Vitaly Bergelson, Ohio State University.

Normal sets in $(\mathbb{N},+)$ and $(\mathbb{N},\times)$.

(the talk is based on joint work with Michał Misiurewicz and Tomasz Downarowicz: "A fresh look at the notion of normality", Annali Scuola Normale Superiore de Pisa, Vol XXI, 2020, pp. 27–88)

Introduction.

Let x be an irrational number in $[0,1]$, let $b\in\mathbb{N}$, $b\geq 2$, and consider the expansion $x=\sum_{i=1}^{\infty}\frac{d_i(x)}{b^i}$.

Identifying x with the sequence $(d_i(x))_{i\in\mathbb{N}}$, let us say that <u>$x$ is normal in base b</u> if any finite word $w=w_1w_2\dots w_k$ with $w_i\in\{0,1,\dots,b-1\}$ appears in the sequence $(d_i(x))_{i\in\mathbb{N}}$ with "correct" frequency.

It was Émile Borel who, in 1909, introduced the notion of normality and (modulo some gaps which were fixed later) prove the following surprising result:

<u>For any fixed $b\in\mathbb{N}$, $b\geq 2$, almost every $x\in[0,1]$ is normal in base b.</u>

Corollary: Almost every $x\in[0,1]$ is base b normal for all $b\geq 2$.

For convenience, we will take $b=2$ and deal with normality of numbers in $[0,1]$ via the binary sequences coming from the expansion $x = \sum_{n=1}^{\infty} \frac{x_n}{2^n}$.

It is also natural to identify elements of $\{0,1\}^{\mathbb{N}}$ with subsets of $\mathbb{N}$: $\forall A \subseteq \mathbb{N}$, $1_A \in \{0,1\}^{\mathbb{N}}$.
(This way, one can talk about normal sets in $\mathbb{N}$).

Given $x = (x_n) \in \{0,1\}^{\mathbb{N}}$ and a finite 0-1 word $w = w_1 w_2 \dots w_k \in \{0,1\}^k$, let $N(w,x,n)$ denote the number of times the word w occurs as a subword of the initial string $x_1 x_2 \dots x_n$ of x :

$$N(w,x,n) := \left| \{ m \in \{1,2,\dots,n-k+1\} : x_m x_{m+1} \dots x_{m+k-1} = w \} \right|$$

<u>Definition</u>. A sequence $x \in \{0,1\}^{\mathbb{N}}$ is <u>normal</u> if $\forall k \in \mathbb{N}$ and every $w \in \{0,1\}^k$ one has

$$\lim_{n \to \infty} \frac{N(w,x,n)}{n} = \frac{1}{2^k} \qquad (*)$$

A sequence $x \in \{0,1\}^{\mathbb{N}}$ is <u>simply normal</u> if formula $(*)$ takes place for any word of length one.

Some equivalent forms of normality:

(i) Number $x \in [0,1]$ is base b normal if and only if it is simply normal in all the bases $b, b^2, b^3, \ldots$ (Pillai, 1940)

(ii) Number $x \in [0,1]$ is base b normal if and only if there exists a set of positive integers $m_1 < m_2 < m_3 < \ldots$ such that x is simply normal to base b^{m_i} for each $i \in \mathbb{N}$ (but no finite set of m's will suffice). (Long, 1957)

(iii) (Borel's original definition; the proof was supplied only in 1951 by Niven and Zuckerman).

A number $x \in [0,1]$ is normal in base b if and only if $\forall k, m \in \mathbb{N}$ the numbers $b^m x \bmod 1$ are simply normal in base b^k.

(iv) Number $x \in [0,1]$ is normal in base b if and only if $(b^n x)_{n \in \mathbb{N}}$ is uniformly distributed mod 1. (Wall, 1949).

Raikov proved in 1936 (without using ergodic theory) that for any $f \in L^1[0,1]$ and any $b \geq 2$, $b \in \mathbb{N}$, one has

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(b^k x \bmod 1) = \int_0^1 f(t)\,dt$$

for a.e. $x \in [0,1]$

In the definition of normality (formula (*)) we used averaging along the intervals $\{1,2,\dots,n\}$.

It is natural to replace the intervals $\{1,2,\dots,n\}$ by a general Følner sequence $(F_n)\subseteq(\mathbb{N},+)$. This leads to the notion of (F_n)-normality (which reduces to the classical normality when $F_n=\{1,2,\dots,n\}$).

Reminder: (F_n) is a Følner sequence in $(\mathbb{N},+)$ if $\forall k\in\mathbb{N}$ one has $\lim\limits_{n\to\infty}\frac{|F_n\cap(F_n-k)|}{|F_n|}=1$

Let $x=(x_n)\in\{0,1\}^{\mathbb{N}}$, let $k\in\mathbb{N}$ and let $w=w_1w_2\dots w_k\in\{0,1\}^k$.

Define:

$$N(w,x,F_n)=|\{m\in\mathbb{N}:\{m,m+1,\dots,m+k-1\}\subset F_n\ \&\ x_m x_{m+1}\dots x_{m+k-1}=w\}|$$

<u>Definition</u>. A sequence $x\in\{0,1\}^{\mathbb{N}}$ is (F_n)-normal if $\forall k\in\mathbb{N}$ and any $w\in\{0,1\}^k$ one has:

$$\lim_{n\to\infty}\frac{N(w,x,F_n)}{|F_n|}=\frac{1}{2^k}$$

Now, let G be a discrete countable cancellative amenable semigroup. By (carefully) imitating the above approach one can introduce, for any Følner sequence $(F_n) \subseteq G$, the notion of (F_n)-normality for elements $x = (x_g)_{g\in G} \in \{0,1\}^G$.

($(F_n) \subseteq G$ is Følner sequence if $\forall g \in G$, $\lim_{n\to\infty} \frac{|F_n \cap g^{-1}F_n|}{|F_n|} = 1$)

Let λ denote the $(\frac{1}{2},\frac{1}{2})$ product measure on $\{0,1\}^G$.

Let σ_g, $g\in G$, denote the "shift" on the symbolic space $\{0,1\}^G$:

for $g \in G$ and $x = (x_h)_{h\in G}$, $\sigma_g(x) := (x_{hg})_{h\in G}$.

<u>Theorem</u>. For any Følner sequence $(F_n) \subseteq G$ with $|F_n| \nearrow \infty$, for any continuous function f on $\{0,1\}^G$ and for λ-almost every $x \in \{0,1\}^G$ we have:

$$\lim_{n\to\infty} \frac{1}{|F_n|} \sum_{g\in F_n} f(\sigma_g x) = \int f \, d\lambda$$

<u>Remark</u>. The general pointwis ergodic theorem for amenable groups (proved by Lindenstrauss in 2001) assumes that the Følner sequence is <u>tempered</u>. In absence of this condition theorem does not work. For example, the sequence of intervals $[n, n+\lfloor\sqrt{n}\rfloor]$ in $(\mathbb{N},+)$ is not tempered and the pointwise ergodic theorem fails along this sequence (Akcoglu & del Junco, 1975). But our situation is rather special: the action is Bernoulli and f is continuous (also $|F_n| \nearrow \infty$)

In total similarity to the classical normality, we have the following theorem:

Theorem. Let G be a countable, cancellative amenable semigroup and let (F_n) be a Følner sequence in G. Then the set of (F_n)-normal elements $x \in \{0,1\}^G$ is of first Baire category in $\{0,1\}^G$.

Corollary. Let G be either $(\mathbb{N},+)$ or $(\mathbb{N},\times)$, and let (F_n) be a Følner sequence in G. Then the set of (F_n)-normal numbers in $[0,1]$ (i.e. numbers which have (F_n)-normal binary expansions) is of first category.

Some notions of largeness in $(\mathbb{N},\times)$.

Here is an example of a Følner sequence in $(\mathbb{N},\times)$:

Let (a_n) be an arbitrary sequence in $\mathbb{N}$.

Let $F_n = \{a_n p_1^{i_1} p_2^{i_2} \cdots p_n^{i_n} : 0 \le i_j \le K_{j,n};\ j=1,2,\dots,n\}$

where $\{p_n\}$ is the set of primes and $\forall j$, $K_{j,n} \xrightarrow[n\to\infty]{} \infty$.

Definition. Let (F_n) be a Følner sequence in $(\mathbb{N},\times)$. Then the upper density of a set $A \subseteq (\mathbb{N},\times)$ w.r. to (F_n) is defined as follows:

$$\bar{d}^{\times}_{(F_n)}(A) = \limsup_{n\to\infty} \frac{|A \cap F_n|}{|F_n|}$$

It is not hard to show that for any $A \subseteq (\mathbb{N},\times)$, any Følner sequence $(F_n) \subseteq (\mathbb{N},\times)$ and any $m \in \mathbb{N}$, $\bar{d}^{\times}_{(F_n)}(A) = \bar{d}^{\times}_{(F_n)}(mA) = \bar{d}_{(F_n)}(A/m)$

A set $S \subseteq G$ is called thick, if it contains a translate of any finite set. In particular:

(i) $S \subseteq (\mathbb{N},+)$ is thick if and only if S contains arbitrarily long intervals

(ii) $S \subseteq (\mathbb{N},\times)$ is thick if and only if S contains arbitrarily large sets of the form $a_n\{1,2,\ldots,n\} = \{a_n, 2a_n, \ldots, na_n\}$.

Remark. Any (F_n)-normal set is thick.

Combinatorial and Diophantine properties of Additively and Multiplicatively normal sets in $\mathbb{N}$.

Convenient abbreviations:

a.n.s. - additively normal set

c.n.s. - "classical" normal set (corresponds to $F_n = \{1,2,\ldots,n\}$ in $(\mathbb{N},+)$)

m.n.s. - multiplicatively normal set.

First (rather cheap) observation: since normal sets are thick, they contain all kinds of configurations which follow from thickness. For example, any a.n.s. contains "finite sums" set $FS(n_i)_{i=1}^{\infty} = \{n_{i_1}+n_{i_2}+\ldots+n_{i_k};\ i_1<i_2<\ldots<i_k;\ k\in\mathbb{N}\}$ and, similarly, any m.n.s. contains sets of the form

$$FP(m_i)_{i=1}^{\infty} = \{m_{i_1} m_{i_2} \ldots m_{i_k};\ i_1<i_2<\ldots<i_k;\ k\in\mathbb{N}\}.$$

Somewhat surprisingly, any m.n.s. contains arbitrarily large sets of the form $FS(n_i)_{i=1}^{k}$ (But not necessarily the infinite finite sums sets $FS(n_i)_{i=1}^{\infty}$).

On the other hand, even if A is a c.n.s. in $(\mathbb{N},+)$, it does not have to contain triples of the form $\{a, b, ab\}$ (A. Fish, 2005).

Still, the following theorems show that c.n.s. do have rather rich structure.

<u>Theorem</u>. Let $A \subset \mathbb{N}$ be a c.n.s. Then, for any Følner sequence (K_n) in $(\mathbb{N},\times)$ there exists a set $E \subset \mathbb{N}$ with

$d^{\times}_{(K_n)}(E) = \frac{1}{2}$, such that for any nonempty finite set $\{n_1, n_2, \dots, n_k\} \subset E$, the set

$A/n_1 \cap A/n_2 \cap \dots \cap A/n_k$ has positive upper density in $(\mathbb{N},+)$.

Invoking some known results about multiplicatively large sets, one has the following corollary.

<u>Theorem</u>. Any "classical" normal set in $(\mathbb{N},+)$ contains, for any $n \in \mathbb{N}$, sets of the form $\{b(a+id)^j ; 0 \le i, j \le n\}$ for some $a, b, d \in \mathbb{N}$.

Some (hard!) open problems.

1. Let $\Omega(n) :=$ number of prime factors of $n \in \mathbb{N}$ (with multiplicities)

Let $E = \{n : \Omega(n) \text{ is even}\}$; $O = \{n : \Omega(n) \text{ is odd}\}$

It is easy to see that for any Følner sequence $(F_n) \subseteq (\mathbb{N}, \times)$

$d^{\times}_{(F_n)}(E) = d^{\times}_{(F_n)}(O) = \frac{1}{2}$.

A much deeper result (which is equivalent to PNT) is this:

$d(E) = d(O) = \frac{1}{2}$.

<u>Conjecture</u> (well known). E (and hence O) is c.n.s. in $(\mathbb{N}, +)$.

2. Let (F_n) be a Følner sequence in $(\mathbb{N}, +)$ and let (K_n) be a Følner sequence in $(\mathbb{N}, \times)$. Assume that $|F_n| \uparrow \infty$ and $|K_n| \uparrow \infty$.

Let λ denote the $(\frac{1}{2}, \frac{1}{2})$ product measure on $\{0,1\}^{\mathbb{N}}$.

We know that λ-almost every $x \in \{0,1\}^{\mathbb{N}}$ is both additively (F_n)-normal and multiplicatively (K_n)-normal.

Can you give a nice Champernowne-like example of such x (for your favorite (F_n) and (K_n))?