Ergodic Theory

José F. Alves

CIMPA School
Recent Advances in Dynamical Systems

November 2021

Centro de Matemática
Universidade do Porto

https://tinyurl.com/alves-cimpa

Measure

Measurable spaces

Given a set X, let $\mathcal{P}(X)$ be the collection of all subsets of X.
We say that $\mathcal{A} \subset \mathcal{P}(X)$ is a σ-algebra if the following conditions hold:

(1) $\emptyset \in \mathcal{A}$;

(2) $A \in \mathcal{A} \implies X \setminus A \in \mathcal{A}$;

(3) $A_1, A_2, \cdots \in \mathcal{A} \implies \cup_{i=1}^{\infty} A_i \in \mathcal{A}$.

$(X, \mathcal{A})$ is called a measurable space and the sets in $\mathcal{A}$ are called measurable sets.

Example 1.1

$\mathcal{A} = \mathcal{P}(X)$ is a σ-algebra.

Exercise 1.2

Show that

1. $A, B \in \mathcal{A} \implies A \cup B \in \mathcal{A}$, $A \cap B \in \mathcal{A}$ and $A \setminus B \in \mathcal{A}$;
2. $A_1, A_2, \cdots \in \mathcal{A} \implies \cap_{i=1}^{\infty} A_i \in \mathcal{A}$.

Borel sets

Given any family $\mathcal{F} \subset \mathcal{P}(X)$, the intersection of all σ-algebras containing $\mathcal{F}$ † is a σ-algebra, called the σ-algebra generated by $\mathcal{F}$.

In case X is a topological space, the σ-algebra $\mathcal{B}_X$ generated by the open sets is called the Borel σ-algebra on X, and its elements are called Borel sets.

Exercise 1.3

Show that

1. $\mathcal{B}_{\mathbb{R}}$ contains the family $\mathcal{I}$ of all intervals in $\mathbb{R}$;
2. the σ-algebra generated by $\mathcal{I}$ coincides with $\mathcal{B}_{\mathbb{R}}$.

Considering $\mathbb{R}$ with the usual structure of topological space, it is a non-obvious fact that

$$\mathcal{B}_{\mathbb{R}} \neq \mathcal{P}(\mathbb{R}). \tag{1}$$

† there exists at least one: $\mathcal{P}(X)$.

Measures

Given a σ-algebra $\mathcal{A} \subset \mathcal{P}(X)$, we say that $\mu : \mathcal{A} \to [0, +\infty]$ is a measure if

1. $\mu(\emptyset) = 0$;
2. if $A_1, A_n, \dots$ are pairwise disjoint sets in $\mathcal{A}$, then

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mu(A_i). \tag{2}$$

We say that μ is a probability measure if $\mu(X) = 1$. We refer to $(X, \mathcal{A}, \mu)$ as measure space or a probability measure space, in case μ is a probability measure.

Exercise 1.4

Show that

1. property (2) holds for finite disjoint unions;
2. given $A, B \in \mathcal{A}$,

$$A \subset B \implies \mu(A) \le \mu(B) \tag{3}$$

$$A \subset B \quad \text{and} \quad \mu(A) < \infty \implies \mu(B \setminus A) = \mu(B) - \mu(A). \tag{4}$$

We say that a property about the elements in X holds μ almost everywhere (a.e.) if the set N for which the property does not hold has $\mu(N) = 0$.

Example 1.5 (Counting measure)

Given a set X, consider $\nu : \mathcal{P}(X) \to [0, +\infty]$ defined as

$$\nu(A) = \begin{cases} \#A, & \text{if } A \text{ is finite;} \\ +\infty, & \text{otherwise.} \end{cases}$$

ν defines a measure, called the counting measure in X.

Example 1.6 (Dirac measure)

Given a set X and $x \in X$, consider $\delta_x : \mathcal{P}(X) \to [0, +\infty]$ defined as

$$\delta_x(A) = \begin{cases} 1, & \text{if } x \in A; \\ 0, & \text{if } x \notin A. \end{cases}$$

δ_x defines a probability measure, called the Dirac measure at x.

Exercise 1.7

Show that ν and δ_x in the previous examples are measures.

Lebesgue measure

Let $\mathcal{I}$ be the family of subintervals of an interval $J \subset \mathbb{R}$. Though $\mathcal{I}$ is not a σ-algebra, we have a notion of length $\ell : \mathcal{I} \to [0, +\infty]$, defined for $I \in \mathcal{I}$ as

$$\ell(I) = \begin{cases} \sup(I) - \inf(I), & \text{if } I \neq \emptyset; \\ 0, & \text{if } I = \emptyset. \end{cases}$$

Theorem 1.8

There exists a unique measure $\lambda : \mathcal{B}_J \to [0, +\infty]$ *such that* $\lambda(I) = \ell(I), \forall I \in \mathcal{I}$.

λ is called the Lebesgue measure on J. See e.g. [Barra 2003] or [Halmos 1950] for a proof of Theorem 1.8. Standard proofs give that, for all $A \in \mathcal{B}_J$,

$$\lambda(A) = \inf \left\{ \sum_{n=1}^{\infty} \lambda(I_n) : I_1, I_2, \cdots \in \mathcal{I} \text{ and } A \subset \bigcup_{n \geq 1} I_n \right\}. \qquad (5)$$

Remark 1.9

Similar conclusions hold in the circle $\mathbb{S}^1$, with length replaced by *arc length*. Lebesgue measure can actually be introduced in any $\mathbb{R}^n$ (or any Riemannian manifold), generalizing our intuitive notion of length, area, volume...

Exercise 1.10

1. Show that Lebesgue measure on $\mathbb{R}$ is *translation invariant*, i.e.

$$\lambda(x+B) = \lambda(B), \quad \text{for all } B \in \mathcal{B}_{\mathbb{R}}.$$

2. Show that there is no measure $\mu \colon \mathcal{P}(\mathbb{R}) \to [0,+\infty]$ such that

$$\mu((0,1]) = 1 \quad \text{and} \quad \mu(x+A) = \mu(A), \quad \text{for all } A \subset \mathbb{R}.$$

Hint: Arguing by contradiction, consider the equivalence relation $\sim$ in $\mathbb{R}$ given by $x \sim y \iff x - y \in \mathbb{Q}$. Define a set $A \subset (0,1]$ choosing a single element from each equivalence class. Denoting by R the set of rational numbers in $(-1,1)$, show that the sets $z+A$, with $z \in R$, are pairwise disjoint and

$$(0,1) \subset \bigcup_{z\in R} (z+A) \subset (-1,2].$$

Deduce that $1 \leq \sum_{z\in R} \mu(z+A) \leq 3$, which is not possible.

The previous exercise shows that there is no reasonable extension of Lebesgue measure to $\mathcal{P}(\mathbb{R})$; recall (1). Standard proofs of Theorem 1.8 give that λ can actually be extended to a σ-algebra $\mathcal{M}$ such that

$$\mathcal{B}_{\mathbb{R}} \subsetneq \mathcal{M} \subsetneq \mathcal{P}(\mathbb{R}).$$

Integration

Measurable functions

Let $(X, \mathcal{A})$ me a measurable space. We say that $f : X \to \mathbb{R}$ is measurable if

$$f^{-1}(B) \in \mathcal{A}, \quad \forall B \in \mathcal{B}_{\mathbb{R}}.$$

Exercise 2.1

Show that the characteristic function χ_A of any measurable set $A \in \mathcal{A}$ is measurable.

Recall that

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A; \\ 0, & \text{if } x \notin A. \end{cases}$$

Exercise 2.2

Show that if $c \in \mathbb{R}$ and $f, g \colon X \to \mathbb{R}$ are measurable functions, then c, cf, $|f|$, $f \pm g$, fg, f/g (when it makes sense), $\max\{f, g\}$ and $\min\{f, g\}$ are measurable.

Exercise 2.3

Show that if X is a topological space and $\mathcal{A}$ is the σ-algebra of Borel sets in X, then any continuous function $f : X \to \mathbb{R}$ is measurable.

Integral of a simple function

Let $(X, \mathcal{A}, \mu)$ be a measure space.
We say that $\varphi : X \to [0, +\infty)$ is a simple function, if it can be written as

$$\varphi = \sum_{i=1}^{n} a_i \chi_{A_i}, \tag{6}$$

with $A_i \in \mathcal{A}$ and $a_i \geq 0$. We define the integral of the simple function φ as

$$\int \varphi d\mu = \sum_{i=1}^{n} a_i \mu(A_i),$$

with the convention that $0 \cdot \infty = 0$.

Exercise 2.4

Show that the value of $\int \varphi d\mu$ does not depend on the representation of φ in (6).

We have in particular for all $A \in \mathcal{A}$

$$\boxed{\int \chi_A d\mu = \mu(A).} \tag{7}$$

Integral of a nonnegative function

We define the integral of a nonnegative measurable function $f : X \to [0, +\infty)$ as

$$\int f \, d\mu = \sup \left\{ \int \varphi d\mu \colon \varphi \text{ simple function and } \varphi \leq f \right\}.$$

It follows that if f and g are nonnegative measurable functions, then

$$f \leq g \quad \Rightarrow \quad \int f \, d\mu \leq \int g \, d\mu.$$

Exercise 2.5

Show that if $f : X \to [0, +\infty)$ is a measurable function, then

$$\int f \, d\mu = 0 \iff f = 0, \quad \mu \text{ a.e.}$$

Integral of a measurable function

Given a measurable function $f : X \to \mathbb{R}$, set

$$f^+ = \max\{f, 0\} \quad \text{and} \quad f^- = \max\{-f, 0\}.$$

It follows that f^+ and f^- are measurable functions (recall Exercise 2.2),

$$f = f^+ - f^- \quad \text{and} \quad |f| = f^+ + f^-.$$

We say that f is integrable if

$$\int f^+ d\mu < \infty \quad \text{and} \quad \int f^- d\mu < \infty.$$

In case at least one of the two integrals above is finite, we define the integral of f

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu.$$

Given $A \in \mathcal{A}$, we say that a measurable function $f : X \to \mathbb{R}$ is integrable on A if $f\chi_A$ is integrable. We define the integral of f in A

$$\int_A f d\mu = \int f\chi_A d\mu.$$

Invariant measures

Invariant measures

Let $(X, \mathcal{A}, \mu)$ be a measure space.
We say that $f \colon X \to X$ is measurable if

$$f^{-1}(A) \in \mathcal{A}, \qquad \text{for all } A \in \mathcal{A}.$$

We say that a measurable function f preserves μ (or μ is f-invariant) if

$$\mu(f^{-1}(A)) = \mu(A), \quad \forall A \in \mathcal{A}.$$

Defining the push-forward $f_*\mu$ as

$$f_*\mu(A) = \mu(f^{-1}(A)), \quad \forall A \in \mathcal{A}, \tag{8}$$

we have

$$\boxed{\mu \text{ is } f\text{-invariant} \iff f_*\mu = \mu.} \tag{9}$$

Example 3.1 (Doubling map)

Let $f\colon [0,1] \to [0,1]$ be given by

$$f(x) = 2x \pmod 1.$$

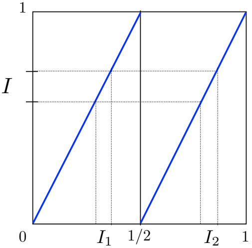
f preserves the Lebesgue measure λ on the Borel sets of $J = [0,1]$.
In fact, let $\mathcal{I}$ be the family of subintervals of J. Given $I \in \mathcal{I}$, we have that $f^{-1}(I)$ is made of two disjoint intervals I_1 and I_2 with

$$\lambda(I_1) = \lambda(I_2) = \frac{1}{2}\lambda(I).$$

Therefore

$$f_*\lambda(I) = \lambda(f^{-1}(I)) = \lambda(I_1 \cup I_2) = \lambda(I_1) + \lambda(I_2) = \frac{1}{2}\lambda(I) + \frac{1}{2}\lambda(I) = \lambda(I).$$

This shows that $\lambda|_{\mathcal{I}} = f_*\lambda|_{\mathcal{I}}$, and so λ and $f_*\lambda$ are both extensions of $\lambda|_{\mathcal{I}} = \ell$ to the Borel sets in $[0,1]$. It follows from Theorem 1.8 that $f_*\lambda = \lambda$.



Example 3.2 (Rotation)

Consider the circle $\mathbb{S}^1 = \{z \in \mathbb{C} \colon |z| = 1\} = \{e^{2\pi i\theta} : \theta \in \mathbb{R}\}$ and, for $\alpha \in \mathbb{R}$, the rotation

$$R_\alpha : \begin{array}{ccc} \mathbb{S}^1 & \longrightarrow & \mathbb{S}^1 \\ e^{2\pi i\theta} & \longmapsto & e^{2\pi i(\theta+\alpha)}. \end{array}$$

Considering the Lebesgue measure λ on the Borel sets of S^1 and $\mathcal{I}$ the family of arcs in $\mathbb{S}^1$, we clearly have

$$R_{\alpha *}\lambda(I) = \lambda(R_\alpha^{-1}(I)) = \lambda(I), \quad \forall I \in \mathcal{I}.$$

This shows that λ and $R_{\alpha *}\lambda$ are both extensions of $\lambda|_{\mathcal{I}}$ to the Borel sets of $\mathbb{S}^1$. It follows from Theorem 1.8 (see also Remark 1.9) that $R_{\alpha *}\lambda = \lambda$.

Exercise 3.3

Let $f \colon [0, 1] \to [0, 1]$ be given by

$$f(x) = \begin{cases} x/2, & \text{if } x \neq 0; \\ 1, & \text{if } x = 0. \end{cases}$$

Show that there is no f-invariant probability measure on the Borel sets of $[0, 1]$.

Hint: Arguing by contradiction, show that the intervals $(1/2^n, 1/2^{n-1}]$ must all have measure zero, for $n \geq 1$. Deduce that the measure of $\{0\}$ is equal to 1. Obtain a contradiction.

Weak* topology

Let $\mathbb{P}(X)$ denote the set of probability measures on the Borel σ-algebra $\mathcal{B}_X$ of a compact metric space X. The weak* topology on $\mathbb{P}(X)$ is characterised as follows: a sequence $(\mu_n)_n$ in $\mathbb{P}(X)$ converges to $\mu \in \mathbb{P}(X)$ if

$$\int \varphi d\mu_n \to \int \varphi d\mu, \quad \text{for all continuous } \varphi \colon X \to \mathbb{R}.$$

Lemma 3.4

$\mathbb{P}(X)$ *is a compact metric space.*

We associate to a measurable map $f \colon X \to X$ a new map

$$f_* \colon \mathbb{P}(X) \to \mathbb{P}(X),$$

assigning to each $\mu \in \mathbb{P}(X)$ the push-forward $f_*\mu \in \mathbb{P}(X)$; recall (8).

Lemma 3.5

f *continuous* $\implies$ f_* *continuous.*

See e.g. [Viana and Oliveira 2016] for a proof of the two lemmas above.

Krylov-Bogolyubov Theorem

Let X be a compact metric space and $f : X \to X$ be a continuous map. Then f has some invariant probability measure.

Proof. Given any $\mu \in \mathbb{P}(X)$ (e.g. a Dirac measure), define the sequence in $\mathbb{P}(X)$,

$$\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} f_*^j \mu.$$

We know by Lemma 3.4 that $\mathbb{P}(X)$ is a compact metric space. Thus, $(\mu_n)_n$ has a subsequence $(\mu_{n_k})_k$ converging to some $\mu_0 \in \mathbb{P}(X)$. We have for all k

$$f_*\mu_{n_k} = \frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^{j+1}\mu = \frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^{j}\mu - \frac{1}{n_k}\mu + \frac{1}{n_k} f_*^{n_k}\mu = \mu_{n_k} - \frac{1}{n_k}\mu + \frac{1}{n_k} f_*^{n_k}\mu.$$

Together with the continuity of f_*, given by Lemma 3.5, this yields

$$f_*\mu_0 = \lim_{k\to\infty} f_*\mu_{n_k} = \lim_{k\to\infty} \left(\mu_{n_k} - \frac{1}{n_k}\mu + \frac{1}{n_k} f_*^{n_k}\mu \right) = \mu_0.$$

Hence, μ_0 is an f-invariant probability measure; recall (9). ∎

Poincaré Recurrence Theorem

Let f preserve a probability measure μ. If A is a measurable set, then for μ almost every $x \in A$, there are infinitely many $n \in \mathbb{N}$ for which $f^n(x) \in A$.

Proof. Set

$$A^r = \{x \in A\colon f^n(x) \in A \text{ for infinitely many } n\text{'s}\},$$

We need to show that

$$\mu(A^r) = \mu(A).$$

Set for each $k \geq 0$,

$$B_k = \left\{x \in A\colon f^k(x) \in A \text{ and } f^{k+n}(x) \notin A, \text{ for all } n \geq 1\right\}.$$

Note that

$$A \setminus A^r = \bigcup_{k \geq 0} B_k.$$

It is enough to show that

$$\boxed{B_k \text{ is measurable and } \mu(B_k) = 0, \text{ for all } k \geq 0.} \tag{10}$$

We have

$$B_k = A \cap f^{-k}(A) \cap f^{-(k+1)}(X \setminus A) \cap f^{-(k+2)}(X \setminus A) \cap \cdots$$

and so B_k is measurable.

For all $k \geq 0$ and $n \geq 1$, we have

$$f^{-n}(B_k) \cap B_k = \emptyset. \tag{11}$$

In fact, if $x \in f^{-n}(B_k)$, then $f^{k+n}(x) \in A$, and so $x \notin B_k$. It follows from (11) that

$$f^{-(n+m)}(B_k) \cap f^{-m}(B_k) = \emptyset$$

for all $n \geq 1$ and $m \geq 0$. Therefore,

$$f^{-n}(B_k) \cap f^{-m}(B_k) = \emptyset$$

if $n \neq m$. It follows that

$$1 \geq \mu\big(\bigcup_{n\geq 1} f^{-n}(B_k)\big) = \sum_{n\geq 1} \mu\big(f^{-n}(B_k)\big).$$

Since f preserves μ, we have $\mu\big(f^{-n}(B_k)\big) = \mu(B_k)$ for all $n \geq 1$, and so

$$\mu(B_k) = 0, \quad \text{for all } k \geq 0.$$

■

Ergodicity

Ergodic measures

Poincaré Recurrence Theorem gives no information on the *asymptotic frequency*

$$\lim_{n\to\infty} \frac{\#\{0 \le j < n \colon f^j(x) \in A\}}{n}. \tag{12}$$

Does this limit exist? Does it depend on x? (almost everywhere...)

The limit clearly depends on x if there is $A \subset X$ such that

$$\begin{cases} \mu(A) > 0 \\ f(A) \subset A \end{cases} \quad \text{and} \quad \begin{cases} \mu(X \setminus A) > 0 \\ f(X \setminus A) \subset X \setminus A \end{cases} \tag{13}$$

In such case

$$x \in A \implies \frac{\#\{0 \le j < n \colon f^j(x) \in A\}}{n} = 1, \quad \forall n \in \mathbb{N}.$$

$$x \in X \setminus A \implies \frac{\#\{0 \le j < n \colon f^j(x) \in A\}}{n} = 0, \quad \forall n \in \mathbb{N}.$$

The *nonexistence* of a set A as in (13) can be translated as

$$f^{-1}(A) = A \quad \implies \quad \mu(A) = 0 \ \text{ or } \ \mu(X \setminus A) = 0.$$

If this condition holds, we say that μ is ergodic for f. Ergodicity is then a necessary condition for the limit in (12) not depend on x. It is also sufficient...

Birkhoff Ergodic Theorem

If $f : X \to X$ preserves a probability measure μ and $\varphi : X \to \mathbb{R}$ is integrable, then there is $\varphi^* : X \to \mathbb{R}$ integrable such that, for μ almost every $x \in X$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x)) = \varphi^*(x).$$

Moreover, if μ is ergodic, then

$$\varphi^*(x) = \int \varphi d\mu,$$

for μ almost every $x \in X$.

See e.g. [Viana and Oliveira 2016] or [Walters 1982] for a proof.
Taking $\varphi = \chi_A$ we get

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n} \chi_A \circ f^j(x) = \lim_{n\to\infty} \frac{\#\{0 \le j < n \colon f^j(x) \in A\}}{n}.$$

Hence, this limit exists for μ almost every $x \in X$ and, if μ is ergodic, it coincides with $\int \chi_A d\mu = \mu(A)$; recall (7).

Circle rotations

Consider the circle

$$\mathbb{S}^1 = \{z \in \mathbb{C} \colon |z| = 1\} = \{e^{2\pi i\theta} : \theta \in \mathbb{R}\}$$

and, for $\alpha \in \mathbb{R}$, the rotation

$$R_\alpha : \begin{array}{ccc} \mathbb{S}^1 & \longrightarrow & \mathbb{S}^1 \\ e^{2\pi i\theta} & \longmapsto & e^{2\pi i(\theta+\alpha)}. \end{array}$$

Theorem 4.1

1. $\alpha \in \mathbb{Q} \implies$ *every orbit is periodic;*
2. $\alpha \in \mathbb{R} \setminus \mathbb{Q} \implies$ *every orbit is dense;*

See [Rechtman 2021] for a proof.

Ergodicity of rotations

We have seen in Example 3.3 that R_α preserves Lebesgue measure λ in $\mathbb{S}^1$.

Exercise 4.2

Show that if $\alpha \in \mathbb{Q}$, then λ is not ergodic for R_α.

Theorem 4.3

λ is ergodic for R_α iff $\alpha \in \mathbb{R} \setminus \mathbb{Q}$.

Proof. The "only if" part corresponds to Exercise 4.2. Assume now that $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and consider a Borel set $A \subset \mathbb{S}^1$ such that

$$R_\alpha^{-1}(A) = A \quad \text{and} \quad \lambda(A) > 0.$$

We need to show that

$$\lambda(A) = 1^\dagger.$$

Fix an arbitrary $0 < \varepsilon < 1$.

---

†Here we assume λ *normalised*, i.e. $\lambda(\mathbb{S}^1) = 1$,

Claim 1. *There is an arc I with $\lambda(I) \leq \varepsilon$ such that*

$$\frac{\lambda(A \cap I)}{\lambda(I)} \geq 1 - \varepsilon$$

Actually, it follows from (5) that there exists a sequence of arcs $I_1, I_2, \dots \subset \mathbb{S}^1$ such that $A \subset \cup_{n \geq 1} I_n$ and

$$\sum_{n \geq 1} \lambda(I_n) \leq \frac{1}{1 - \varepsilon} \lambda(A)$$

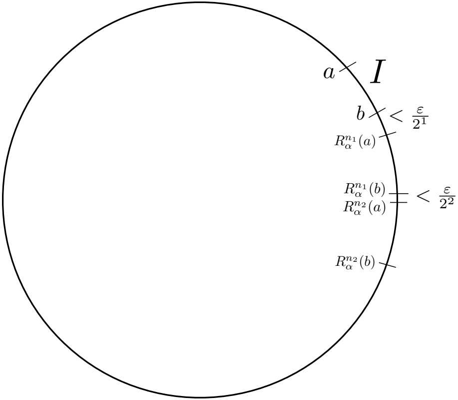
It is no restriction to assume these arcs pairwise disjoint and each of length less than ε. Since

$$\sum_{n \geq 1} \lambda(A \cap I_n) = \lambda(A) \geq (1 - \varepsilon) \sum_{n \geq 1} \lambda(I_n)$$

we must have $\lambda(A \cap I_n) \geq (1 - \varepsilon)\lambda(I_n)$ for some $n \geq 1$. Take $I = I_n$.

Claim 2. *There exist integers $n_1, \ldots, n_k \geq 1$ such that the sets $R_\alpha^{n_1}(I), \ldots, R_\alpha^{n_k}(I)$ are pairwise disjoint and*

$$\lambda(\cup_{i=1}^{k} R_\alpha^{n_i}(I)) \geq 1 - 2\varepsilon.$$



Since the orbits of the endpoints of I are dense in $\mathbb{S}^1$, by Theorem 4.1, these integers may be chosen.

Now, since R_α is invertible with a measurable inverse, R_α preserves λ and A is invariant, from Claim 1 we get for all $1 \leq i \leq k$

$$\begin{gathered} \lambda(A \cap I) \geq (1-\varepsilon)\lambda(I) \\ \Downarrow \\ \lambda(R_\alpha^{n_i}(A \cap I)) \geq (1-\varepsilon)\lambda(R_\alpha^{n_i}(I)) \\ \Downarrow \\ \lambda(A \cap R_\alpha^{n_i}(I)) \geq (1-\varepsilon)\lambda(R_\alpha^{n_i}(I)) \end{gathered} \tag{14}$$

Since the sets $R_\alpha^{n_1}(I), \dots, R_\alpha^{n_k}(I)$ are pairwise disjoint

$$\begin{aligned} \lambda(A) &\overset{(3)}{\geq} \lambda(A \cap \cup_{i=1}^k R_\alpha^{n_i}(I)) = \lambda(\cup_{i=1}^k (A \cap R_\alpha^{n_i}(I))) \\ &= \sum_{i=1}^k \lambda(A \cap R_\alpha^{n_i}(I)) \overset{(14)}{\geq} (1-\varepsilon) \sum_{i=1}^k \lambda(R_\alpha^{n_i}(I)) \\ &= (1-\varepsilon)\lambda(\cup_{i=1}^k R_\alpha^{n_i}(I)) \overset{\text{Claim 2}}{\geq} (1-\varepsilon)(1-2\varepsilon) \end{aligned}$$

Since $0 < \varepsilon < 1$ is arbitrary, we get $\lambda(A) = 1$.

Ergodicity of the doubling map

We have seen in Example 3.1 that the doubling map $f\colon [0,1] \to [0,1]$, given by

$$f(x) = 2x \pmod 1,$$

preserves the Lebesgue measure λ on the Borel sets of $[0,1]$.

Theorem 4.4

λ is ergodic for the doubling map f.

Proof. Let A be a Borel set in $[0,1]$ such that $f^{-1}(A) = A$. We need to show that

$$\lambda(A) = 0 \quad \text{or} \quad \lambda(A) = 1. \tag{15}$$

Consider the dyadic intervals $E_0 = [0,1/2]$ and $E_1 = [1/2,1]$. Since $f^{-1}(A) = A$, then

$$\lambda(A \cap E_0) = \lambda(A \cap E_1).$$

It follows that for $i = 0,1$

$$\lambda(A) = \lambda(A \cap E_0) + \lambda(A \cap E_1) = 2\lambda(A \cap E_i) = \frac{\lambda(A \cap E_i)}{\lambda(E_i)}. \tag{16}$$

Using that $f^{-n}(A) = A$, we similarly prove that for any dyadic interval

$$E_{k,n} = \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right], \quad n \geq 1, \quad 1 \leq k \leq 2^n,$$

we have

$$\lambda(A \cap E_{k,n}) = \lambda(A)\lambda(E_{k,n})$$

If $E = \cup_{k,n} E_{k,n}$ is a disjoint union of dyadic intervals, then

$$\lambda(A \cap E) = \sum_{k,n} \lambda(A \cap E_{k,n}) = \sum_{k,n} \lambda(A)\lambda(E_{k,n}) = \lambda(A)\lambda(E).$$

Now, consider an arbitrary $\varepsilon > 0$.

Exercise 4.5

Given any interval $I \subset [0,1]$, there is a disjoint union of dyadic intervals $\cup_{k,n} E_{k,n}$ such that $I \subset \cup_{k,n} E_{k,n}$ and $\lambda(\cup_{k,n} E_{k,n} \setminus I) < \varepsilon$.

From (5) and Exercise 4.5, there is a disjoint union E of dyadic intervals such that

$$A \subset E \quad \text{and} \quad \lambda(E \setminus A) < 2\varepsilon.$$

Hence

$$0 \leq \lambda(A) - \lambda(A)^2 = \overbrace{\lambda(A) - \lambda(A)\lambda(E)}^{=0} + \lambda(A)\lambda(E) - \lambda(A)^2$$

$$= \lambda(A)\left[\lambda(E) - \lambda(A)\right] \leq \lambda(E) - \lambda(A) \overset{(4)}{=} \lambda(E \setminus A) < 2\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we get $\lambda(A) - \lambda(A)^2 = 0$, and so (15) holds. ■

Normal numbers

A number in $x \in [0,1]$ is said to be normal (in base 2), if the digits 0 and 1 have the same asymptotic frequency in the binary expansion of $x^\dagger$.

Theorem 4.6 (Borel)

Lebesgue almost every $x \in [0,1]$ is normal.

We are going to prove this result using Birkhoff.

With no loss of generality, we may exclude (the countable set of) points in [0,1] having more than one binary expansion.

Exercise 4.7

Every countable set in the real line has zero Lebesgue measure.

† This is a simplified version of the definition, which is more restrictive.

Translating normality

Consider the binary expansion of a number $x \in [0,1]$:

$$x = 0.a_1 a_2 a_3 \dots, \quad a_i \in \{0,1\}.$$

The asymptotic frequency of a digit $d \in \{0,1\}$ is

$$\lim_{n\to\infty} \frac{\#\{1 \le j \le n\colon a_j = d\}}{n}$$

Does this limit exist?

The number x is normal if

$$\lim_{n\to\infty} \frac{\#\{1 \le j \le n\colon a_j = d\}}{n} = \frac{1}{2}, \quad d = 0,1$$

Proof of Theorem 4.6

Considering $x \in [0,1]$ in the binary expansion,

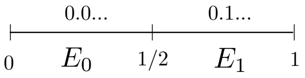
$$x = 0.a_1 a_2 a_3 \cdots \implies 2x = a_1.a_2 a_3 \ldots.$$

Therefore

$$2x \pmod 1 = 0.a_2 a_3 \ldots$$

Letting f be the doubling map, we have for all $j \geq 1$ and $d = 0, 1$

$$a_j = d \quad \Leftrightarrow \quad f^{j-1}(x) \in E_d.$$



Hence

$$\frac{\#\{1 \leq j \leq n \colon a_j = d\}}{n} = \frac{\#\{0 \leq j < n \colon f^j(x) \in E_d\}}{n}.$$

Since the Lebesgue measure λ is ergodic for f, by Theorem 4.4, it follows from Birkhoff Ergodic Theorem that, for λ almost every $x \in [0,1]$,

$$\frac{\#\{0 \leq j < n \colon f^j(x) \in E_d\}}{n} = \frac{1}{n} \sum_{j=0}^{n} \chi_{E_d} \circ f^j(x) \xrightarrow[n \to \infty]{} \int \chi_{E_d} \, d\lambda = \lambda(E_d) = \frac{1}{2}.$$

References

Barra, G. de (2003). *Measure theory and integration*. Horwood Publishing Series. Mathematics and Its Applications. Revised edition of the 1981 original. Horwood Publishing Limited, Chichester, p. 239.

Halmos, P. R. (1950). *Measure Theory*. D. Van Nostrand Company, Inc., New York, N. Y., pp. xi+304.

Rechtman, A. (2021). *Introduction to circle dynamics*. CIMPA School on Recent Advances in Dynamical Systems.

Viana, M. and K. Oliveira (2016). *Foundations of ergodic theory*. Vol. 151. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, pp. xvi+530.

Walters, P. (1982). *An introduction to ergodic theory*. Vol. 79. Graduate Texts in Mathematics. New York: Springer-Verlag, pp. ix+250.