A LITTLE TOPOLOGICAL COUNTERPART OF BIRKHOFF'S ERGODIC THEOREM

REINHARD WINKLER

Dedicated to the memory of Professor Edmund Hlawka (1916-2009)

Abstract. For a compact metric space X and a continuous transformation $T : X \to X$ with at least one transitive and recurrent orbit, there is a set $M_0(T)$ of T-invariant probability measures on X such that for a comeager set of starting points the set of limit measures is exactly $M_0(T)$.

1. Introduction

For a compact metric space X and $T : X \to X$ a point $x \in X$ is called transitive resp. recurrent if its T-orbit $(T^n x)_{n\in\mathbb{N}}$ is dense resp. meets every neighbourhood of x infinitely many times. Furthermore $\mathcal{M}(X,T)$ denotes the set of those $\mu \in \mathcal{M}(X)$ (the compact metrizable space of all Borel probability measures) which are T-invariant, and $M(T,x)$ the set of all limit measures of the sequence $\mathbf{x} = (x_n)_{n\in\mathbb{N}}$ with $x_n = T^n x$. By definition, a limit measure of $\mathbf{x}$ is an accumulation point of the measures $\mu_{\mathbf{x},n} = \frac{1}{n}\sum_{n=0}^{n-1}\delta_{x_n} \in \mathcal{M}(X)$, $n = 1,2,3,\ldots$, δ_{x_n} denoting the point measures concentrated in $x_n \in X$. We will prove:

Theorem *Let X be a compact metric space, $T : X \to X$ a continuous transformation, $x_0 \in X$ a transitive and recurrent point and $M_0(T)$ the union of all $M(T,x)$ with transitive x. Then $M(T,x) = M_0(T)$ for most $x \in X$, i.e. for all $x \in X \setminus E$ where the exceptional set E is meager (of first Baire category).*

Note that for infinite X the assumptions of the Theorem imply that X is perfect (i.e. has no isolated points), hence uncountable, and that every dense orbit is recurrent.

For an ergodic measure $\mu \in \mathcal{M}(X,T)$, Birkhoff's ergodic theorem yields that for μ-almost all $x \in X$ the sequence $\mathbf{x} = (T^n x)_{n\in\mathbb{N}}$ is uniformly distributed with respect to μ, i.e. $M(T,x) = \{\mu\}$. Thus the Theorem above

2000 *Mathematics Subject Classification.* Primary 37B05; Secondary 37A05.

Key words and phrases. Birkhoff's ergodic theorem, Baire category, topological dynamics, distribution of sequences.

The research for this paper has been supported by the Austrian Science Foundation FWF through Project no. S9612-N23.

can be considered as a topological counterpart in the sense of Baire categories (see also the classical textbook [O 80]) where the singleton $\{\mu\}$ has to be replaced by the set $M_0(T)$. The proof of the Theorem is the content of Section 2. Section 3 is a short discussion including examples where $M_0(T) = \mathcal{M}(X,T)$, i.e. where most points have maximal oscillation in the sense of (21.17) in [DGS 76]. Section 4 shows that $M_0(T) \neq \mathcal{M}(X,T)$ is possible as well.

Related properties of the topologically typical distribution behaviour of orbits have already been observed in [D 53], for arbitrary sequences by Prof. Hlawka in his seminal paper [H 56]. For more recent investigations cf. [Wi 97], [GSW 00], [GSW 07] and [TZ 10].

2. Proof of the Theorem

Let in this Section X, T and x_0 be as in the assumptions of the Theorem. For any sequence of mappings $\phi_n : X \to Y$ (Y metric space) and $y \in Y$ the set $X((\phi_n), y)$ of all $x \in X$ such that y is an accumulation point of $(\phi_n(x))_{n\in\mathbb{N}}$ can be written as

$$X((\phi_n), y) = \bigcap_{N,k\in\mathbb{N}} \bigcup_{n\geq N} \phi_n^{-1}(B(y, \frac{1}{k+1})).$$

($B(y,r)$ denotes the open ball with center y and radius r.) For any topological space X and continuous ϕ_n the sets $\phi_n^{-1}(B(y, \frac{1}{k+1}))$ are open. This shows that for continuous $\phi_n : X \to Y$, $X((\phi_n), y)$ is a G_δ-set and that $X((\phi_n), y)$ is residual if and only if for all $N, k \in \mathbb{N}$ the set

$$\bigcup_{n\geq N} \phi_n^{-1} \left(B(y, \frac{1}{k+1}) \right)$$

is dense in X. Take $Y = \mathcal{M}(X)$, $\phi_n : x \mapsto \mu_{(T^k x)_{k\in\mathbb{N}}, n} = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{T^k x}$ and $y = \mu \in M(T, x_0)$. Since then $\mu \in M(T, T^n(x_0))$ for all $n \in \mathbb{N}$, $\bigcup_{n\geq N} \phi_n^{-1}(B(\mu, \frac{1}{k+1}))$ is dense in X for all $N, k \in \mathbb{N}$ (the balls taken w.r.t. any metric for the topology on $\mathcal{M}(X)$). Hence:

Proposition 1. *If $\mu \in M(T, x_0)$ then $\mu \in M(T, x)$ for most $x \in X$.*

As a subset of the compact metric space $\mathcal{M}(X)$, $M_0(T)$ contains a countable dense subset $\{\mu_n : n \in \mathbb{N}\}$. Let X_{μ_n} denote the set of all $x \in X$ with $\mu_n \in M(T, x)$. By Proposition 1 each X_{μ_n}, $n \in \mathbb{N}$, is residual in X. Hence also the countable intersection $X_1 = \bigcap_{n\in\mathbb{N}} X_{\mu_n}$ is residual. For all $x \in X_1$ and $n \in \mathbb{N}$ we have $\mu_n \in M(T, x)$. Since $M(T, x)$ is closed this implies:

Proposition 2. *$M_0(T) \subseteq \overline{M(T)} \subseteq M(T, x)$ for most $x \in X$.*

It is well-known that in transitive systems most orbits are dense (cf. for instance [DGS 76], 6.11). By the definition of $M_0(T)$ as the union of the $M(T, x)$ with transitive x and since $M(T, x) \subseteq \mathcal{M}(X, T)$ (cf. for instance [GSW 07], Lemma 2.17 (1)) this implies the converse inclusion $M(T, x) \subseteq M_0(T)$ for most $x \in X$, proving the Theorem.

3. Discussion

Trivial examples for the Theorem are uniquely ergodic transformations where $M_0(T) = \mathcal{M}(X,T)$ is a singleton. A less trivial example with $M_0(T) = \mathcal{M}(X,T)$ is the full shift, i.e. $X = A^{\mathbb{N}}$, the set of all sequences over a finite alphabet A, and $T = \sigma : x = (a_n)_{n\in\mathbb{N}} \mapsto (a_{n+1})_{n\in\mathbb{N}}$ (cf. [DGS 76] chapter 21, in particular 21.18). The full shift also shows that the residual set of all $x \in X$ with $M(T,x) = M_0(T)$, in general, does not coincide with the set of all transitive and recurrent x: Take any sequence x which contains all finite words, separated by sufficiently long blocks of 0's. Then $M(T,x) = \{\delta_{0^\infty}\} \neq \mathcal{M}(X,T) = M_0(T)$ while x is transitive and recurrent.

It is clear that the transitivity assumption in the Theorem cannot be omitted. (Otherwise we might have disjoint open sets with disjoint T-orbits such that the Theorem must fail. Most trivial example: T the identity on X where X contains at least two points.) However, some kind of generalization of the Theorem to the intransitive case is possible. But since this requires a much broader framework I do not go into this direction here.

Similarly to the transitivity assumption, the Theorem does not hold in general if we omit the recurrence condition. Consider $X = X_1 \cup \{x_0\}$, the compact space $X_1 = \{0,1\}^{\mathbb{N}}$ of all binary sequences plus an isolated point x_0. Let $T = T_0 \cup T_1$ with the shift $T_1 = \sigma : (a_n)_{n\in\mathbb{N}} \to (a_{n+1})_{n\in\mathbb{N}}$ on X_1 and $T_0 : x_0 \mapsto x_1$ with some $x_1 \in X_1$. If x_1 contains each binary word of finite length, then x_0 is a transitive point (but not recurrent). As already mentioned above there is a $M_0(T_1)$, namely the set of all invariant measures on X_1. If x_1 is suitably chosen (see above), then $M(T,x_0) = M(T,x_1)$ does not contain all invariant measures. $\{x_0\}$, as an open set, is not meager. Thus, provided $M(T,x_1) \neq \mathcal{M}(X,T_1)$, there is no Baire-typical $M_0(T)$ for T considered as a transformation on the whole space $X = X_1 \cup \{x_0\}$.

4. An example with $M_0(T) \neq \mathcal{M}(X,T)$

Our example is the subshift generated by the binary sequence

$$x_0 = \mathbf{a} = (a_n)_{n\in\mathbb{N}} = a_0a_1a_2\ldots = 0^1 1^1 a_0 0^2 1^2 a_0 a_1 0^3 1^3 a_0 a_1 a_2 \ldots .$$

Let X be the orbit closure of x_0 under σ and T the restriction of σ to X. Note that x_0 is defined in such a way that each finite initial word of x_0 occurs infinitely many times in x_0. Hence x_0 is recurrent. Since X is the orbit closure, x_0 is also transitive. So the Theorem applies and $M(T,x) = M_0(T)$ for most $x \in X$. Note that $1^\infty = 111\ldots \in X$ since x_0 contains all 1-blocks 1^n, $n \in \mathbb{N}$. Furthermore the point measure δ_{1^∞} is shift invariant, hence $\delta_{1^\infty} \in \mathcal{M}(X,T)$. So the proof of $M_0(T) \neq \mathcal{M}(X,T)$ will be complete as soon as we have shown $\delta_{1^\infty} \notin M_0(T)$.

The definition of x_0 induces a partition of $\mathbb{N}$ into subintervals $I_k = I_k^{(0)} \cup I_k^{(1)} \cup I_k^{(r)}$, $k = 1,2,3,\ldots$, in such a way that $I_1 < I_2 < \ldots$ elementwise, $I_k^{(0)} < I_k^{(1)} < I_k^{(r)}$ elementwise and $|I_k^{(0)}| = |I_k^{(1)}| = |I_k^{(r)}| = k$. Clearly this determines the partition uniquely. Note that $a_n = 0$ for all $n \in I_k^{(0)}$, $a_n = 1$ for all $n \in I_k^{(1)}$ and $a_n = a_j$ if $n = m + j \in I_k^{(r)} = \{m, m+1, \ldots, m+k-1\}$ with $j < k$.

Let W_l be the set of all words $w = (a'_n, a'_{n+1}, \ldots, a'_{n+l-1})$ of length l occuring in x_0 and $W = \bigcup_{l \in \mathbb{N}} W_l$. Let us write $\mu(0|w)$ and $\mu(1|w) = 1 - \mu(0|w)$ for the relative frequency of 0's resp. 1's in a nonempty word w. Formally: For $w = (a'_0, \ldots, a'_{l-1})$, $\mu(i|w) = \frac{1}{l}|\{n : \ 0 \le n \le l-1, \ a'_n = i\}| \in [0, 1]$, $i \in \{0, 1\}$, $l = 1, 2, \ldots$.

Proposition 3. *In every initial word $w = (a_0, a_1, \ldots, a_{l-1})$ of x_0, $l \in \mathbb{N}$, we have $\mu(0|w) \ge \frac{1}{2} \ge \mu(1|w)$.*

Proof of Proposition 3: One sees immediately that the assertion holds for $l = 0, 1, 2, 3$, hence it holds for the word induced by x_0 on I_1, hence on I_2 etc., hence on all concatenations of these finite words. □

Proposition 4. *If $w = 10w' \in W$, then $\mu(0|0w') \ge \frac{1}{2} \ge \mu(1|0w')$.*

Proof of Proposition 4: Let us, by contradiction, suppose that the claim fails. Then there is a minimal $n_0 \in \mathbb{N}$ and an $l \ge 2$ such that

$$w = (a_{n_0}, a_{n_0+1}, \ldots, a_{n_0+l-1}) = 10w'$$

is a counterexample to the proposition. Since $n_0 \notin I_k^{(0)}$ for any $k \in \mathbb{N}$ it suffices to distinguish two cases for n_0.

Case 1, $n_0 \in I_{k_0}^{(1)}$: In the first subcase ($l - 1 \le k_0$) we have $w = 10w' = 10a_1 \ldots a_{l-1} = 1a_0a_1 \ldots a_{l-1}$, contradicting Proposition 3. In the other subcase ($l - 1 > k_0$) we have $w = 1a_0 \ldots a_{k_0-1}b_{k_0+1}b_{k_0+2} \ldots b_{k_1-1}b_{k_1}$ where the b_j are the finite words induced by x_0 on I_j, b_{k_1} being only an initial segment. This again contradicts Proposition 3.

Case 2, $n_0 \in I_{k_0}^{(r)}$: If (first subcase) $n_0 + l - 1 \in I_{k_0}^{(r)}$ then w occurs already as $w = a_{m_0} \ldots a_{m_0+l-1}$ with $m_0 < n_0$, contradicting the minimal choice of n_0. Otherwise (second subcase) we can write w as a concatenation $w = b_{k_0}b_{k_0+1} \ldots b_{k_1-1}b_{k_1}$ of words b_i with $k_1 > k_0$ in such a way that $b_{k_0} = 10w''$ comes from an end word of $I_{k_0}^{(r)}$, the b_{k_0+j} with $0 < j < k_1 - k_0$ come from the corresponding I_{k_0+j} and b_{k_1} is an initial word. By the minimality of n_0 the claim of the lemma holds for $0w''$ instead of $0w'$. For the tail $b_{k_0+1} \ldots b_{k_1-1}b_{k_1}$ of w, Proposition 3 implies that there are at least as many 0's as 1's, contradiction. □

Proposition 5. *Every $x = (a'_n)_{n \in \mathbb{N}} \in X$ is either of the form $x = w1^\infty$ (type 1) with a finite initial word w or the upper density $\overline{\mu}(1|x)$ of the set $\{n : \ a'_n = 1\}$ is at most $\frac{1}{2}$ (type 0). (Here $\overline{\mu}(1|x)$ denotes the upper limit of $\mu(1|w_n)$ for $n \to \infty$ where w_n is the n-th initial word of x.) In particular, $\delta_{1^\infty} \notin M(T, x)$ for every x of type 0.*

Proof of Proposition 5: If $x \in X$ is not of type 1 then x contains infinitely many 0's. For $x = 0^\infty$ the claim is obvious, otherwise there is a finite word w_0 and an infinite sequence x' such that $x = w_0 10x'$. For all finite initial words w' of x' we have $10w' \in W$. In combination with Proposition 4 this implies that $\overline{\mu}(1|x) \leq \frac{1}{2}$ for all x of type 0. Thus for such x and any $\mu \in M(T, x)$ we have $\mu(X_1) \leq \frac{1}{2}$, X_1 denoting the set of all sequences in X starting with the digit 1. Since $\delta_{1^\infty}(X_1) = 1$ this implies $\delta_{1^\infty} \notin M(T, x)$ for all x of type 0. $\Box$

Proof of $M_0(T) \neq \mathcal{M}(X, T)$: Since X has no isolated points, each of the points of type 1 (in the sense of Proposition 5), as a singleton, forms a nowhere dense subset. Since there are not more than countably many points of type 1, most points are of type 0. By Proposition 5 we have $\delta_{1^\infty} \notin M(T, x)$ for every x of type 0, hence $\delta_{1^\infty} \notin M_0(T)$. $\Box$

References

[DGS 76] M. Denker, C. Grillenberger and K. Sigmund, *Ergodic Theory on Compact Spaces*, Lecture Notes in Mathematics, Volume 527, Springer (1976).

[D 53] Y. N. Dowker, *The mean and transitive points of homeomorphisms*, Ann. of Math. (2) 58 (1953), 123–133.

[GSW 00] M. Goldstern, J. Schmeling and R. Winkler, *Metric, fractal dimensional and Baire results on the distribution of subsequences*, Math. Nachr. 219 (2000), 97–108.

[GSW 07] M. Goldstern, J. Schmeling and R. Winkler, *Further Baire results on the distribution of subsequences*, Uniform Distribution Theory 2 (2007), no.1, 127–149.

[H 56] E. Hlawka, *Folgen auf kompakten Räumen*, Abh. Semin. Univ. Hamb. 20 (1956), 223-241.

[O 80] J. C. Oxtoby, *Measure and Category*, Graduate Texts in Mathematics, Springer (1980).

[TZ 10] R. F. Tichy and Martin Zeiner, *Baire results of multisequences*, Uniform Distribution Theory (the present volume of Uniform Distribution Theory).

[Wi 97] R. Winkler, *On the distribution behaviour of sequences*, Math. Nachr. 186 (1997), 303–312.