# APPLICATIONS OF THE BIRKHOFF ERGODIC THEOREM

## 1. Transitivity of discrete dynamical systems

Let $\big(X, m\big)$ be a finite measure space, and let $T : X \longrightarrow X$ a measure preserving map. If $T$ also happens to be ergodic, then we recall the following consequence of the Birkhoff ergodic theorem: if

$$A \subset X$$

is measurable, then for almost all $x \in X$ we have the relation

$$\lim_{N\to\infty} \frac{1}{N} \cdot \big|\{j \in [0, N-1],\, T^j(x) \in A\}\big| = \frac{m(A)}{m(X)}. \tag{1.1}$$

We shall use this information to deduce a stronger statement about dense orbits for ergodic mappings in the more specialized context of *measure spaces which are also metric spacea*:

**Proposition 1.1.** *Let $(X, m)$ be a finite measure space and further assume $d : X \times X \to \mathbb{R}_+$ is a metric, such that the metric space*

$$\big(X, d\big)$$

*has a countable basis of open sets. Further, assume that all open subsets are measurable, and that for each non-empty open set $U \subset X$ we have*

$$m(U) > 0.$$

*Then if $T : X \longrightarrow X$ is ergodic, there is a zero measure set $N \subset X$ with the property that each $x \in X \backslash N$ has a dense orbit, i. e.*

$$\overline{\mathcal{O}^+(x)} = X$$

*for each $x \in X \backslash N$.*

**Corollary 1.2.** *Let $T : S^1 \longrightarrow S^1$ the doubling map. Then the set of points $x \in S^1$ which have a dense orbit has measure one.*

*Proof.* (Prop.) We follow the argument in Zehnder. Thus let $\{V_k\}_{k\geq 1}$ a countable basis of open sets for $X$. By (1.1) with $A = V_k$, we can find exceptional zero sets $N_k$, $k \geq 1$, such that if $x \in X \backslash N_k$, we have

$$\lim_{N\to\infty} \frac{1}{N} \cdot \big|\{j \in [0, N-1],\, T^j(x) \in V_k\}\big| = \frac{m(V_k)}{m(X)}.$$

In particular, there are infinitely many $j \geq 0$ such that $T^j(x) \in V_k$.
Setting $N = \bigcup_{k\geq 1} N_k$, we infer that for $x \in X \backslash N$, we have

$$T^j(x) \in V_k$$

for infinitely many $j \geq 0$, and this *for each* $k \geq 1$. But then if

$$V \subset X$$

is an arbitrary open set, there is a $V_k \subset V$, and so we know that

$$T^j x \in V$$

holds for infinitely many $j \geq 0$. This of course implies that

$$\overline{\mathcal{O}^+(x)} = X,\, x \in X \backslash N.$$

By definition we have $m(N) = 0$. $\square$

## 2. Towards mixing; a way to refine ergodicity

At this point we have identified both the irrational circle rotations and the doubling map of the circle as ergodic. In light of their obviously very disparate nature, it appears desirable to introduce an abstract property that tells them apart. The following definition achieves this:

**Definition 2.1.** *Let $(X, m)$ be a finite measure space with $m(X) = 1$, and let $T : X \longrightarrow X$ a measure preserving map. Then we say that $T$ is mixing, provided we have that for every pair of measurable stes $A, B \subset X$, we have*

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

We observe right away that if $T$ is mixing, it is ergodic. For if $T^{-1}A = A$, then we have

$$m(A) = \lim_{n\to\infty} m\big(T^{-n}A \cap A\big) = m(A)^2,$$

which of course implies either $m(A) = 0$ or $m(A) = 1$. However, not every ergodic map is mixing, and so this concept is a refinement of ergodicity in the context of measure spaces of measure $m(X) = 1$. We note right away the simple

**Lemma 2.2.** *A map $T : X \longrightarrow X$ is mixing iff*

$$\lim_{n\to\infty} \int_X (f \circ T^n) \cdot g \, dm = \big(\int_X f \, dm\big) \cdot \big(\int_X g \, dm\big)$$

*for arbitrary $f, g \in L^2(X, dm)$.*

*Proof.* If the preceding identity holds, then it applies to $f = \chi_A, g = \chi_B$, which implies the defining property of mixing.
If $T$ is mixing, then we can approximate $f, g$ by step functions $\widetilde{f} = \sum_i c_i \chi_{A_i}, \widetilde{g} = \sum_j d_j \chi_{B_j}$. Then we have

$$\begin{aligned}\lim_{n\to\infty} \int_X \widetilde{f} \circ T^n \cdot g \, dm &= \sum_{i,j} c_i b_j \lim_{n\to\infty} \int_X \chi_{T^{-n}A_i} \cdot \chi_{B_j} \, dm \\ &= \sum_{i,j} c_i b_j m(A_i) \cdot m(B_j) = \big(\int_X \widetilde{f} \, dm\big) \cdot \big(\int_X \widetilde{g} \, dm\big).\end{aligned}$$

The result for $f, g$ then follows by letting $\widetilde{f} \to f$, $\widetilde{g} \to g$ in $L^2(X, dm)$.

□

**Lemma 2.3.** *(i) No rotation map $T_\alpha : S^1 \longrightarrow S^1$ is mixing. (ii) On the other hand, the doubling map $T : S^1 \longrightarrow S^1$ is mixing.*

*Proof. (i)* Let $\alpha \in \mathbb{R}$ and $T_\alpha : S^1 \longrightarrow S^1$ the corresponding rotation. Then if $\alpha = \frac{p}{q} \in \mathbb{Q}$ set $n_k := q \cdot k$, $k \geq 1$, while if $\alpha \in \mathbb{R} \backslash \mathbb{Q}$, pick a sequence $n_k \in \mathbb{N}$ with the property that $n_k\alpha - \lfloor n_k\alpha \rfloor \longrightarrow 0$ as $k \to \infty$. Then for any measurable set we have

$$m\big(T^{-n_k}A \cap A\big) \longrightarrow m(A \cap A) = m(A)$$

as $k \to \infty$, but for $m(A) \in (0, 1)$ this does not equal $m(A)^2$.

*(ii)* Now assume that $T$ is the doubling map. Identifying $S^1 \cong [0, 1)$, let $A = (a_1, a_2)$ with $a_1 < a_2$. Then

$$T^{-n}A = \bigcup_{k=0}^{2^n-1} (2^{-n}(a_1 + k), 2^{-n}(a_2 + k)) =: \bigcup_{k=0}^{2^n-1} I_k.$$

Then if $B = (b_1, b_2)$ with $b_1 < b_2$ is another interval, we have that

$$\lim_{n\to\infty} 2^{-n} \cdot \big|\{k \in [0, \dots, 2^n - 1], (\frac{k}{2^n}, \frac{k+1}{2^n}) \subset B\}\big| = m(B)$$

But then since each interval $(2^{-n}(a_1+k), 2^{-n}(a_2+k)) \subset (\frac{k}{2^n}, \frac{k+1}{2^n})$ has measure $2^{-n} \cdot m(A)$, we get

$$\lim_{n\to\infty} m\big(T^{-n}A \cap B\big) = \lim_{n\to\infty} m\Big(\bigcup_{k=0}^{2^n-1} (2^{-n}(a_1+k), 2^{-n}(a_2+k)) \cap B\Big)$$
$$= m(A) \cdot m(B).$$

If $A, B$ are arbitrary open sets, fill them up with countably many open intervals $A_i, B_j$, respectively, write (with convergence in the $L^2$-sense, say)

$$\chi_A = \sum_i \chi_{A_i},\ \chi_B = \sum_j \chi_{B_j},$$

and use the preceding observation for each $A_i, B_j$.
If $A$ is closed and $B$ is open, use

$$\lim_{n\to\infty} m\big(T^{-n}A \cap B\big) = m(B) - m\big(T^{-n}A^c \cap B\big) = m(B) - m(A^c) \cdot m(B) = m(A) \cdot m(B),$$

and similarly if $A$ is open but $B$ is closed. The result for both sets closed then also easily follows.
If $A, B$ are arbitrary Lebesgue measurable subsets of $S^1$, we approximate them from within by closed sets and from without by open sets and use the preceding. □

While general ergodic maps fail to be mixing, we have a weaker and similar property, which replaces the strong limiting relation in Definition 2.1 by an *averaged one*:

**Proposition 2.4.** *Let $\big(X, m\big)$ be a probability measure space, i. e. $m(X) = 1$, and let $T : X \longrightarrow X$ be measure preserving. Then $T$ is ergodic iff*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} m\big(T^{-j}A \cap B\big) = m(A) \cdot m(B). \tag{2.1}$$

*for any pair of measurable sets $A, B$.*

*Proof.* First assume (2.1). Then if $A$ is invariant under $T$, i. e. $A = T^{-1}A$, we conclude that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} m\big(T^{-j}A \cap A\big) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} m\big(A \cap A\big) = m(A).$$

On the other hand, this must equal $m(A)^2$. But

$$m(A) = m(A)^2$$

implies either $m(A) = 0$ or $m(A) = 1$. It follows that $T$ is ergodic.

Next assume that $T$ is ergodic. Then by the Birkhoff ergodic theorem, we now that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} \chi_{T^{-j}A}(x) = m(A)$$

for almost every $x \in X$. It follows that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} \chi_{T^{-j}A}(x) \cdot \chi_B(x) = m(A) \cdot \chi_B(x)$$

for almost every $x \in X$. Since the functions

$$f_N(x) := \frac{1}{N} \sum_{j=0}^{N-1} \chi_{T^{-j}A}(x) \cdot \chi_B(x)$$

are uniformly bounded in absolute value by 1, the Lebesgue dominated convergence theorem implies that

$$\int_X f_N\, dm \longrightarrow m(A) \cdot \int_X \chi_B(x)\, dm = m(A) \cdot m(B).$$

But we have

$$\int_X f_N\, dm = \frac{1}{N} \sum_{j=0}^{N-1} m\big(T^{-j} A \cap B\big),$$

and (2.1) follows. $\square$