METRIC, FRACTAL DIMENSIONAL AND BAIRE RESULTS ON THE DISTRIBUTION OF SEQUENCES AND SUBSEQUENCES

MARTIN GOLDSTERN, JÖRG SCHMELING, AND REINHARD WINKLER

Abstract. Let X be a locally compact metric space. One important object connected with the distribution behavior of an arbitrary sequence $\mathbf{x}$ on X is the set $M(\mathbf{x})$ of limit measures of $\mathbf{x}$. It is defined as the set of accumulation points of the sequence of the discrete measures induced by $\mathbf{x}$.

Using binary representation of reals one gets a natural bijective correspondence between infinite subsets of the set $\mathbb{N}$ of positive integers and numbers in the unit interval $I = [0, 1)$. Hence to each sequence $\mathbf{x} = (x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}}$ and every $a \in I$ there corresponds a subsequence denoted by $a\mathbf{x}$.

We investigate the set $M(a\mathbf{x})$ for given $\mathbf{x}$ with emphasis on the behavior for "typical" a in the sense of Baire category, Lebesgue measure and Hausdorff dimension. We also treat the special case of the sequence $(n\alpha)$ mod 1 and its subsequences.

1. Introduction

Notations. For the whole paper we fix the following abbreviations.

Let $\mathbb{N}$ denote the set of positive integers $1, 2, \ldots$ and let $\mathbb{R}/\mathbb{Z}$ be the one dimensional torus group, identified with the half open unit interval $I = [0, 1)$ in the natural way. That is, we identify the remainder class $x + \mathbb{Z}$ with the real number $x - [x]$ (where $[x]$ denotes the the largest integer $k \in \mathbb{Z}$ with $k \le x$). So the canonical epimorphism $\mathbb{R} \to \mathbb{R}/\mathbb{Z}$ can be identified with the map $x \mapsto x - [x]$.

For any set X we write $X^{\mathbb{N}}$ for the set of all sequences $\mathbf{x} = (x_n)_{n\in\mathbb{N}}$, $x_n \in X$. Sequences (both infinite and finite) are always denoted by boldface letters. If $\mathbf{x} = (x_n)_{n\in\mathbb{N}}$ is an infinite sequence, we write $\mathbf{x}\restriction N$ for the finite sequence $(x_1, \ldots, x_N)$.

For a finite sequence $(\varepsilon_1, \ldots, \varepsilon_n)$ of elements of X we write $[\varepsilon_1, \ldots, \varepsilon_n]$ for the "cylinder set"

$$\{(x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}} : x_1 = \varepsilon_1, \ldots, x_n = \varepsilon_n\}.$$

If X is discrete (in particular if $X = \{0, 1\}$), then we may write $C_n(\mathbf{x})$ for $[\varepsilon_1, \ldots, \varepsilon_n]$. In this case the family $(C_n(\mathbf{x}))_{n\in\mathbb{N}}$ is a local basis at $\mathbf{x}$ in the product topology on $X^{\mathbb{N}}$.

Each number $a \in I$ has a representation $a = \sum_{n\in\mathbb{N}} a_n 2^{-n}$ with $a_n \in \{0, 1\}$. This representation is unique if one requires that the set $S_a = \{n \in \mathbb{N} \mid a_n = 1\}$ is infinite. (Note that to $0 \in I$, by identification, there corresponds the set $S_0 = S_1 = \mathbb{N}$.) Thus the correspondence $a \mapsto S_a$ is a bijection between I and the system of all infinite

---

Date: 1998-03-04.

1991 *Mathematics Subject Classification.* Primary 11K55; Secondary 28A33.

The first author was supported by an Erwin Schrödinger Fellowship from the Austrian Science Foundation (FWF).

The second author was partially supported by the Leopoldina Forschungspreis.

subsets of $\mathbb{N}$ or, equivalently, the set of all strictly increasing infinite sequences of integers. If $S_a = \{k_1 < k_2 < \dots\}$ and $\mathbf{x} = (x_n)_{n\in\mathbb{N}}$, then we write $a\mathbf{x}$ for the sequence $(x_{k_n})_{n\in\mathbb{N}} = (x_i)_{i\in S_a}$. If $\mathbf{x} = (x_1, \dots, x_n)$ is a finite sequence, then $a\mathbf{x}$ will denote the subsequence $(x_i)_{i\in\{1,\dots,n\}\cap S_a}$.

We usually use the Euclidean metric and topology on $[0, 1)$, in particular when we consider measure and Hausdorff dimension. For arguments dealing with Baire category, however, we identify $[0, 1)$ with a subspace of $\mathbb{N}^{\mathbb{N}}$ and use the Tychonoff topology, which we call "the 0-dimensional topology." Note that the notions "meager" and "residual" are the same in the 0-dimensional topology and the Euclidean topology.

Each $p \in I$ induces the p-Bernoulli measure λ_p on I which is the Borel measure uniquely defined by the requirement $\lambda_p(\{a \mid a_n = 1\}) = p$ for all $n \in \mathbb{N}$. If $p \in \{0, 1\}$, λ_p is concentrated on $0 = 1 \in I$. Hence the case $0 < p < 1$ is the interesting one and will be assumed for the rest of the paper.

Let X be a locally compact Hausdorff space with countable topological base. Thus there is a complete metric d_X inducing the topology.

With $A(\mathbf{x})$ we denote the set of accumulation points of $\mathbf{x}$. This is the set of all $y \in X$ such that in each neighborhood O of y the relation $x_n \in O$ holds for infinitely many $n \in \mathbb{N}$.

Let now $\mathcal{M}(X)$ denote the set of all nonnegative Borel measures μ on X with $\mu(X) = 1$ if X is compact and $\mu(X) \le 1$ if X is not compact. Let $\mathcal{K}(X)$ denote the space of all continuous $f : X \to \mathbb{R}$ with compact support

$$\operatorname{supp}(f) = \overline{\{x \in X \mid f(x) \neq 0\}}.$$

We consider the weak topology on $\mathcal{M}(X)$ which is the weakest topology such that all functionals $\mu \mapsto \int_X f d\mu$, $f \in \mathcal{K}(X)$, are continuous. $\mathcal{M}(X)$ is a compact space with countable topological base, hence also induced by a metric $d_{\mathcal{M}(X)}$. In most cases it will cause no confusion to write d for both d_X and $d_{\mathcal{M}(X)}$. We will need the fact that $d_{\mathcal{M}(X)}$ can be chosen to be compatible with convex combinations, i.e.,

$$d_{\mathcal{M}(X)}(\mu_0, (1-t)\mu_0 + t\mu_1) = t d_{\mathcal{M}(X)}(\mu_0, \mu_1)$$

for $0 \le t \le 1$.

(One possibility of defining d is the following: Let $f_n : \mathbb{R} \to [0, 1]$, $n \in \mathbb{N}$, be a countable family spanning a dense subspace of $\mathcal{K}(X)$. Then the function $d : \mathcal{M}(X)^2 \to \mathbb{R}$, defined by

$$d(\mu_1, \mu_2) = \sum_{n\in\mathbb{N}} 2^{-n} \Big| \int f_n \, d\mu_1 - \int f_n \, d\mu_2 \Big|$$

satisfies the above requirements.)

Each finite sequence $\mathbf{x} = (x_1, \dots, x_N)$ induces a probability measure $\mu_{\mathbf{x}} = \frac{1}{N}\sum_{n=1}^{N} \delta_{x_n}$, where δ_x is the Dirac measure concentrate at x. For an infinite sequence $\mathbf{x} = (x_n)_{n\in\mathbb{N}}$ we let

$$\mu_{N,\mathbf{x}} = \mu_{\mathbf{x}\restriction N}, \qquad \mu_{N,\mathbf{x}}(B) = \frac{1}{N}\#\{i \le N : x_i \in B\}$$

($\#A$ denotes the cardinality of a set A.)

With $M(\mathbf{x})$ we denote the set of all accumulation points of the sequence $(\mu_{N,\mathbf{x}})_{N\in\mathbb{N}}$ in $\mathcal{M}(X)$.

It follows from [Si] (for the compact case) or [W2] that for an arbitrary subset $M \subseteq \mathcal{M}(X)$ the following holds: There is a sequence $\mathbf{x}$ with $M = M(\mathbf{x})$ if and only if M is nonempty, closed and connected. $\mathbf{x}$ is called uniformly distributed w.r.t. μ (μ-u.d.) if $M(\mathbf{x}) = \{\mu\}$. If X is a compact group and μ is the Haar measure on X, we call $\mathbf{x}$ simply u.d. In the other extreme case $M(\mathbf{x}) = \mathcal{M}(X)$, the sequence $\mathbf{x}$ is called maldistributed (md.). Thus we use the notation from [W2] which is a slight modification of the one in [My].

If $A \subseteq X$ is a closed subset of X, then the set $\mathcal{M}(A)$ can be considered as a subspace of $\mathcal{M}(X)$ in an obvious way: Identify every $\mu \in \mathcal{M}(A)$ with its extension $\bar{\mu} \in \mathcal{M}(X)$, defined by $\bar{\mu}(B) = \mu(B \cap A)$ for arbitrary Borel sets $B \subseteq X$.

Recall that a subset of a topological space is called nowhere dense if its topological closure has empty interior. A set is of first category or meager if it is the union of at most countably many nowhere dense sets. All sets which are not meager are said to be of second category. A set is called co-meager or residual if its complement is meager. Complete metric spaces (as $I = \mathbb{R}/\mathbb{Z}$, X and $\mathcal{M}(X)$) are Baire spaces, i.e., nonempty open subsets are of second category.

Content of the paper. In section 2 we investigate the set $M(a\mathbf{x})$ for given $\mathbf{x}$ with emphasis on the behavior for "typical" a in the sense of Baire category, Lebesgue measure and Hausdorff dimension. In section 3 we consider sequences depending on an additional parameter. In section 4 we focus on the special case of $n\alpha$-sequences. We give Baire results on the typical distribution of $(k_n\alpha)_{n\in\mathbb{N}}$ depending on the growth rate of $(k_n)_{n\in\mathbb{N}}$. For special sequences $(k_n)_{n\in\mathbb{N}}$ cf. also [Be].

The reader who is interested in connections to ergodic theory may want to consult the survey paper [R-W]. The standard textbook on uniform distribution is [K-N]. Many interesting results and an up-to-date bibliography can found in [D-T].

2. Subsequences

We fix the general assumption that X be a locally compact and second countable (hence complete metric) space. Given a sequence $\mathbf{x}$ on X we look, in dependence on a, at its subsequences $a\mathbf{x}$ and the corresponding sets $M(a\mathbf{x})$. The natural questions on this sets are:

1. Which a guarantee $M(a\mathbf{x}) = M(\mathbf{x})$ for arbitrary $\mathbf{x}$? (density 1, theorem 1.)
2. What can $M(a\mathbf{x})$ look like? (Anything reasonable, theorem 2.)
3. What is the typical situation in the Baire sense? (Maldistribution, theorem 3.)
4. What is the typical situation in the Lebesgue sense? ($M(a\mathbf{x}) = M(x)$, theorem 5.)
5. What is the (Hausdorff) dimension of all a with $M(a\mathbf{x}) \neq M(\mathbf{x})$? (Dimension 1, theorem 14.)

For theorem 1 we recall the notion of density. For $S \subseteq \mathbb{N}$ and $n \in \mathbb{N}$ one considers the numbers $d(n,S) = \frac{1}{n}\#S \cap (0,n]$. Since $0 \leq d(n,S) \leq 1$ the sequence $(d(n,S))_{n\in\mathbb{N}}$ has an upper and lower limit, the "upper density $\overline{d}(S)$" and the "lower density $\underline{d}(S)$." If they coincide, their common value is called the density of S, $d(S)$.

Theorem 1.

a) *If S_a has density 1 then $M(a\mathbf{x}) = M(\mathbf{x})$.*
b) *If $\overline{d}(S) < 1$, and $\mathbf{y}$ is an arbitrary sequence on X, $\#X \geq 2$, then there exists a sequence $\mathbf{x}$ such that $a\mathbf{x} = \mathbf{y}$ and $M(a\mathbf{x}) = M(\mathbf{y}) \neq M(\mathbf{x})$.*

Proof. a): Use the fact that $d(n, S_a) \geq 1-\varepsilon$ implies $|\mu_{N,\mathbf{x}}(B) - \mu_{N,a\mathbf{x}}(B)| \leq 2\varepsilon$ for any set B.
b): First case: $M(\mathbf{y}) = \mathcal{M}(X)$, i.e., $\mathbf{y}$ is maldistributed. Take an arbitrary $x \in X$ and define $x_n = x$ for all $n \notin S_a$ and $x_{a_n} = y_n$ for $a_n \in S_a$ such that $a\mathbf{x} = \mathbf{y}$. Since $\#X \geq 2$ there exists a $\mu \in \mathcal{M}(X)$ with $\mu(\{x\}) = 0$. Note that

$$\liminf \mu_{N,\mathbf{x}}(\{x\}) \geq 1 - \overline{d} > 0.$$

Thus $\mu \notin M(\mathbf{x})$ proving $M(\mathbf{x}) \neq \mathcal{M}(X) = M(\mathbf{y}) = M(a\mathbf{x})$.

Second case: If there is a $\mu_0 \in \mathcal{M}(X) \setminus M(\mathbf{y})$ consider $\mathbb{N} \setminus S_a = \{b_1 < b_2 < \dots\}$. Let $\mathbf{z}$ be μ_0-u.d. and define $x_{b_n} = z_n$ and $x_{a_n} = y_n$. $M(a\mathbf{x})$ is a nonempty compact set. Hence there is a $\mu_1 \in M(a\mathbf{x})$ such that $\varepsilon = d(\mu_0, \mu_1) > 0$ takes a the minimal value. It follows that there is a convex combination $\mu_t = t\mu_0 + (1-t)\mu_1$, $0 < t < 1$, such that $\mu_t \in M(\mathbf{x})$. Here we used $\overline{d} < 1$. Since $d(\mu_0, \mu_t) < d(\mu_0, \mu_1)$ we conclude $\mu_t \in M(\mathbf{x}) \setminus M(a\mathbf{x})$. □

If lower and upper density $\underline{d}$ and $\overline{d}$ of S_a satisfy $\underline{d} < \overline{d} = 1$ general assertions as in theorem 1 fail to be true. To illustrate the situation look at the following examples where we use the notations of theorem 1:

If $\mathbf{z}$ is μ_0-u.d. and $\mathbf{y}$ is μ_1-u.d. $\mu_0 \neq \mu_1$, then in the second case of part b) one gets $M(\mathbf{x}) = \{\mu_t \mid t \in [\underline{d}, \overline{d}]\}$ with $\varepsilon = \underline{d} > 0$. Thus $M(\mathbf{x}) = M(a\mathbf{z})$. On the other hand, for any a with $\overline{d}(S_a) = 1$, it is possible to construct a maldistributed sequence $\mathbf{y}$ such that $a\mathbf{x} = \mathbf{y}$ implies $M\mathbf{x}) = M(\mathbf{y})$. A more detailed analysis will be the object of a forthcoming paper.

Theorem 2. *Let $\mathbf{x} \in X^{\mathbb{N}}$ and $M \subseteq \mathcal{M}(X)$. There exists an $a \in I$ with $M(a\mathbf{x}) = M$ if and only if M is closed and connected with $\emptyset \neq M \subseteq \mathcal{M}(A(\mathbf{x}))$.*

Proof. First assume $M = M(a\mathbf{x})$. Then, by theorem 3.1.1 in [W2], M is nonempty, closed and connected. To derive the remaining relation $M \subseteq \mathcal{M}(A(\mathbf{x}))$ it suffices to show that every $x \in X \setminus A(\mathbf{x})$ has a neighborhood U with $\lim_{N\to\infty} \mu_{N,a\mathbf{x}}(U) = 0$. To show this take any neighborhood U with compact closure $\overline{U}$ and with $\overline{U} \cap A(\mathbf{x}) = \emptyset$. Such a U exists since X is a normal space. If $x_n \in \overline{U}$ for infinitely many n, $\overline{U}$ would, by compactness, contain an accumulation point of $\mathbf{x}$, contradiction. Thus $x_n \notin U$ for all $n \geq N_0$. This proves

$$\lim_{N\to\infty} \mu_{N,a\mathbf{x}}(U) \leq \lim_{N\to\infty} \frac{N_0}{N} = 0.$$

For the other direction assume that $M \neq \emptyset$ is closed and connected. By [W2, theorem 3.1.2], there is sequence $\mathbf{z} = (y_n)_{n\in\mathbb{N}}$ in $A(\mathbf{x})$ with $M(\mathbf{z}) = M$. By the definition of $A(\mathbf{x})$ there is a subsequence $a\mathbf{x}$ of $\mathbf{x}$ with $a = (a_1, a_2, \dots)$ such that $\lim_{n\to\infty} d(y_n, x_{a_n}) = 0$. This subsequence, obviously, has the required properties. □

Theorem 3. $M(a\mathbf{x}) = \mathcal{M}(A(\mathbf{x}))$ *holds for all $a \in R$ from a residual set $R \subseteq I$, i.e., most subsequences are maldistributed in $A(\mathbf{x})$.*

Proof. Let R denote the set of all a with $M(a\mathbf{x}) = \mathcal{M}(A(\mathbf{x}))$, i.e., $a\mathbf{x}$ is maldistributed in $A(\mathbf{x})$:

$$R = \{a : M(a\mathbf{x}) = \mathcal{M}(A(\mathbf{x}))\}$$

The inclusion $\subseteq$ follows from theorem 2. For the other one we have to prove that for every $\mu \in \mathcal{M}(A(\mathbf{x}))$ the set

$$R_\mu = \{a \in I : \mu \in M(a\mathbf{x})\}$$

is residual. Since $\mathcal{M}(X)$ is second countable, one may choose a sequence $(\mu_n)_{n\in\mathbb{N}}$ which is dense in $\mathcal{M}(A(\mathbf{x}))$. We now check that

$$R = \bigcap_{n\in\mathbb{N}} R_{\mu_n}$$

is a countable intersection of residual sets, hence residual. In the above equation the inclusion $\subseteq$ is trivial. But also the converse inclusion is clear: Let a be an element from the right hand side, then the set $M(a\mathbf{x})$ contains the topological closure of the set of all μ_n which is $\mathcal{M}(A(\mathbf{x}))$, hence $a \in R$.

To show that R_μ is residual for $\mu \in \mathcal{M}(A(\mathbf{x}))$, just observe

$$R_\mu = \bigcap_{n,k\in\mathbb{N}} O_{n,k}$$

with the sets $O_{n,k}$ of all $a \in I$ such that there exists an $n' > n$ with $d(\mu_{n' a\mathbf{x}}, \mu) < \frac{1}{k}$. Standard arguments show that each $O_{n,k}$ is dense and open. [$O_{n,k}$ is open since the map $a \mapsto d(\mu_{n,a\mathbf{x}}, \mu)$ is continuous (with respect to the 0-dimensional topology on I).

$O_{n,k}$ is dense: Take any μ-u.d. $\mathbf{y} = (y_n)_{n\in\mathbb{N}}$, $y_n \in A(\mathbf{x})$. Any subsequence $(x_{n_k})_{k\in\mathbb{N}}$ with $d(y_k, x_{n_k}) \to 0$ can be written as $a\mathbf{x}$ with $a \in O_{n,k}$.] Thus the countable intersection R_μ of these sets is residual, and the theorem is proved. □

Remark 4. A general principle containing all Baire arguments of the above type can be found in [Gr].

Theorem 5. *$M(a\mathbf{x}) = M(\mathbf{x})$ for λ_p-almost all a, $0 < p \leq 1$.*

Proof. We have to show that $\lambda_p(A) = 1$ for

$$A = \{a \mid M(\mathbf{x}) = M(a\mathbf{x})\}.$$

The idea of the proof is to show that, for almost all $a \in I$, for sufficiently large N the distance between the following two measures becomes small:

- the measure $\mu_{N,\mathbf{x}}$ which is induced by the finite sequence $(x_1, \dots, x_n)$,
- the measure $\mu_{a(\mathbf{x}\restriction N)}$, induced by the subset consisting of those x_n, $n \leq N$, with $n = a_k$ for some k

This can be done by considering, for any $f \in \mathcal{K}(X)$, the sets

$$A_f = \left\{a \mid \lim_{N\to\infty} |s(f,N) - s(f,N,a)| = 0\right\}$$

where

$$s(f,N) = \frac{1}{N}\sum_{n\leq N} f(x_n) \quad \text{and} \quad s(f,N,a) = \frac{1}{\#(S_a \cap (0,N])}\sum_{n\leq N, n\in S_a} f(x_n).$$

Since X is second countable, $\mathcal{K}(X)$ is separable w.r.t. the sup-norm. Hence there is a dense sequence $f_n, n \in \mathbb{N}$, in $\mathcal{K}(X)$. Observe that, by definition of the weak topology, this implies

$$A \supseteq \bigcap_{f\in\mathcal{K}(X)} A_f = \bigcap_{n\in\mathbb{N}} A_{f_n}.$$

Thus it suffices to prove $\lambda_p(A_f) = 1$ for arbitrary $f \in \mathcal{K}(X)$ where we may, w.l.o.g., assume $0 \leq f(x) < 1$ for all $x \in X$.

Fix therefore any continuous $f : X \to [0,1)$ and an integer $m \geq 1$. Consider the partition of $\mathbb{N}$ into subsets T_i, $i = 1, \dots, m$, defined in such a way that $n \in T_i$ if and only if $\frac{i-1}{m} \leq f(x_n) < \frac{i}{m}$. One easily computes

$$s(f,N) - s(f,N,a) \leq \frac{1}{m} \left(\sum_{i=1}^{m} i\alpha(i,N) - (i-1)\beta(i,N,a)) \right)$$

where

$$\alpha(i,N) = \frac{\#(T_i \cap (0,N])}{N} \quad \text{and} \quad \beta(i,N,a) = \frac{\#(T_i \cap S_a \cap (0,N])}{\#(S_a \cap (0,N])}.$$

Note

$$i\alpha(i,N) - (i-1)\beta(i,N,a) = (i-1)(\alpha(i,N) - \beta(i,N,T)) + \alpha(i,N)$$

and

$$\alpha(i,N) - \beta(i,N,a) = \alpha(i,N)\,(1 - \gamma(i,N,a)\delta(N,a))$$

with

$$\gamma(i,N,a) = \frac{\#(T_i \cap S_a \cap (0,N])}{\#(T_i \cap (0,N])} \quad \text{and} \quad \delta(N,a) = \frac{N}{\#(S_a \cap (0,N])}.$$

By a simple version of the strong law of large numbers we get that, if T_i is infinite, almost all a satisfy

$$\lim_{N\to\infty} \gamma(i,N,a) = p \quad \text{and} \quad \lim_{N\to\infty} \delta(N,a) = \frac{1}{p}.$$

If T_i is finite we have $\lim_{N\to\infty} \alpha(i,N) = \lim_{N\to\infty} \beta(i,N,a) = 0$ for almost all a. Combining all these facts we get

$$\limsup_{N\to\infty} s(f,N) - s(f,N,a) \leq \frac{1}{m} \limsup_{N\to\infty} \sum_{i=1}^{m} \alpha(i,N) = \frac{1}{m}$$

for almost all a. Similarly one gets $\lim_{N\to\infty} \sup s(f,N,a) - s(f,N) \leq \frac{1}{m}$, thus $\limsup_{N\to\infty} |s(f,N) - s(f,N,a)| \leq \frac{1}{m}$ for almost all $a \in I$. Since m was chosen arbitrarily from the countable set $\mathbb{N}$ we even get $\lim_{N\to\infty} s(f,N) - s(f,N,a) = 0$ for almost all a which means $\lambda_p(A_f) = 1$. This, by the arguments above, proves the theorem. □

Remark 6. For other metric results on the behavior of subsequences consult [L-T] and the papers quoted there.

For the rest of this section we will fix a sequence $\mathbf{x} = (x_n)_{n\in\mathbb{N}}$ and we will investigate the Hausdorff dimension of sets considered above.

Definition 7. For any non-empty closed connected set $M \subseteq M(A(\mathbf{x}))$ we define

$$\rho(M) = \rho(M, \mathbf{x}) = \sup\{\underline{d}(a) \mid M(a\mathbf{x}) = M\}$$

where $\underline{d}(a)$ denotes the lower density of the sequence a in $\mathbb{N}$.

Remark 8. Rainer Siegmund–Schultze has pointed out that the supremum in the definition is actually a maximum.

In view of theorem 1 this definition implies immediately

Corollary 9. $\rho(M) = 1 \iff M = M(\mathbf{x})$.

Definition 10. We set

$$\phi(M) = \phi(M, \mathbf{x}) = \dim_{\mathrm{H}}\{a \mid M(a\mathbf{x}) = M\}.$$

For $x \in (0,1)$ let

$\tau(x) = -x\log(x) - (1-x)\log(1-x)$ ("log" is the logarithm to the base 2)

Note that $0 \le \tau(x) \le 1$ for all $x \in (0,1)$, $\lim_{x\to 0}\tau(x) = 0$, and: $\tau(x) = 1$ only if $x = \frac{1}{2}$.

We are going to prove the following theorem.

Theorem 11. *Fix* $\mathbf{x}$*, and let* $M \subseteq M(A(\mathbf{x}))$ *be a non-empty closed connected set. Then*

$$\phi(M) = 0 \iff \rho(M) = 0$$

and

$$\phi(M) = 1 \iff \rho(M) = 1$$

Proof. **1.** $(\rho(M) = 0 \Longrightarrow \phi(M) = 0)$ In fact we will show that for small δ we have $\rho(M) \le \delta \Rightarrow \phi(M) \le \tau(\delta)$.

Let $A_0 = \{\underline{d}(a) = 0\}$. Note that $\rho(M) = 0$ implies that $\{a : M(a\mathbf{x}) = M\} \subseteq A_0$.

So fix δ. For any n let

$$\mathcal{C}_n(\delta) = \{[\varepsilon_1 \cdots \varepsilon_n] : \varepsilon_i \in \{0,1\}, \frac{\sum_{j=1}^n \varepsilon_j}{n} \le \delta\},$$

and let $C_n(\delta)$ be the union of all elements of $\mathcal{C}_n(\delta)$. Clearly, then,

$$A_0 \subseteq \bigcup_{k=N}^{\infty} C_k(\delta) \qquad \text{for each } N \in \mathbb{N}$$

Notice that the set $\mathcal{C}_n(\delta)$ consists of approximately $\sum_{k\le\delta n}\binom{n}{k} = o\big(n2^{-n\tau(\delta)}\big)$ many cylinders, each of diameter 2^{-n}. Hence for $s > \tau(\delta)$ the s-dimensional Hausdorff measure of A_0 can be estimated by

$$\lim_{N\to\infty}\sum_{n=N}^{\infty} n2^{-n\tau(\delta)}(2^{-n})^2 = 0$$

Hence $\dim_H(A_0) \le \tau(\delta)$. This completes the proof of part 1.

2. $(\rho(M) = 1 \Longrightarrow \phi(M) = 1)$ By theorem 1 we have that $M = M(\mathbf{x})$. Then theorem 5 implies that the set $\{a \mid M(a\mathbf{x}) = M(\mathbf{x})\}$ has full Lebesgue measure and hence Hausdorff dimension 1.

3.$(\phi(M) = 1 \Longrightarrow \rho(M) = 1)$

Let us assume that $\rho(M) < 1$. In view of the definition of $\rho(M)$ there is a measure $\mu \in M(\mathbf{x})\triangle M$. Let $\varepsilon = \frac{1}{10}d(\mu, M)$ if $\mu \in M(\mathbf{x})\setminus M$, and $\varepsilon = \frac{1}{10}d(\mu, M(\mathbf{x}))$ if $\mu \in M \setminus M(\mathbf{x})$.

To estimate the Hausdorff dimension of

$$A_0 := \{a : M(a\mathbf{x}) = M\}$$

we first claim that $a \in A_0$ implies that there are infinitely many $N \in \mathbb{N}$ such that $d(\mu_{\mathbf{x}\restriction N}, \mu_{a(\mathbf{x}\restriction N)}) \le 8\varepsilon$. To check this claim we assume without loss of generality $\mu \in M(\mathbf{x}) \setminus M$ (the other case is similar). There are infinitely many N such that

$d(\mu, \mu_{\mathbf{x}\restriction N}) \leq \varepsilon$, and for all large enough N we have $d(\mu_{a(\mathbf{x}\restriction N)}, M(a\mathbf{x})) \leq \varepsilon$, hence $d(\mu_{\mathbf{x}\restriction N}, \mu_{a(\mathbf{x}\restriction N)}) \geq 8\varepsilon$.

Hence we can find finitely many functions $f_1, \ldots, f_k$ such that the set A_0 is contained in

$$A_1 := \bigcup_{j=1}^{k} \{a : \exists^\infty N \ |s(f_j, N) - s(f_j, a, N)| \geq 4\varepsilon\}$$

It is therefore enough to show that for any continuous f and any $\varepsilon > 0$ the set $\{a : \exists^\infty N \ \big|s(f, N) - s(f, N, a)\big| \geq 4\varepsilon\}$ has Hausdorff dimension < 1.

So fix f and ε. Let $m > 1/\varepsilon$, and define $T_1, \ldots, T_m$ as in the proof of theorem 5.

Write $A_{N,\varepsilon}$ for the set

$$A_{N,\varepsilon} := \{a : |s(f, N) - s(f, N, a)| \geq 4\varepsilon\}$$

Note that $a \in A_{N,\varepsilon}$ implies that there is some $i \in \{1, \ldots, m\}$ such that

$$\Big|\frac{1}{N} \sum_{\substack{j \in T_i \\ j \leq N}} f(x_j) \ - \ \frac{1}{\# S_a \cap \{1, \ldots, N\}} \sum_{\substack{j \in T_i \cap S_a \\ j \leq N}} f(x_j)\Big| \geq 4\varepsilon \frac{\# T_i}{N}$$

Replacing $f(x_j)$ by $\frac{i}{m}$ introduces an error $\leq \frac{1}{m} < \varepsilon$ in each term, hence

$$\Big|\frac{1}{N} \sum_{\substack{j \in T_i \\ j \leq N}} \frac{i}{m} \ - \ \frac{1}{\# S_a \cap \{1, \ldots, N\}} \sum_{\substack{j \in T_i \cap S_a \\ j \leq N}} \frac{i}{m}\Big| \geq 3\varepsilon \frac{\# T_i}{N}$$

Employing the functions γ and δ that we defined in the proof of theorem 5 and using $\frac{i}{m} \leq 1$ we conclude that $a \in A_{N,\varepsilon}$ implies

$$1 - \delta(N, a) \cdot \gamma(i, N, a) \geq 3\varepsilon \qquad \text{for some } i$$

so $\delta(N, a) \leq 2 - \varepsilon$ or $\gamma(i, N, a) \geq \frac{1}{2} - \varepsilon$.

Hence

$$A_1 = \bigcap_{n=1}^{\infty} \bigcup_{N=n}^{\infty} A_{N,\varepsilon} \subseteq \bigcap_{n=1}^{\infty} \bigcup_{N=n}^{\infty} \{a : \delta(N, a) < 2 - \varepsilon\} \cup \bigcap_{n=1}^{\infty} \bigcup_{N=n}^{\infty} \{a : \gamma(i, N, a) < \frac{1}{2} - \varepsilon\}$$

We will only estimate the Hausdorff dimension of the sets

$$\bigcap_{n=1}^{\infty} \bigcup_{N=n}^{\infty} \{a : \gamma(i, N, a) \leq \frac{1}{2} - \varepsilon\}$$

leaving the second (similar) estimate to the reader.

As in the first step of the proof of theorem 5, let

$$\mathcal{C}_N(\varepsilon) := \{[\varepsilon_1, \ldots, \varepsilon_n] : \frac{1}{N_i} \sum_{\substack{j \leq n \\ j \in T_i}} \varepsilon_j \leq \frac{1}{2} - \varepsilon\} \qquad (\text{where } N_i := \#(T_i \cap [1, \ldots, N])),$$

and let $C_N(\varepsilon)$ be the union of all the cylinders in $\mathcal{C}_N(\varepsilon)$. Again, $\mathcal{C}_N(\varepsilon)$ contains at most $\binom{N_i}{(\frac{1}{2} - \varepsilon)N_i} = o\big(N 2^{-N\tau(\frac{1}{2} - \varepsilon)}\big)$ many cylinders, each of them of diameter 2^{-N_i}. Since A_1 is covered by $\bigcup_{n=N}^{\infty} C_n(\varepsilon)$ for any $N \in \mathbb{N}$, we can as before bound the the s-dimensional Hausdorff measure of A_1 for any $s > \tau(\frac{1}{2} - \varepsilon)$ by

$$\lim_{N \to \infty} \sum_{n=N}^{\infty} n 2^{n\tau(\frac{1}{2} - \varepsilon)} 2^{-ns} = 0$$

Hence the Hausdorff dimension of A_0 is $\leq \tau(\frac{1}{2}-\varepsilon) < 1$.

4. **($\phi(M) = 0 \Longrightarrow \rho(M) = 0$)** Let us assume that $\rho(M) = d > 0$. In view of the definition of $\rho(M)$ there is a sequence a with $M(a\mathbf{x}) = M$ and positive lower density $\underline{d}(a) > d - \delta$. We are going to consider the subset in $[0,1]$ corresponding to all subsequences of a. For this we consider for $N \in \mathbb{N}$ the union of cylinder sets $\mathcal{J}_N$ defined by

$$[\varepsilon_1 \cdots \varepsilon_N] \in \mathcal{J}_N \quad \Longleftrightarrow \quad \varepsilon_j = 0 \text{ unless } j = a_k \text{ for } j \in [0,N] \text{ and some } k.$$

We set

$$\Lambda_a = \bigcap_{N=1}^{\infty} \mathcal{J}_N = \{b : S_b \subseteq S_a\}.$$

There is a natural identification of Λ_a with the space $\{0,1\}^{\mathbb{N}}$ of all $\{0,1\}$-sequences. We define a measure μ_a supported on Λ_a as the push–forward of the Bernoulli measure on $\{0,1\}^{\mathbb{N}}$ generated by the probability vector $(\frac{1}{2},\frac{1}{2})$

$$\mu_a(C_N(b)) = \begin{cases} \left(\frac{1}{2}\right)^{\#(S_a \cap [1,N])} & \text{if } b \in \Lambda_a \\ 0 & \text{if } C_N(b) \cap \Lambda_a = \emptyset \end{cases}$$

This measure is a probability measure supported on Λ_a and has the following property

$$\liminf_{N\to\infty} -\frac{\log \mu_a(C_N(b))}{N \log 2} = \liminf_{N\to\infty} \frac{\#(S_a \cap [0,N])}{N} = \underline{d}(a) > 0 \qquad \text{for } \mu_a\text{–a.e. } b$$

Frostman's Lemma (see [Fal, p.221]) tells us that

$$\dim_{\mathrm{H}} \mu_a = \underline{d}(a).$$

Now we want to apply theorem 5 to the case $\mathbf{y} = a\mathbf{x}$. In order to do this we identify the subsequences of a with points in the interval $[0,1)$ in the standard way. Under this identification the measure μ_a transforms into the normalized Lebesgue measure on $[0,1)$. Hence, theorem 5 implies that μ_a–a.e. subsequence b of a has the same limit behavior as a — i.e., $M(a\mathbf{x}) = M(b\mathbf{x}) = M$. By the above considerations the set of these subsequences b has Hausdorff dimension $\underline{d}(a)$. This yields $\phi(M) \geq \underline{d}(a) > 0$. ☐

We actually proved slightly stronger results than stated in theorem 11. These results are stated below.

Corollary 12.

$$\dim_{\mathrm{H}}\{b \mid M(b\mathbf{x}) = M(a\mathbf{x})\} \geq \underline{d}(a).$$

Remark 13. The estimate in corollary 12 is not very effective for $\underline{d}(a) \leq \frac{1}{2}$. In this case the estimate

$$\dim_{\mathrm{H}}\{b \mid M(b\mathbf{x}) = M(a\mathbf{x})\} \geq \tau(\underline{d}(a)).$$

is better. One can derive this estimate along similar lines as in the proof of part 4 of the theorem.

We also want to state another consequence of theorem 11:

Theorem 14. *Let $\mathbf{x}$ be an arbitrary sequence on X. Assume that $M(\mathbf{x})$ contains a non-atomic measure. Then the set of all a such that $M(a\mathbf{x}) \neq M(\mathbf{x})$ has Hausdorff dimension* 1.

Proof. By corollary 9

$$\{a \mid M(a\mathbf{x}) \neq M(\mathbf{x})\} \supset \{a \mid \rho(M(a\mathbf{x})) < 1\}.$$

Let $\mu \in M(\mathbf{x})$ be a non-atomic measure. We fix $\varepsilon > 0$ and choose an open subset $U \subset X$ with measure $\mu(U) = \varepsilon$. We note that such subsets U exist since μ is non-atomic. Let a be the subsequence defined by

$$k \in a \quad \Longleftrightarrow \quad x_k \notin U.$$

By the definition of the lower density and of $M(a\mathbf{x})$ this sequence has a subsequence a' with density $\underline{d}(a') = 1 - \varepsilon$ and $\mu \notin M(a'\mathbf{x})$. By corollary 12 we get

$$\dim_{\mathrm{H}}\{b \mid \rho(M(b\mathbf{x})) < 1\} \geq \dim_{\mathrm{H}}\{b \mid M(b\mathbf{x}) = M(a'\mathbf{x})\} \geq \underline{d}(a') = 1 - \varepsilon$$

Since ε was arbitrary this concludes the proof. □

3. Sequences of maps

Let now $\mathbf{f}$ be a sequence of maps $f_n : X \to Y$. Thus $\mathbf{f} : \mathbb{N} \times X \to Y$, $\mathbf{f}(n, x) = f_n(x)$. For notational convenience we also write $\mathbf{f}(x)$ for the sequence $(f_n(x))_{n \in \mathbb{N}}$. To $\mathbf{f}$ there corresponds a map $\Phi = \Phi_{\mathbf{f}} : I^2 \to X^{\mathbb{N}}$, $(a, x) \mapsto a\mathbf{f}(x)$ where $a\mathbf{f}(x)$ is the subsequence $(f_{k_n}(x))_{n \in \mathbb{N}}$ of $(f_n(x))_{n \in \mathbb{N}}$ which corresponds to a in the sense described in the first section. This point of view is similar to that in [W1], where several related metric results can be found.

Let $T \subseteq I \times X$, $a \in I$, $x \in X$, then ${}_aT = \{x \in X \mid (a, x) \in T\}$ and $T_x = \{a \in I \mid (a, x) \in T\}$ denote the vertical and horizontal cross-section of T. Important informations about T and its cross-sections are their Hausdorff dimensions $\dim_H$, their d-dimensional Hausdorff measures m_d and their Baire categories as subsets of $I \times X$ resp. X or I. We turn to the Baire aspect.

Theorem 15. *Let X, Y be complete metric spaces, Y second countable, and $\mathbf{f}$ a sequence of continuous maps $f_n : X \to Y$. Suppose $D \subseteq Y$ is a dense set such that for every $y \in D$ the sets*

$$D_y = \bigcup_{n \in \mathbb{N}} f_n^{(-1)}(y)$$

are dense in X in a strong sense, meaning that every nonempty open set in X has an infinite intersection with D_y. Then the following statements hold:
(i) There is a residual set $R \subseteq I \times X$ such that $a\mathbf{f}(x)$ is maldistributed for all $(a, x) \in R$.
(ii) There is a residual set $R \subseteq I$ such that for all $a \in R$ the sequence $a\mathbf{f}(x)$ is maldistributed for all x from a residual set $R(a) \subseteq X$.
(iii) There is a residual set $R \subseteq X$ such that for all $x \in X$ $a\mathbf{f}(x)$ is maldistributed for all a from a residual set $R(x) \subseteq I$.

Proof. By the theorem of Kuratowski-Ulam (cf. [Ox]) the three statements of the theorem are equivalent. Hence it suffices to prove (iii). For an arbitrary nonempty open set $O \subseteq Y$ define

$$R(O) = \bigcup_{n \in \mathbb{N}} f_n^{-1}(O).$$

By assumption $R(O)$ is open and dense. Since Y is second countable it has a countable topological base of nonempty open sets O_k, $k \in \mathbb{N}$. Since X is a complete metric space this implies that

$$R = \bigcap_{k \in \mathbb{N}} R(O_k)$$

is residual in X. If $x \in R$ and U is a nonempty open set in Y we have $O_k \subseteq U$ for some k and, since $x \in R(O_k)$, $x \in f_n^{-1}(O_k)$ for some $n \in \mathbb{N}$, i.e. $f_n(x) \in O_k \subseteq U$. This proves that $(f_n(x))_{n \in \mathbb{N}}$ is dense in Y, i.e., $A(\mathbf{f}(x)) = Y$. Thus (iii) follows from theorem 3. □

Remark 16. To get an analogous metric result [with "maldistributed" replaced by "uniformly distributed"] one has to use Fubini's theorem instead of Kuratowski-Ulam's. A sufficient condition on the sequence of functions for this case is mentioned in Koksma's metric theorem (cf. [K-N, p.34]).

Elementary examples 17. If the maps f_n are expanding in an appropriate sense the condition of theorem 15 might be fulfilled, maybe under additional hypotheses on X and Y. Concrete examples are: $X = T \subseteq \mathbb{R}$, $Y = I$, $f_n(x) = a_n x$, $f_n(x) = x^{a_n}$, $f_n(x) = a_n^x$ etc., or similar sequences on the s-dimensional torus $X = Y = I^s$.

Examples 18. The most natural application is to consider subsequences of ergodic iterations–i.e. sequences T^{n_k} where $T : X \to X$ is a map. The most interesting case (from the point of view of uniform distributions) would be if T is ergodic. This means

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(A \cap T^{-k} B) = \mu(A)\mu(B)$$

for all measurable sets A and B. This property (for subsequences) obviously implies the above assumptions for the support of the measure. Unfortunately, this property is not inherited by subsequences. (This is one reason why subsequences of $n\alpha$ are so interesting.) Therefore it might be better to consider systems which are topologically mixing. These are maps such that for all pairs of open sets U and V there is a number N such that for all $n \geq N$ the intersection $U \cap T^N(V)$ is not empty. This condition follows if for all measurable sets A and B

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

holds. Maps satisfying the latter condition are called mixing. There are many systems which are known as mixing. For instance any transitive expanding map of the n–torus is topologically mixing. Also a broad class of systems decomposes into mixing components. Unfortunately, group translations, and hence $n\alpha$–sequences, are not mixing.

4. Results on $n\alpha$-sequences

We now turn to a special case of the sequences $(f_n(x))_{n \in \mathbb{N}}$ discussed in the previous section, namely, the sequence $(n\alpha)_{n \in \mathbb{N}}$ and its subsequences.

If $a \in [0,1)$, $S_a = \{k_1 < k_2 < \cdots\}$, then we will identify a with the sequence $(k_1, k_2, \ldots)$. For any $\alpha \in \mathbb{R}/\mathbb{Z}$ we let $\boldsymbol{\alpha}$ denote the sequence $(\alpha, 2\alpha, 3\alpha, \ldots)$, so $a\boldsymbol{\alpha}$ will be the subsequence $(k_1\alpha, k_2\alpha, \ldots)$.

In this section we will use the "forcing" notation, which is a *façon de parler* that allows us to talk about "the typical element" of a Polish (i.e., complete separable metric) space.

(We use this notation only for proofs, not for the theorems themselves. The reader who does not want to learn the forcing calculus will have no difficulty translating our proofs into more conventional ones.)

Definition 19. 1. Let X be a Polish space. We call the nonempty open set of X "conditions" and we view them as giving us information about the "typical element $\dot{x} \in p$". If p, q are conditions with $p \supseteq q$, we write $p \leq q$ to emphasize that "$\dot{x} \in q$" contains more information than "$\dot{x} \in p$."

For any $B \subseteq X$ we write $p \Vdash B$ ("p forces B") as an abbreviation for "except for a meager set, every element of p is in B" or more formally: "$p \setminus B$ is meager in X".

2. For each space X we reserve a special variable $\dot{x}$ to denote the "generic" or "typical" element of X. The "typical" real number will be denoted $\dot{\alpha}$.
   For any formula (or property) φ we write
   $$p \Vdash \varphi(\dot{x})$$
   iff $p \Vdash \{x : \varphi(x)\}$, i.e., if "most" elements of p have the property φ (or equivalently, if most elements x of X have the property: "if $x \in p$, then $\varphi(x)$.")

Fact 20. *If B is a nonmeager Borel set, then there is a condition (i.e, a nonempty open set) p with $p \Vdash B$.*

Proof. All Borel sets have the Baire property (cf. [Ox, chapter 4]). □

Fact 21. *Assume that $\varphi(x)$ and $\psi(x)$ are formulas such that the corresponding sets $\{x : \varphi(x)\}$ and $\{x : \psi(x)\}$ are Borel sets.*

1. *If $p \Delta q$ is meager, then $p \Vdash \varphi(\dot{x})$ iff $q \Vdash \varphi(\dot{x})$. (We will not need the converse, which is also true: If "$p \Vdash \varphi(\dot{x})$ implies $q \Vdash \varphi(\dot{x})$" for every Borel formula $\varphi(\dot{x})$, then $p \setminus q$ is meager.)*
2. *$p \Vdash \varphi(\dot{x}) \& \psi(\dot{x})$ iff: $p \Vdash \varphi(\dot{x})$ and $p \Vdash \psi(\dot{x})$.*
3. *"$p \Vdash \varphi(\dot{x})$" is false iff there is $q \geq p$ such that $q \Vdash \neg\varphi(\dot{x})$. (Apply fact 20 to $p \setminus \{x : \varphi(x)\}$.)*
4. *If $p \Vdash \exists k \in \mathbb{N} \varphi(\dot{x}, k)$ then there is $q \geq p$ and $k_0 \in \mathbb{N}$ such that $q \Vdash \varphi(\dot{x}, k_0)$. (This is a thinly disguised version of Baire's theorem.)*
5. *It is impossible that $p \Vdash \varphi(\dot{x})$ and $p \Vdash \neg\varphi(\dot{x})$.*
6. *If $p \Vdash \varphi(\dot{x})$, $q \geq p$, then also $q \Vdash \varphi(\dot{x})$.*

Theorem 22. *Let $a = (k_0, k_1, k_2, \dots)$ be a sequence of natural numbers. If $\liminf_n (k_{n+1}/k_n) > 1$, then the set*
$$\{\alpha : a\dot{\boldsymbol{\alpha}} \text{ is u.d. }\}$$
is meager.

This theorem answers a question of Šalát [Sa], who has proved a similar result under the assumption $\liminf_n (k_{n+1}/k_n) \geq 2$.

We prepare ourselves for the proof of theorem 22 with a definition and a lemma.

Definition 23. We say that a family of functions $(f_1, f_2, \dots)$ is ε-mixing if: for all intervals J of length ε and all $k \in \mathbb{N}$:
$$\bigcap_{n=1}^{k} f_n^{-1}(J) \neq \emptyset$$

More generally we say that $(f_1, f_2, \dots)$ is (ε, δ)-mixing if: for all intervals J of length ε, and all $k \in N$, and all intervals J' of length δ:

$$J' \cap \bigcap_{n=1}^{k} f_n^{-1}(J) \neq \emptyset$$

Remark 24. Although we work in $I = [0, 1)$, we identify elements in I with their equivalence classes in $\mathbb{R}/\mathbb{Z}$, so an (open) interval can either be of the form (a, b) or of the form $[0, a) \cup (b, 1)$ (for $0 \leq a < b \leq 1$). However, since we are mainly concerned with very short intervals, it is no loss of generality to only consider intervals of the first form.

Lemma 25. *Let f_n be the function mapping α to $k_n\alpha$ modulo 1, where $(k_n)_{n \in \mathbb{N}}$ is a sequence of natural numbers satisfying*

1. $k_{n+1} > \frac{2}{\varepsilon} k_n$ *for all n.*
2. $k_1 > \frac{\varepsilon}{\delta}$.

Then $(f_1, f_2, \dots)$ is (ε, δ)-mixing.

Proof. Let J be an interval of length ε, J' an interval of length δ.

We will show (by induction on m) that each set

$$J' \cap \bigcap_{i=1}^{m} f_i^{-1}(J)$$

contains in fact an interval I_m of length ε/k_m. This is clear for $m = 0$, as the length of J' is $\delta > \varepsilon/k_0$.

Consider $m > 0$. Note that $f_m^{-1}(J) = \bigcup_{j=1}^{k_m} \{\alpha : k_m \cdot \alpha - j \in J\}$ is a union of k_m many disjoint intervals, each of length ε/k_m.

By inductive assumption, the set $J' \cap \bigcap_{i=1}^{m-1} f_i^{-1}(J)$ contains an interval I_{m-1} of length ε/k_{m-1}. Since $\varepsilon/k_{m-1} > 2/k_m$, we can find a natural number j such that the interval

$$[\frac{j}{k_m}, \frac{j+1}{k_m}] = \{\alpha : k_m \cdot \alpha - j \in [0, 1]\}$$

is contained in I_{m-1}. Hence the set

$$I_m := \{\alpha : k_m \cdot \alpha - j \in J\},$$

an interval of length ε/k_m, is also contained in I_m. □

Proof of theorem 22 Let $1 < q < \liminf_n (k_{n+1}/k_n)$. Without loss of generality we have $k_{n+1}/k_n > q$ for *all* k.

Let $\dot\alpha$ be the "generic" real number. Towards a contradiction, assume that "$a\dot{\boldsymbol\alpha}$ is not u.d." is not forced, then by fact 21(3) there is a condition p such that

$$p \Vdash a\dot{\boldsymbol\alpha} \text{ is u.d.}$$

Pick an integer $c > 1$ such that $q^c \geq 9c$, and let $\varepsilon = \frac{1}{3c}$. Let J be any interval of length ε. So we have

$$p \Vdash \frac{1}{n} \#\{j < n : k_j \dot\alpha \in J\} \to \frac{1}{3c}$$

and consequently:

$$p \Vdash \exists n_0 \, \forall n \geq n_0 : \; \#\{j < n : k_j \dot\alpha \in J\} \leq \frac{1}{2c} \cdot n$$

Without loss of generality (replacing p by some $p' \geq p$ if necessary) we may assume p is an interval and by fact 21(4) we may without loss of generality assume that there is n_0 such that

$$\forall n \geq n_0: \quad p \Vdash \#\{j < n : k_j \dot\alpha \in J\} \leq \frac{1}{2c} \cdot n.$$

Replacing n by nc we get

$$\forall n \geq n_0: \quad p \Vdash \#\{j < nc : k_j \dot\alpha \in J\} \leq \frac{1}{2c} \cdot nc = \frac{1}{2}n \qquad (*)$$

Let δ be the length of the interval p. By increasing n_0, if necessary, we may assume $k_{n_0 c} > \frac{\varepsilon}{\delta}$. Define for $i = 0, 1, 2, \ldots, 100n_0$:

$$f_i(x) = k_{(n_0+i)c} x.$$

Note that $k_{(n_0+i+1)c}/k_{(n_0+i)c} \geq q^c \geq 9c > 6c = 2/\varepsilon$. Hence (by lemma 25) the functions $f_0, f_1, \ldots$ are (ε, δ)-mixing, so there is an interval $q \subseteq p$ of positive length such that

$$q \subseteq \bigcap_{i=0}^{100n_0} f_i^{-1}(J)$$

and hence:

$$q \Vdash \forall i \in \{0, \ldots, 100n_0\} : k_{(n_0+i)c} \dot\alpha \in J.$$

But this means that

$$q \Vdash \#\{j \leq 100n_0 : k_j \alpha \in J\} \geq \#\{n_0 c, (n_0+1)c, \ldots, 100n_0 c\} \geq 99n_0 > \frac{1}{2}(100n_0)$$

a contradiction to $(*)$.

Theorem 26. *If $a = (k_n)_{n\in\mathbb{N}}$ with $\lim k_{n+1}/k_n = \infty$, then the set $\{\alpha : a\boldsymbol{\alpha} m.d.\}$ is residual.*

Proof. Similar to the above (but easier). ☐

Obviously, "lim inf" is not an appropriate concept for questions dealing with uniform distribution, since for example a modification of a very thin subsequence of a sequence of reals can change the limes inferior without affecting the distribution behavior of the sequence. A more appropriate concept would be the following:

Definition 27. Let $\mathcal{J}$ be an ideal of subsets of $\mathbb{N}$, $x = (x_n)_{n\in\mathbb{N}}$ a sequence of real numbers. We define the lower limit (with respect to $\mathcal{J}$) of this sequence as follows:

$$\liminf{}^{\mathcal{J}} x := \sup\{t \in \mathbb{R} \mid \{n : x_n < t\} \in \mathcal{J}\},$$

or equivalently,

$$\liminf{}^{\mathcal{J}} x := \sup\{t \in \mathbb{R} \mid \{n : x_n \leq t\} \in \mathcal{J}\}.$$

We define $\limsup^{\mathcal{J}}$ dually: $\limsup^{\mathcal{J}} x := \inf\{t \in \mathbb{R} \mid \{n : x_n \geq t\} \in \mathcal{J}\}$.

Thus, the usual notion of lim inf is $\liminf^{\mathrm{fin}}$, where fin is the ideal of finite subsets of $\mathbb{N}$. In our context the ideal of sets of density 0 is the more interesting notion.

For example, we can refine the previous theorem to the following:

Remark 28. Let $\mathcal{J}$ be the ideal of sets with density 0. Let $a = (k_1, k_2, \dots)$, $k_1 < k_2 < \cdots$. If $\liminf_n^{\mathcal{J}} (k_{n+1}/k_n) > 1$, then the set

$$\{\alpha : a\alpha \text{ is u.d. }\}$$

is meager.

We leave the (easy) proof of this fact, as well as the formulation of similar generalizations for other theorems, to the reader.

If $\lim \frac{k_{n+1}}{k_n} = 1$, then the behavior of $(k_n\alpha)$ is more complicated. Of course there are cases (such as the trivial case $k_n = n$) where the "typical" behavior of $(k_n\alpha)_{n\in\mathbb{N}}$ will be uniform distribution, but we will show that there are also "many" sequences $(a_n)_{n\in\mathbb{N}}$ for which k_{n+1}/k_n converges to 1 rather quickly, and still $(k_n\alpha)$ is maldistributed for all irrational α.

For $k_{n+1}/k_n \to 1$ we will measure the convergence speed by considering the asymptotic behavior of $c_n := (k_{n+1}/k_n - 1) * n$. This is justified by the dichotomy shown in theorems 30 and 34.

Definition 29. For any positive number c let

$$X_c := \{(k_n)_{n\in\mathbb{N}} \in \mathbb{N}^{\mathbb{N}} : \forall n\ k_n < k_{n+1} \leq (1 + \frac{c}{n})k_n\}$$

Note that as a closed subset of $\mathbb{N}^{\mathbb{N}}$, X_c is a complete metric space (hence not meager).

Theorem 30. *Let $c \geq 2$. Then X_c is a perfect Polish space, and for every irrational number α the set*

$$\{a \in X_c : a\boldsymbol{\alpha} \text{ is m.d.}\}$$

is residual in X_c.

This result is similar to a theorem of Erdős and Taylor [E-T, Theorem 11], who proved that there is a sequence $a = (k_n)$ in which the differences $k_{n+1} - k_n$ are even bounded, such that for an uncountable set of α's the sequence $a\boldsymbol{\alpha}$ is not u.d. By relaxing $k_{n+1} - k_n \leq C$ to $k_{n+1}/k_n \leq 1 + c/n$ we get a different space, in which a stronger assertion holds: Instead of a single sequence there is now a residual set of such sequences, and instead of a large (uncountable, even residual) set of α's we get our result for all irrational α.

Definition 31. We let X_c^{fin} be the set of all finite sequences $(k_1, \dots, k_\ell)$ satisfying $\frac{k_{n+1}}{k_n} \leq 1 + \frac{c}{n}$ for all $n < \ell$.

We need an easy lemma about extending sequences in X_c^{fin}. Part (1) of the following lemma is basically due to the fact that the product $\prod(1 + \frac{2}{n})$ diverges.

Lemma 32.

1. *Assume that $(k_1, \dots, k_n) \in X_c^{\text{fin}}$, and let $R \in \mathbb{N}$. Then there is $m \in \mathbb{N}$, $m > n$ and a sequence $(k_{n+1}, \dots, k_m)$ such that*
   - $(k_1, \dots, k_m) \in X_c^{\text{fin}}$
   - $k_m \geq 2R \cdot m$
2. *Let n, R, m be as above, and assume that $k_{m+\ell+1} - k_{m+\ell} \in [R, 2R]$ for all $\ell \geq 0$. Then*
   (a) *for all i: $k_{m+i} \geq R(m+i)$*

(b) $(k_1, k_1, \dots) \in X_c$

Proof of (1). We may assume $n \geq 4R$ [otherwise we start by replacing $(k_1, \dots, k_n)$ by $(k_1, \dots, k_n, k_n+1, k_n+2, \dots, k_n+(4R-n))$]. Let $m = n^3$, and define

- $k_{n+\ell} = k_n + 2\ell$ for $n + \ell \leq n^2$.
- $k_{jn^2+\ell} = k_{kn^2} + (k+1)\ell$ for $j = 1, \dots, n-1$, $\ell = 1, \dots, n^2$.

We first notice $k_i \geq i$ for all i, hence

$$\frac{k_{n+\ell+1}}{k_{n+\ell}} \leq 1 + \frac{2}{k_{n+\ell}} \leq 1 + \frac{2}{n+\ell}$$

for all $\ell < n^2 - n$. Next, we see by induction that

$$k_{jn^2} \geq \frac{j(j+1)}{2} n^2 \quad \text{for } j = 1, \dots, n.$$

so

$$\frac{k_{jn^2+\ell+1}}{k_{jn^2+\ell}} \leq 1 + \frac{j+1}{k_{jn^2+\ell}} = 1 + \frac{j+1}{k_{jn^2} + \ell(j+1)} \leq$$
$$\leq 1 + \frac{j+1}{\frac{1}{2}j(j+1)n^2 + \ell(j+1)} = 1 + \frac{2}{jn^2 + 2\ell} \leq 1 + \frac{2}{jn^2 + \ell}.$$

Finally, we have $k_{n^3} \geq \frac{n(n+1)}{2} n^2 > n^3 \cdot \frac{n}{2} \geq Rn^3$.

The proof of (2) is trivial. □

Fact 33. *If $\mu_{N,\mathbf{x}} \to \mu_0$ (weakly, i.e., $\mathbf{x}$ is μ_0-u.d.), $U \subseteq \mathcal{M}(X)$ an open set containing μ_0, then there is $\varepsilon > 0$ such that for all $\mathbf{y} \in X^{\mathbb{N}}$:*

$$\textit{If } \|\mathbf{x} - \mathbf{y}\|_\infty < \varepsilon, \textit{ then } M(\mathbf{y}) \subseteq U$$

Proof of theorem 30:

Fix an irrational number α, and assume that the set

$$\{a \in X_c : a\boldsymbol{\alpha} \text{ is m.d.}\}$$

is not residual. So there is an open set p such that

$$p \Vdash \dot{a}\boldsymbol{\alpha} \text{ is not m.d.}$$

Now "$\dot{a}\boldsymbol{\alpha}$ is not m.d." is equivalent to "$M(\dot{a}\boldsymbol{\alpha}) \neq \mathcal{M}(X)$", so since $M(\dot{a}\boldsymbol{\alpha})$ must be closed this means

$$p \Vdash \text{there is a basic open set } U \text{ such that } U \cap M(\dot{a}\boldsymbol{\alpha}) = \emptyset.$$

By 21(4) can we can find a basic open set $U \subseteq \mathcal{M}(I)$ and a condition $p' \geq p$ such that

$$p' \Vdash U \cap M(\dot{a}\boldsymbol{\alpha}) = \emptyset$$

By strengthening p' to a condition q and by shrinking U we can even get

$$q \Vdash \forall n : \mu_{n,\dot{a}\boldsymbol{\alpha}} \notin U$$

Now we will use the fact that q gives only finitely much information about $\dot{a}$.

Let $\mu_0 \in U$, and let $\mathbf{y} = (y_n)_{n\in\mathbb{N}}$ be a sequence with $\mu_{n,\mathbf{y}} \to \mu_0$.

Choose ε as in fact 33. The condition q is a cylinder $q = [q_1, \dots, q_n]$. Without loss of generality $1/n < \varepsilon$.

Choose R so large that the set $\{\alpha, 2\alpha, \dots, R\alpha\}$ meets every interval of length ε (mod 1).

By lemma 32 we may extend the sequence $(q_1, \dots, q_n)$ to a sequence $(q_1, \dots, q_m)$ satisfying $q_m \geq 2Rm$. For all $i > m$ we can now choose

$$q_i \in [q_{i-1} + R, q_{i-1} + 2R]$$

such that $\alpha q_i \in [y_i - \varepsilon, y_i + \varepsilon] \pmod 1$, and $(q_n)_{n\in\mathbb{N}} \in X_c$. Let $\mathbf{q} = (q_n)_{n\in\mathbb{N}}$. We have $M(\mathbf{q}) \subseteq U$, so for large enough n we have $\mu_{n,\mathbf{q}\boldsymbol{\alpha}} \in U$. Since trivially $(q_1, \dots, q_n) \Vdash \dot{k}_1 = q_1, \dots, \dot{k}_n = q_n$, we have

$$(q_1, \dots, q_n) \Vdash \mu_{n,\dot{a}\boldsymbol{\alpha}} \in U,$$

a contradiction.

We now deal with the (easier) case where $c < 2$. It turns out that in this case the space X_c has a trivial structure: It contains a countable dense set of isolated points.

Theorem 34. *Let $c < 2$. Then for every irrational number α the set*

$$\{a = (k_n)_{n\in\mathbb{N}} \in X_c : a\boldsymbol{\alpha} \text{ is u.d.}\}$$

contains a countable dense open subset of X_c.

Proof. Let

$$X_c^* := \{a \in X_c : \exists \exists N \, \forall n \geq N \; k_n = n+\}$$

Clearly X_c^* is a countable dense subset of X_c, and for any irrational α we know that $a\dot{\alpha}$ is u.d. for every $a \in X_c^*$. It remains to show that X_c^* is open.

We will show that all points of X_c^* are isolated in X_c.

Assume that for all $n > N$ we have $k_n = n+$. Without loss of generality we have $N \cdot (\frac{2}{c} - 1) >$, so $(n+) \cdot \frac{c}{n} < 2$ for all $n \geq N$. Let

$$U := \{\mathbf{y} = (y_n)_{n\in\mathbb{N}} \in X_c : y_1 = k_1, \dots, y_N = k_N\}$$

Then U is a basic open neighborhood of a, and a is the only point in U. $\square$

Consider the following question: given a, what is the behavior of $a\boldsymbol{\alpha}$ for "typical" α? From theorem 3 we know that for a residual set of a's the typical behavior of $a\boldsymbol{\alpha}$ is maldistribution. (An explicit example of such an a is mentioned in theorem 26.) For some a's (e.g., $a = (1, 2, 3, \dots)$) the typical behavior is uniform distribution.

In general, the following dichotomy holds:

Theorem 35. *For any $a = (k_n)_{n\in\mathbb{N}}$ the set*

$$A = \{\alpha \in \mathbb{R}/\mathbb{Z} : a\boldsymbol{\alpha} \text{ is u.d.}\}$$

is either meager or residual.

Proof. Assume that A is not meager, then there is an interval $[a, b]$ such that $A \cap [a, b]$ is residual in $[a, b]$. Clearly, $\alpha \in A$ implies $n\alpha \in A$ for any $n \in \mathbb{N}$. Now for $n > \frac{1}{b-a}$ the set $n \cdot A$ is residual in $n \cdot [a, b] = \mathbb{R}/\mathbb{Z}$. $\square$

Open Problem 36. Is it true that

for all a, one of the sets

$$\{\alpha : a\boldsymbol{\alpha} \text{ is u.d.}\} \qquad \{\alpha : a\boldsymbol{\alpha} \text{ is m.d.}\}$$

is residual?

The proof of theorem 22 suggests a negative answer.

References

[Be] D. Berend. IP-sets on the circle. *Can. J. Math.*, **XLII**:575–589, 1990.

[D-T] M. Drmota and R. F. Tichy. *Sequences, Discrepancies and Applications*, volume 1651 of *Lecture Notes in Mathematics*. Springer, 1997.

[E-T] P. Erdős and S. J. Taylor. On the set of points of convergence of a lacunary trigonometric series and the equidistribution properties of related sequences. *Proc. Lond. Math. Soc.*, **7**:598–615, 1957.

[Fal] Falconer. *Fractal Geometry*. Mathematical Foundations and Applications. John Wiley & Sons, Chichester, 1990.

[Gr] P. M. Gruber. In most cases approximation is irregular. *Rend. Sem. Mat. Univers. Politecn. Torino*, **41.1**, 1983.

[K-N] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences*. Wiley, New York, 1974.

[L-T] V. Losert and R. F. Tichy. On uniform distribution of subsequences. *Prob. Th. Rel. Fields*, **72**:517–528, 1986.

[My] G. Myerson. A sampler of recent developments in the distribution of sequences. *Lecture Notes in pure and applied Mathematics*, **147**:163–190, 1993.

[Ox] J. C. Oxtoby. *Measure and Category*. Graduate texts in Mathematics. Springer, 1980.

[R-W] J. M. Rosenblatt and M. Wierdl. Pointwise ergodic theorems via harmonic analysis. In *Ergodic theory and its connections with harmonic analysis*, volume 205 of *London Math. Soc. Lecture Notes*, pages 3–151, Cambridge, 1995. Cambridge University Press. Proceedings of conference in Alexandria, Egypt, May 24–28, 1993.

[Si] K. Sigmund. On minimal centers of attraction and generic points. *J. Reine Angew. Math.*, **295**:72–79, 1977.

[Sa] T. Šalát. A remark on normal numbers. *Rev. Roumaine Math. Pures Appl.*, **11**:53–56, 1966.

[W1] R. Winkler. Strong Weyl property. *Lecture Notes (Number-Theoretic Analysis)*, **1452**:208–220, 1990.

[W2] R. Winkler. On the distribution behavior of sequences. *Mathematische Nachrichten*, **186**:303–312, 1997.

E-mail address: `Martin.Goldstern@tuwien.ac.at`
URL: `http://info.tuwien.ac.at/goldstern/`

E-mail address: `shmeling@math.fu-berlin.de`

E-mail address: `Reinhard.Winkler@oeaw.ac.at`

Department of Algebra, Wiedner Hauptstr. 8-10/118, Technische Universität, A-1040 Wien, Austria, Europe,

Fachbereich Mathematik, Freie Universität Berlin, Arnimallee 3, D-14195 Berlin Germany

Österreichische Akademie der Wissenschaften, Ignaz Seipel-Platz 2, A-1010 Wien, Austria, Europe