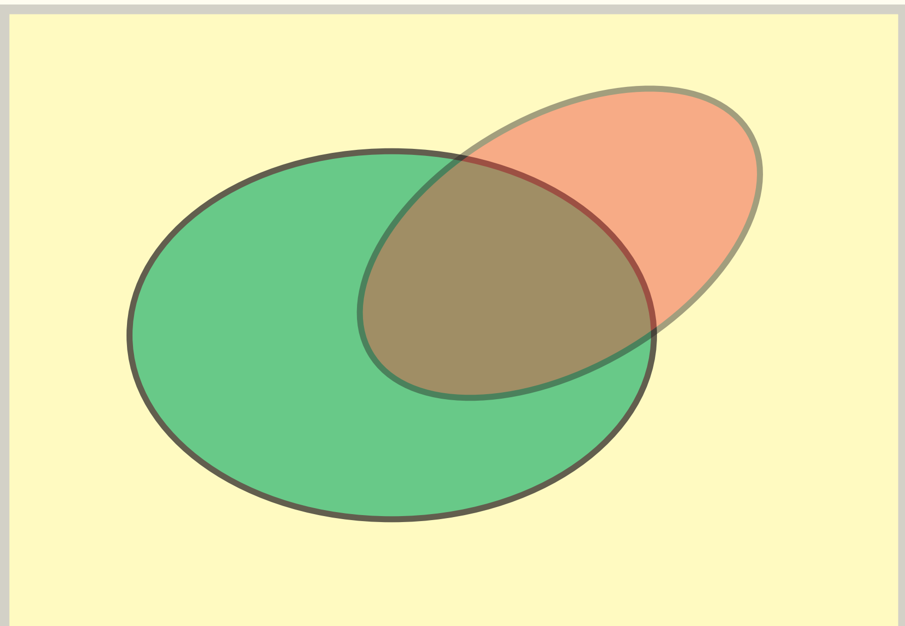
Dynamical systems and the Carathéodory definition of a measurable set

Rod Nillsen, Wollongong, NSW



Introduction

The main results in this talk have arisen from a project to see what results in analysis, which traditionally have required measure theory, can be done without measure theory. To do this, new techniques of proof have had to be developed. The book "Randomness and recurrence in dynamical systems" (MAA, 2011) is a result of this project.

However, the question arises as to whether new techniques of proof that avoid measure theory lead to new insights and/or new results.

In a paper of 1988, Igor Kluvánek referred to the paper of F. Riesz "Sets of measure zero and their rôle in analysis". He said:

> "The characterization of null sets and functions goes back to F. Riesz. He uses convergence of monotonic sequences though, rather than absolute summabiity. Riesz used his discovery of the rôle of null sets to great advantage in the simplification of integration theory."

Note that, at least on the real line, it is easy to give a definition of a set of measure zero. Namely, $A \subseteq \mathbb{R}$ has measure zero if for any $\varepsilon > 0$, there is a sequence (J_n) of intervals such that

$$A \subseteq \bigcup_{n=1}^{\infty} J_n \text{ and } \sum_{n=1}^{\infty} \mu(J_n) < \varepsilon.$$

Borel's Theorem without measure theory

Given a finite set A, $|A|$ denotes the number of elements in the set A.

Borel's Theorem (single digit version). *For $x \in [0,1)$, let $d_n(x)$ denote the n^{th} digit in the expansion of x to the base 2. Then, there is a set $\mathcal{Z}$ of measure zero such that, if $x \notin \mathcal{Z}$ we have:*

$$\lim_{n\to\infty} \frac{1}{n}|\{j : 1 \leq j \leq n \text{ and } d_j(x) = 0\}| = \frac{1}{2}.$$

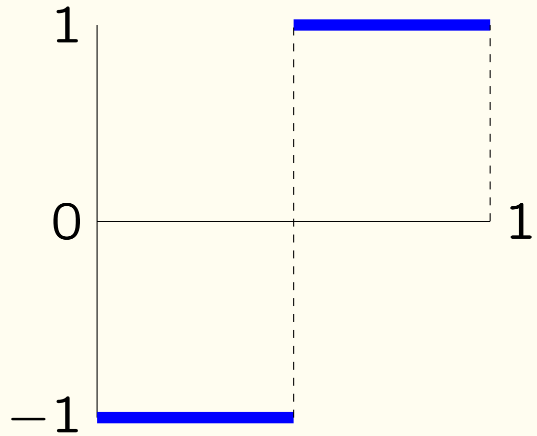
That is, for almost all numbers in $[0,1)$, there is an asymptotically equal proportion of 0s and 1s in the binary expansion of the number.

COMMENTS. The paper by F. Riesz (1939), uses the Rademacher functions (an approach he attributes to Khintchine). This is also the approach taken by Mark Kac in his "Statistical Independence in Probability, Analysis and Number Theory", MAA (1959).

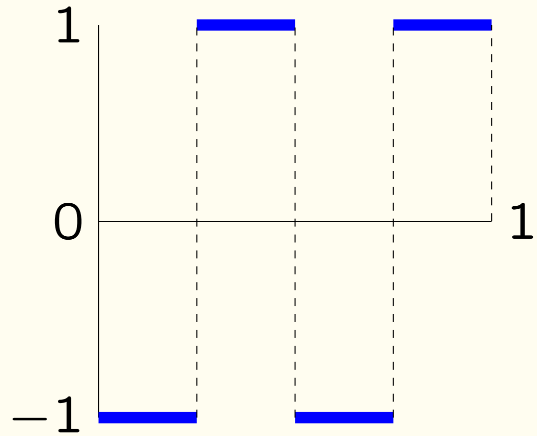
Also, G. Goodman has this approach in his paper (Amer. Math. Monthly, 1999) on the 40th anniversary of Kac's book. These use measure theory in an essential way but, in fact, measure theory can be avoided. I'll sketch the idea as it appeared in the Amer. Math. Monthly in 2000.

Let $r_n : [0,1) \longrightarrow \{-1,1\}$, be the n^{th} Rademacher function, given by

$$r_n(x) = \begin{cases} -1, \text{if } x \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right) \text{ and } k \text{ is odd,} \\ \\ 1, \text{if } x \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right) \text{ and } k \text{ is even.} \end{cases}$$



Rademacher function r_1



Rademacher function r_2

Now if $x \in [0,1)$, and if $d_n(x)$ denotes the n^{th} digit in the binary expansion of x, we have

$$d_n(x) = 0 \Longleftrightarrow r_n(x) = -1,$$

and

$$d_n(x) = 1 \Longleftrightarrow r_n(x) = 1.$$

Hence, Borel's result is equivalent to saying that for almost all $x \in [0,1)$,

$$\lim_{n\to\infty} \frac{1}{n}\Big(r_1(x) + r_2(x) + \cdots + r_n(x)\Big) = 0.$$

A routine calculation shows that

$$\int_0^1 \left(\frac{r_1(x) + \cdots + r_n(x)}{n}\right)^4 dx = \frac{3n-2}{n^3},$$

so that

$$\sum_{n=1}^{\infty} \int_0^1 \left(\frac{r_1(x) + \cdots + r_n(x)}{n}\right)^4 dx < \infty.$$

Now we use the following (proved without Beppo Levi).

Lemma. *Let* (ϕ_n) *be a sequence of real valued step functions on* $[0, 1)$ *such that*

$$\sum_{n=1}^{\infty}\left(\int_0^1 |\phi_n(x)|\,dx\right) < \infty.$$

Then, for almost all $x \in [0, 1)$,

$$\lim_{n\to\infty} \phi_n(x) = 0.$$

The Lemma applies to the functions $\phi_n = ((r_1 + \cdots + r_n)/n)^4$, so we deduce that for almost all $x \in [0, 1)$,

$$\lim_{n\to\infty}\left(\frac{r_1(x) + r_2(x) + \cdots + r_n(x)}{n}\right) = 0,$$

and Borel's result follows.

Proof of the lemma. Let (a_n) be such that $a_n \to \infty$ and

$$\sum_{n=1}^{\infty} a_n \int_{[0,1)} |\phi_n| d\mu < \infty.$$

We let $\mathcal{Z}$ be the set of points x such that $(\phi_n(x))$ does not converge to 0, and put

$$A_k = \left\{x : x \in [0,1) \text{ and } |\phi_k(x))| > \frac{1}{a_k}\right\}. \text{ Then}$$

$$\mathcal{Z} \subseteq \bigcap_{n=1}^{\infty} \left(\bigcup_{k=n}^{\infty} A_k \right). \text{ This gives}$$

$$\mu(A_k) \;=\; \int_{A_k} 1 \; d\mu \leq a_k \int_{A_k} |\phi_k| \; d\mu, \text{ so } \sum_{k=1}^{\infty} \mu(A_k) < \infty.$$

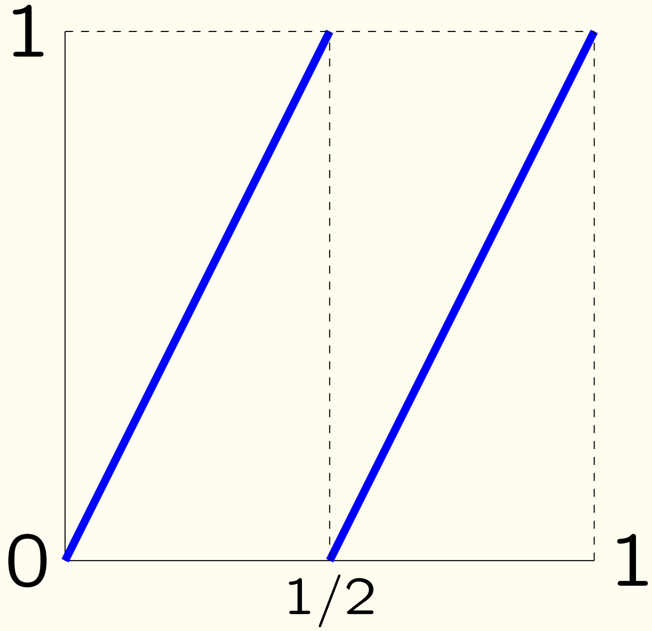
Each A_k is a finite union of intervals, so $\mathcal{Z}$ has measure zero.

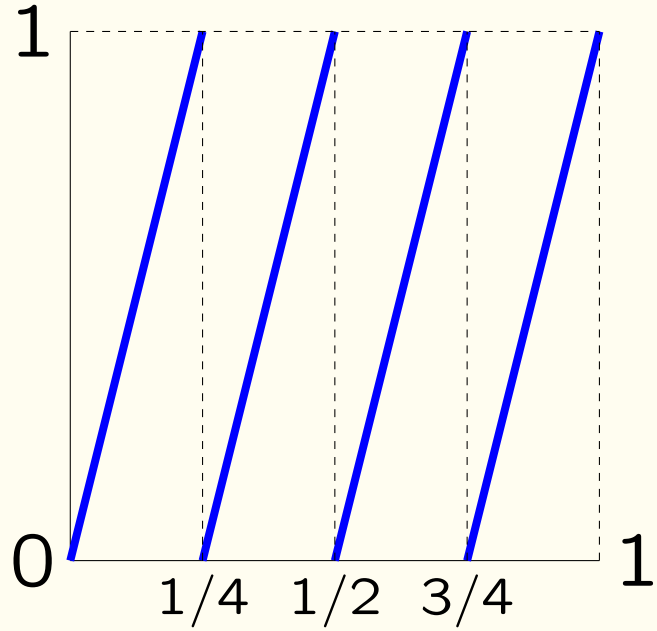
The dynamical systems version of Borel's Theorem seems to be due to D. D. Wall (1948). We define $f : [0,1) \longrightarrow [0,1)$ by

$$f(x) = \begin{cases} 2x, \text{if } 0 \leq x < 1/2, \\ \\ 2x - 1, \text{if } 1/2 \leq x < 1 \end{cases}$$

The connection is due to the fact that, denoting the n^{th} binary digit of x by $d_n(x)$,

$$d_n(f(x)) = d_{n+1}(x).$$





On the left is depicted the function f; on the right the function $f^2 = f \circ f$.

Borel's Theorem (dynamical systems form). For almost all $x \in [0, 1)$,

$$\lim_{n\to\infty} \frac{1}{n} \Big| \Big\{ x : 1 \leq j \leq n \text{ and } f^j(x) \in [0, 1/2) \Big\} \Big| = \frac{1}{2}.$$

Measurable sets

In the general theory of integration, measure theory and, in particular, the notion of a *measurable set*, play a central rôle.

Various definitions of the notion of a measurable set have been used, but perhaps the most famous is the one given by Carathéodory in 1914. Caratheodory's definition depends on the notion of outer measure.

Given a set X an **outer measure** on X is a function μ_* whose domain is the set of all subsets of X and which has the following properties (A^c is the complement of A):

(i) $\mu_*(\emptyset) = 0$, where $\emptyset$ denotes the empty set,

(ii) $\mu_*(C) \in [0, \infty]$ for all subsets C of X, and

(iii) if (C_n) is any sequence of subsets of X, then

$$\mu_* \left(\bigcup_{n=1}^{\infty} C_n \right) \leq \sum_{n=1}^{\infty} \mu_*(C_n).$$

Then, a subset A of X is called **measurable** if, for all subsets B of X,

$$\mu_*(B) = \mu_*(A \cap B) + \mu_*(A^c \cap B).$$

The significance of this definition lies in the facts that the set of all measurable subsets of X forms a sigma algebra upon which the outer measure is countably additive.

Note, that for the time being, we assume no knowledge of this fact.

Carathéodory's definition has frequently been the subject of comment or defence in a way unusual for the definition of a mathematical concept.

In *Real and Abstract Analysis* Edwin Hewitt and Karl Stromberg write:

"How Carathéodory came to think of this definition seems mysterious, since it is not in the least intuitive. Carathéodory's definition has many useful implications".

Also, Paul Halmos comments in his *Measure Theory*: “It is rather difficult to get an understanding of the meaning of ...... measurability except through familiarity with its implications...The greatest justification of this apparently complicated concept is, however, its possibly surprising but absolutely complete success as a tool in proving the important and useful extension theorem”.

However, not everyone has accepted that the definition of a measurable set can be justified by appealing to a usefulness which may be at the time quite unclear but which will be amply justified in the future. In his *Conjectures and Refutations* Imre Lakatos is critical about Michel Loève’s treatment of measurable sets.

Although the motivation for the work here did not derive from the objections of Lakatos, the work here shows how Carathéodory's definition of a measurable set arises from a specific "problem situation", a pedagogical approach recommended by Lakatos.

DYNAMICAL SYSTEMS

Let X be a set and let $T : X \longrightarrow X$ be a function on X. Such a function is often called a *transformation* on X. A **dynamical system** is then such a pair (X, T).

Let T^0 denote the identity transformation, and if $n \in \mathbb{N}$, the set of natural numbers, T^n denotes the composition of T with itself n times. If A is a subset of X and $n = 0$ or $n \in \mathbb{N}$, the sets $T^n(A)$ and $T^{-n}(A)$ are the sets given by

$$T^n(A) = \{T^n(x) : x \in A\} \text{ and}$$

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

Let (X,T) be a given dynamical system. Then if $A \subseteq X$, A is called T-**invariant** or **invariant** if $T^{-1}(A) = A$.

Clearly $\emptyset$ and X are invariant. If T is invertible,

$$A \text{ is invariant} \iff A = T^{-1}(A) \iff A = T(A).$$

ERGODIC TRANSFORMATIONS

Let $T : X \longrightarrow X$ be a given transformation. Let $\mathcal{S}$ be a σ-algebra of subsets of X , and let ν be a probability measure on X that is $\mathcal{S}$-measure preserving. Then T is **ergodic** if:

$$A \in \mathcal{S} \text{ and } A \text{ is invariant } \Longrightarrow \nu(A) = 0 \text{ or } \nu(A) = 1.$$

The unit circle $\mathbb{T}$ is the set of all complex numbers of modulus one.

A **rotation** on $\mathbb{T}$ is a function $\rho : \mathbb{T} \longrightarrow \mathbb{T}$ for which there exists $z \in \mathbb{T}$ such that

$$\rho(w) = zw, \quad \text{for all} \quad w \in \mathbb{T}.$$

A rotation is invertible.

If $z = e^{2\pi i\alpha}$ where α is irrational, the rotation is called **irrational**, otherwise the rotation is **rational**.

There are two main approaches for proving that an irrational rotation ρ is ergodic and both assume a background in measure theory.

The first uses Fourier analysis techniques in $L^2(\mathbb{T})$ to show that if χ_A is the characteristic function of a measurable ρ-invariant set A, χ_A is constant, up to a set of measure zero. It follows that either χ_A is 0 or 1 in $L^2(\mathbb{T})$, so A or A^c has measure zero.

The second approach uses of points of density.

The first assumes a knowledge of measure and integration, complete orthonormal sets and Fourier expansions in L^2-spaces; while the second presupposes a knowledge of measure theory at a level of refinement which is not generally attained in an undergraduate course.

The ergodicity problem on the unit circle

As formulated the ergodicity of an irrational rotation on the circle group requires a knowledge of measures on σ-algebras. If we do not not wish to assume this knowledge, we might ask the following.

QUESTION. Let A be an invariant set in $\mathbb{T}$ for the irrational rotation ρ. Is the **outer measure** of A necessarily 0 or 1?

The answer is "no", but the answer is "yes" if A satisfies a type of "Carathéodory condition". Thus, by considering the outer measure of an invariant set in $(\mathbb{T}, \rho)$, we are led to the Carathéodory definition.

Outer measure, rotations and invariant sets

A subset V of $\mathbb{T}$ is called an *arc* either if there is a half open subinterval J of $[0,1)$ such that

$$V = \Big\{\exp(2\pi it) : t \in J\Big\},$$

or if there are disjoint half open subintervals K, L of $[0,1)$ such that K is of the form $[0,b)$, L is of the form $[c,1)$, and

$$V = \Big\{\exp(2\pi it) : t \in K \cup L\Big\}.$$

We put

$$\mu(V) = \text{length of } J \text{ or length of } K \cup L.$$

Thus,

$$\mu(\mathbb{T}) = 1.$$

We introduce the outer measure μ_* on $\mathbb{T}$ in the usual way. That is, if $A \subseteq \mathbb{T}$,

$$\mu_*(A) = \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) : A \subseteq \bigcup_{n=1}^{\infty} A_n \text{ and each } A_n \text{ is an arc} \right\}.$$

Arcs of density The following result establishes the existence of "arcs of density" for sets of positive outer measure.

LEMMA. *Let $A \subseteq \mathbb{T}$, let $\mu_*(A) > 0$ and let $\varepsilon > 0$. Then there is a non-empty arc J of $\mathbb{T}$ such that*

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

Proof. Assume that the result is false. Then, for any non-void arc W of t,

$$\mu_*(A \cap W) \leq (1 - \varepsilon)\mu(W).$$

Now let $\delta > 0$. By the definition of $\mu_*(A)$ there is a sequence (W_n) of non-void arcs of $\mathbb{T}$ such that

$$A \subseteq \bigcup_{n=1}^{\infty} W_n \text{ and } \sum_{n=1}^{\infty} \mu(W_n) < \mu_*(A) + \delta.$$

Then $A \subseteq \cup_{n=1}^{\infty}(A \cap W_n)$, so it follows that

$$\begin{aligned}
\mu_*(A) &\leq \sum_{n=1}^{\infty} \mu_*(A \cap W_n), \\
&\leq (1 - \varepsilon) \sum_{n=1}^{\infty} \mu(W_n), \\
&\leq (1 - \varepsilon)(\mu_*(A) + \delta).
\end{aligned}$$

Hence

$$\mu_*(A) \leq \frac{\delta(1-\varepsilon)}{\varepsilon},$$

and this is true for all $\delta > 0$. Hence, if the result is false, $\mu_*(A) = 0$, which contradicts the assumed fact that $\mu_*(A) > 0$. $\square$

The following result is a refinement of the preceding one.

LEMMA. *Let $A \subseteq \mathbb{T}$, let $\mu_*(A) > 0$ and let $\varepsilon > 0$. Then there is a number $\theta > 0$ which has the following property: if $\eta \in (0, \theta)$, there is an arc J of $\mathbb{T}$ such that*

$$\mu(J) = \eta \quad \textit{and} \quad \frac{\mu_*(A \cap J)}{\mu(J)} \; > \; 1 - \varepsilon.$$

LEMMA. *Let z be an element of $\mathbb{T}$ which is not a root of unity, let J, K be non-empty arcs in $\mathbb{T}$ which are of equal length, and let $\varepsilon > 0$. Then there is some $n \in \mathbb{N}$ such that*

$$\mu\left(z^n J \Delta K\right) < \varepsilon.$$

Proof. Ideas only. The conclusion of the Lemma may seem more obvious if it is thought of in purely geometric terms. For, if we have two arcs of equal length, there is a rotation which exactly rotates one arc into the other–but by Kronecker's Theorem this rotation may be approximated by some iterate of the given irrational rotation, so that such an iterate will "approximately" rotate one arc into the other. $\square$

The following result implies the usual ergodicity of an irrational rotation on $\mathbb{T}$. It is clearly related to the Carathéodory definition.

Theorem. *Let $z \in \mathbb{T}$, z not a root of unity, and let $\rho : \mathbb{T} \longrightarrow \mathbb{T}$ be the rotation $u \longmapsto zu$. Let A be a ρ-invariant subset of $\mathbb{T}$. Suppose that there is $1 \leq \theta < 2$ such that for all arcs J of $\mathbb{T}$,*

$$\mu_*(A \cap J) + \mu_*(A^c \cap J) \leq \theta\mu(J).$$

Then, either A or A^c is a set of outer measure zero.

Proof. Let $\rho(A) = A$, and assume that $\mu_{*(A)} > 0$ and that $\mu_*(A^c) > 0$. Let $\varepsilon > 0$. By the Lemmas, there are arcs J, K such that $\mu(J) = \mu(K)$ and

$$\frac{\mu(A \cap J)}{\mu(J)} > 1 - \varepsilon \text{ and } \frac{\mu(A^c \cap K)}{\mu(K)} > 1 - \varepsilon.$$

As J, K have equal length, there is $n \in \mathbb{N}$ such that

$$\mu(z^{-n} J \Delta K) < \delta.$$

Now observe that

$$A^c \cap K \subseteq \left(A^c \cap z^{-n} J\right) \bigcup \left(z^{-n} J \Delta K\right),$$

so that

$$\mu_*(A^c \cap K) \leq \mu_* \left(A^c \cap z^{-n} J\right) + \mu \left(z^{-n} J \Delta K\right).$$

Using this fact we now have

$$\begin{aligned}
\mu_*(A^c \cap J) &= \mu_*(\rho_{z^{-n}}(A^c \cap J)), \\
&= \mu_*(z^{-n}(A^c \cap J)), \\
&= \mu_*\left(z^{-n}A^c \cap z^{-n}J\right), \\
&= \mu_*(A^c \cap z^{-n}J), \quad \text{as } A^c \text{ is } \rho_z\text{-invariant}, \\
&\geq \mu_*(A^c \cap K) - \mu(z^{-n}J \Delta K), \\
&\geq \mu_*(A^c \cap K) - \delta, \\
&\geq (1-\varepsilon)\mu(K) - \delta, \\
&= (1-\varepsilon)\mu(J) - \delta,
\end{aligned}$$

as $\mu(J) = \mu(K)$. Note that A and J are independent of δ and that this holds for *all* $\delta > 0$. Hence,

$$\mu_*(A^c \cap J) \geq (1-\varepsilon)\mu(J).$$

We now have

$$\mu_*(A \cap J) + \mu_*(A^c \cap J) \geq 2(1 - \varepsilon)\mu(J).$$

Recall that ε was *any* number such that $\varepsilon > 0$, so that $2(1 - \varepsilon)$ may be as close to 2 as specified in advance. So, summarizing the discussion we have:

if z is an element of $\mathbb{T}$ which is not a root of unity, and if A is a ρ_z-invariant subset of $\mathbb{T}$ with the properties that $\mu_*(A) > 0$ and $\mu_*(A^c) > 0$; then for each positive number $\eta < 2$ there is an arc J of $\mathbb{T}$, this arc depending upon η, such that

$$\mu_*(A \cap J) + \mu_*(A^c \cap J) \geq \eta \ \mu(J).$$

The result follows. $\square$

Now $(\mathbb{T}, \rho)$ is a dynamical system and ρ is not week mixing. Are there other dynamical systems that lead to the Carathéodory definition?

If $\xi : \mathbb{T} \longrightarrow \mathbb{T}$ is given by $z \longmapsto z^2$, then ξ is strong mixing: we have the following.

Theorem. *If A is a subset of $\mathbb{T}$ that is ξ-invariant, then $\mu_*(A) = 0$ or $\mu_*(A) = 1$.*

Hence, the system $(\mathbb{T}, \xi)$ does not lead to the Carathéodory definition.

Question: is there an example of a weakly mixing transformation which leads to the definition of Carathéodory?

Answer: The example of a weakly mixing but not strongly mixing transformation, given in 1973 by S. Kakutani, but which he states goes back to a conversation in 1941 with John von Neumann, is an example of such a transformation.

This is joint work with G. Williams and A. Koeller (Coll. Math. 2007).

Mixing transformations

Let X be a set, let $\mathcal{S}$ be a σ-algebra of subsets of X and let ν be a probability measure on $\mathcal{S}$. Let $T : X \longrightarrow X$. Then, T is **measure-preserving** if for all $A \in \mathcal{S}$,

$$T^{-1}(A) \in \mathcal{S} \text{ and } \nu(T^{-1}(A)) = \nu(A).$$

We shall say that a measure preserving transformation $T : X \longrightarrow X$ is **weakly mixing** if for all $A, B \in \mathcal{S}$,

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

If we assume that $T^{-1}(A) = A$ and take $B = A$, we deduce that $\mu(A) - \mu(A)^2 = 0$, so that $\mu(A) \in \{0, 1\}$. Thus, a weakly mixing transformation is ergodic.

Strongly mixing transformations

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

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

We see that a strongly mixing transformation is weakly mixing.

An irrational rotation on the unit circle $\mathbb{T}$ is ergodic but is not weakly mixing. The transformation $z \longmapsto z^2$ is strongly mixing on $\mathbb{T}$. An example of a weakly mixing transformation which is not strongly mixing is due to Kakutani. We proceed to give a description of this example.

Induced systems

Let X be a set, let $\mathcal{B}$ be a σ-algebra of subsets of X, let μ be a measure on X, 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 $\mu(A^c) > 0$. Let A' be a set disjoint from X and such that there is a one-to-one mapping τ from A onto A'. Make the definitions that

$$\widetilde{X} = X \cup A',$$

$$\widetilde{\mathcal{B}} = \{B : B \subseteq \widetilde{X},\ B \cap X \in \mathcal{B} \text{ and } \tau^{-1}(B \cap A') \in \mathcal{B}\}, \text{ and}$$

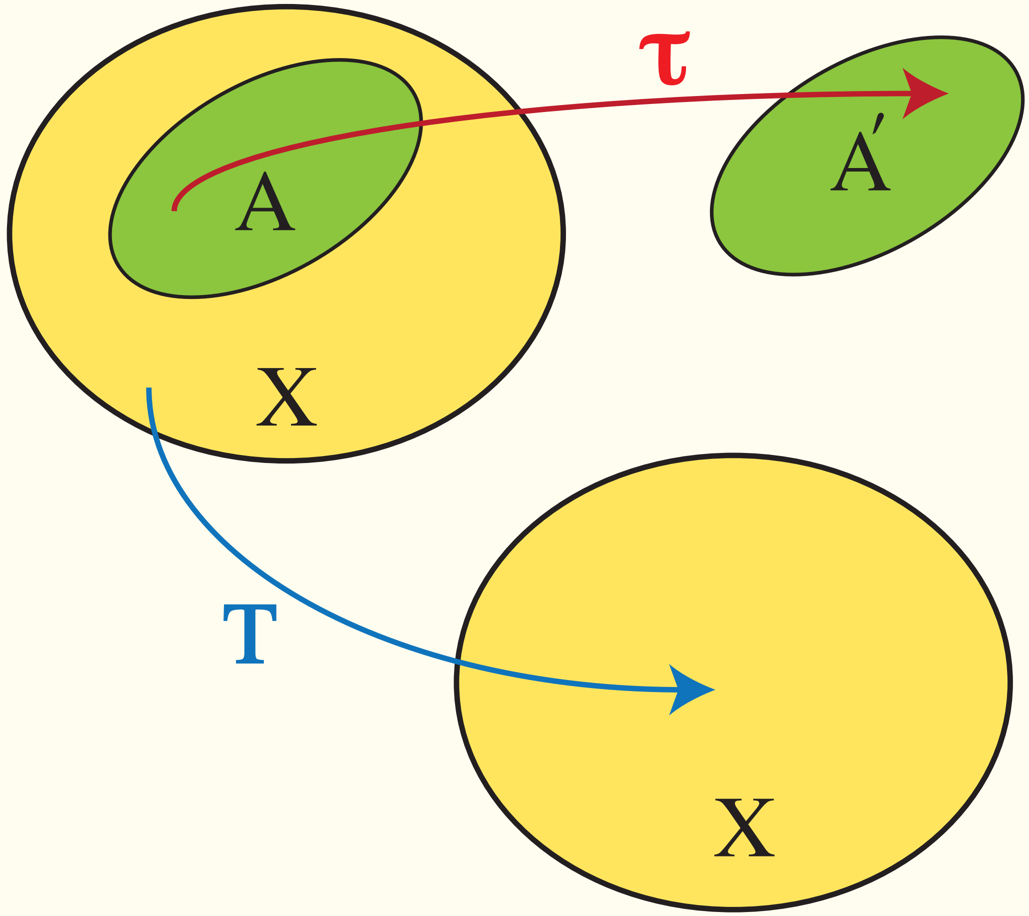
$$\widetilde{\mu}(B) = \mu(B \cap X) + \mu(\tau^{-1}(B \cap A')), \text{ for all } B \in \widetilde{\mathcal{B}}.$$

Then, it is easy to check that $\widetilde{\mathcal{B}}$ is a σ-algebra of subsets of $\widetilde{X}$ and that $\widetilde{\mu}$ is a measure on $\widetilde{\mathcal{B}}$.

Moreover, if we define the transformation $\tilde{T}$ on $\widetilde{X}$ by

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

it is easy to check that $\tilde{T}$ is a measure preserving transformation on $\widetilde{X}$.



The transformation $\widetilde{T} : X \cup A' \longrightarrow X \cup A'$.

That is, $\widetilde{T} : \widetilde{X} \longrightarrow \widetilde{X}$.

Given A and A' as described, $(\widetilde{X}, \tilde{\mathcal{B}}, \tilde{\mu}, \tilde{T})$ can be constructed from $(X, \mathcal{B}, \mu, T)$, as described. This construction is a special case of an induced measure preserving transformation, introduced by Kakutani, in the sense that the action of T on X induces the action of $\tilde{T}$ on $\widetilde{X}$. It can be seen that if T is one-to-one, then $\tilde{T}$ is one-to-one; that the range of $\tilde{T}$ is the union of the range of T with A'; and that if T is ergodic then $\tilde{T}$ is ergodic.

Kakutani's example arises from a special case of the construction above.

Henceforth, we take X to be the unit interval $[0, 1)$, we take $\mathcal{B}$ to be the family of Borel subsets of X.

We take μ to be the usual Lebesgue measure on $\mathcal{B}$. For $n = 0, 1, 2, \ldots,$ put

$$I_n = \left[\, 1 - \frac{1}{2^n}, 1 - \frac{1}{2^{n+1}}\right),$$

and observe that

$$X = [0, 1) = \bigcup_{n=0}^{\infty} I_n.$$

We take A to be the subset of X given by

$$A = \bigcup_{n=0}^{\infty} I_{2n},$$

and we 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

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

The transformation ψ on X is measure preserving. Then, as in the construction described above, we construct $(\widetilde{X}, \widetilde{\mathcal{B}}, \widetilde{\mu}, \widetilde{\psi})$ from $(X, \mathcal{B}, \mu, \psi)$. Note that $\widetilde{\psi}$ is $\widetilde{\mu}$-measure preserving.

Note that ψ maps the interval I_n linearly onto the interval $[1/2^{n+1}, 1/2^n)$, and that

$$(0,1) = \bigcup_{n=0}^{\infty} \Big[\frac{1}{2^{n+1}}, \frac{1}{2^n}\Big),$$

where this union is disjoint. It follows from the above that ψ is a function mapping $[0,1)$ onto $(0,1)$ and that ψ is one-to-one. The transformation ψ on X is ergodic.

However, ψ is not weakly mixing, for $\psi(0, 1/2) = (1/2, 1)$ and $\psi(1/2, 1) = (0, 1/2)$. Consequently, if we take $A = B = (0, 1/2)$ we have

$$\lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ |\mu(\psi^{-i}(A) \cap B) - \mu(A)\mu(B)| \quad = \frac{1}{4}$$

so that ψ is not weakly mixing.

Now, because ψ is one-to-one, $\tilde{\psi}$ is also one-to-one. Also, because ψ is ergodic, $\tilde{\psi}$ is ergodic. The range of $\tilde{\psi}$ is $(0,1) \cup A'$. That is, the range of $\tilde{\psi}$ is $\widetilde{X}$ except for the single point 0. The function $\tilde{\psi}$ was defined by Kakutani, and he showed that $\tilde{\psi}$ is measure preserving and weak mixing on $\widetilde{X}$, but that $\tilde{\psi}$ is not strong mixing.

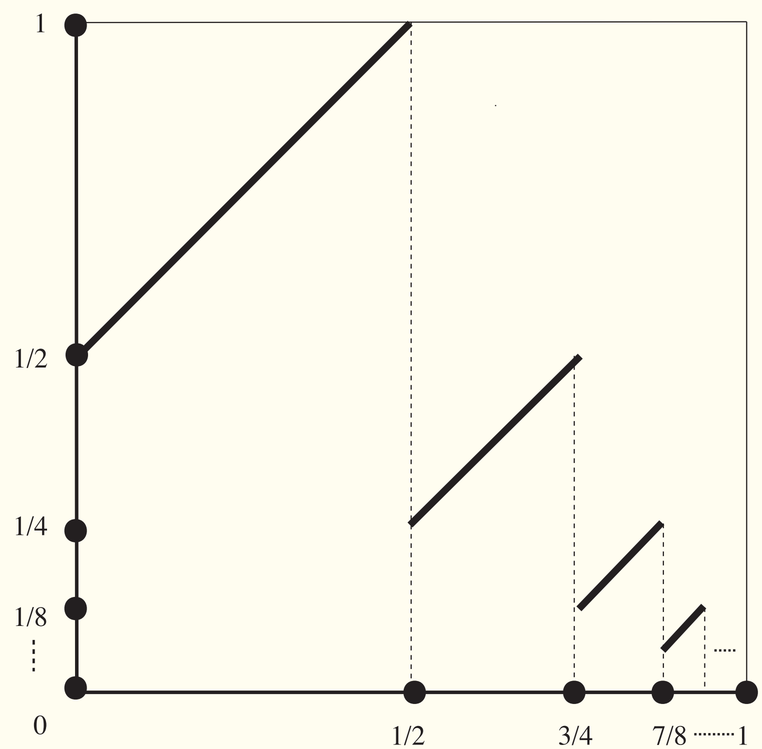


Figure 1. The graph of the transformation ψ. The picture illustrates the first four of the infinite sequence of line segments.

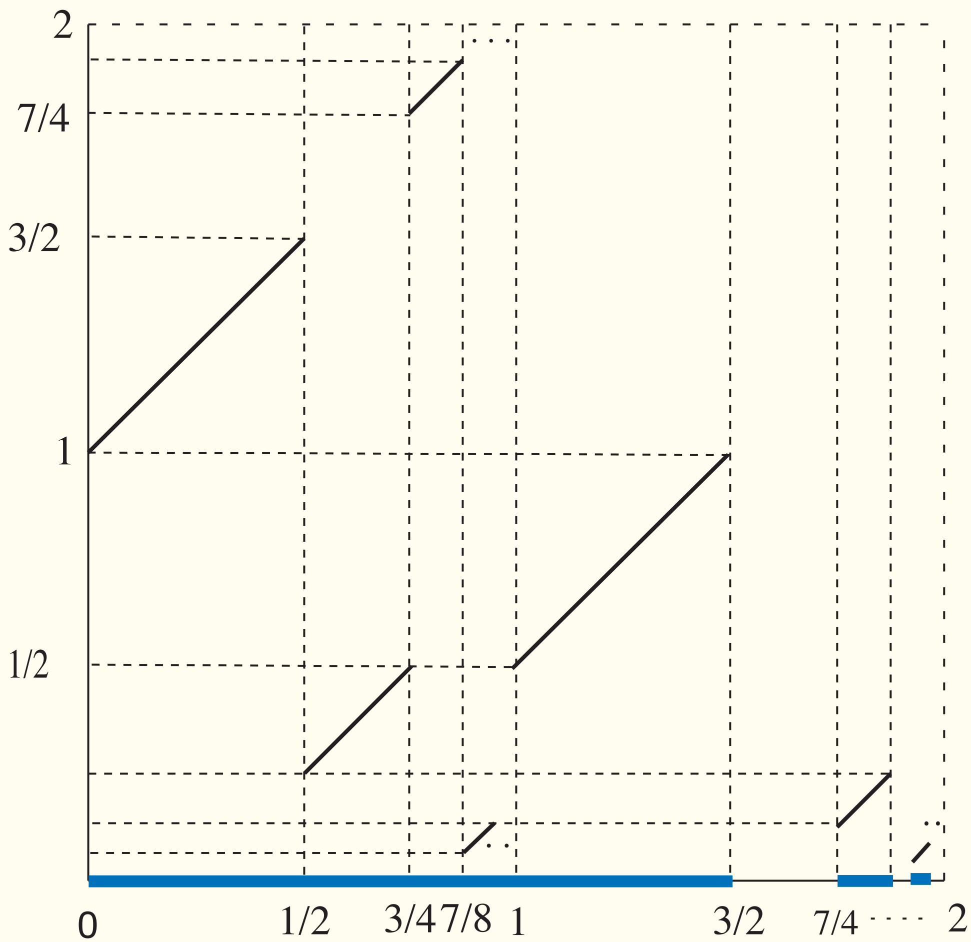


Figure 2. The graph of the induced transformation $\widetilde{\psi}$, obtained when we take $\tau(x) = x + 1$.

We started with the system $(X, \mathcal{B}, \mu, \psi)$ and we have induced the system $(\widetilde{X}, \tilde{\mathcal{B}}, \tilde{\mu}, \tilde{\psi})$. The outer measure from $\tilde{\mu}$ is denoted by $\tilde{\mu}_*$.

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

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

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

Proof. This is along the lines of the proof for rotations, but is considerably more technical. We prove by contradiction.

IDEAS. We assume that $\widetilde{\mu}_*(B) > 0$ and that $\widetilde{\mu}_*(B^c) > 0$. An open subinterval J of $[0,1)$ is called **dyadic** if there are $n, q \in \mathbb{N}$ such that

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

In this case n is the **order** of J. Put

$$\varepsilon = \frac{2-\theta}{3} > 0.$$

FACT I. There is n_0 such that if $n > n_0$ with $n \in \mathbb{N}$, there are dyadic intervals J_1, J_2 of order n such that

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

[Idea is that $\mu_*(B \cap X) > 0$ and $\mu_*(B^c \cap X) > 0$ and $\widetilde{\mu}_*(C) = \mu_*(C)$ when $C \subseteq X$.

FACT II. As J_1, J_2 are dyadic intervals of the same order n, they belong to the $\widetilde{\psi}$-orbit of $(0, 1/2^n)$. There are r, s such that

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

FACT III. If $B \subseteq \widetilde{X}$, then

$$\widetilde{\mu}_* \left(\widetilde{\psi}^{-1}(B)\right) = \widetilde{\mu}_*(B).$$

NOW, assume that $r \geq s$. Then,

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

WE NOW HAVE: $\mu(J_1) = \mu(J_2)$, $\widetilde{\mu}_*$ is $\widetilde{\psi}$-invariant, and B, B^c are $\widetilde{\psi}$-invariant sets. So, ...

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

Hence, recalling that

$$\theta < 2 \text{ and } \varepsilon = (2-\theta)/3,$$

we have

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

a contradiction.

THE END

THANK YOU