Equidistribution Course

Lecturer: Vitaly Bergelson
Scribes: James Leng, Jiazhen Tan, Jet Chung, and Daniel Shapiro

Ross 2020

Contents

Lecture #1

1.1 Uniform Distribution of Sequences

Recall that a subset A of $\mathbb{R}$ is *dense* if every $x \in \mathbb{R}$ is arbitrarily close to A. That is, for every $\epsilon > 0$, every $x \in \mathbb{R}$ is ϵ-close to some element of A. More precisely:

$$\forall x \in \mathbb{R} \text{ and } \forall \epsilon > 0, \text{ there exists } a \in A \text{ such that } |x - a| < \epsilon.$$

Exercise 1.1. *Examples of dense subsets:*

1. *$\mathbb{Q}$ is dense in $\mathbb{R}$.*
2. *$\mathbb{Q}^2 = \mathbb{Q} \times \mathbb{Q}$ is dense in $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$.*

   That is: $\forall \epsilon > 0,\ \forall (x, y) \in \mathbb{R}^2,\ \exists (r, s) \in \mathbb{Q}^2$ such that distance$\big((x, y), (r, s)\big) < \epsilon$.

   Here we use the Euclidean distance.
3. *(optional) Let C be the classical middle-thirds Cantor set. Verify that the set of endpoints of all the removed intervals is a dense subset of C.*

If $x \in \mathbb{R}$ define its *fractional part* to be $\{x\} = x - \lfloor x \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function. Then $\{x\} \in [0, 1)$. We sometimes write (x mod 1) in place of $\{x\}$.

MOTIVATING EXAMPLE.

Proposition 1.2. $\boxed{\textit{If } \alpha \in \mathbb{R} \textit{ is irrational, then the sequence } (n\alpha)_{n\in\mathbb{N}} \textit{ is dense in } [0, 1].}$

How do we define a dense *sequence*? We defined denseness above only for subsets.

Corollary 1.3. *(Easy exercise.) Uncountably many sequences are dense in* $[0, 1]$.

We will prove Proposition 1.2 in several steps. Before starting the proof, we recall some terminology. Given $\epsilon > 0$, a sequence (a_n) in $[0, 1]$ is called ϵ-dense if:

$$\forall x \in [0, 1],\ \exists n \in \mathbb{N} \text{ such that } |a_n - x| < \epsilon.$$

This says: For any subinterval $J \subseteq [0, 1]$ of length 2ϵ, some terms of the sequence (a_n) are in J.

Check that a sequence (a_n) is dense in $[0, 1]$ if it is $\frac{1}{k}$-dense for every $k \in \mathbb{N}$.

To prove the Proposition, we need to show: $\forall k \in \mathbb{N}$ the sequence $\{na\}_{n\in\mathbb{N}}$ is $\frac{1}{k}$-dense.

Let's consider the case $k = 10$.

Claim 1.4. *If there exists m such that $\{m\alpha\} < \frac{1}{10}$, then the sequence $\{n\alpha\}_{n\in\mathbb{N}}$ is $\frac{1}{10}$-dense.*

Proof. Given such m, the sequence $\{m\alpha\}, \{2m\alpha\}, \{3m\alpha\}, \dots$ must enter every subinterval of length $\frac{1}{10}$. □

Claim 1.5. *Among $\{\alpha\}, \{2\alpha\}, \dots, \{11\alpha\}$ in $[0,1]$, there are two terms with distance $< \frac{1}{10}$.*

Proof. Pigeonhole with the subintervals $\left[0, \frac{1}{10}\right], \left[\frac{1}{10}, \frac{2}{10}\right], \cdots, \left[\frac{9}{10}, 1\right]$. □

Proof of Proposition 1.2. By Claim 1.5, there exist $i > j$ such that $\left|\{i\alpha\} - \{j\alpha\}\right| < \frac{1}{10}$. If $\{i\alpha\} > \{j\alpha\}$ in $[0,1]$, then $m = i - j$ satisfies $\{m\alpha\} = \{i\alpha - j\alpha\} = \{i\alpha\} - \{j\alpha\} < \frac{1}{10}$, and Claim 1.4 applies.

What if $\{i\alpha\} < \{j\alpha\}$ in $[0,1]$? Show that $\frac{9}{10} < \{m\alpha\} < 1$. Can we conclude that the sequence is $\frac{1}{10}$-dense in this case? *[Yes. Successive multiples of $\{m\alpha\}$ differ by less than $\frac{1}{10}$, so they also enter every length $\frac{1}{10}$ subintervals.]*

Finish the proof by replacing 10 with arbitrarily large integer k. □

Question 1.6. *Can you generalize the above for $[0, \alpha)$ where $\alpha \neq 1$?*

Question 1.7. *When is the sequence $(\{n\alpha\}, \{n\beta\})_{n\in\mathbb{N}}$ dense in $[0,1] \times [0,1]$?*

(Much easier variant) When is the sequence $\big(\{m\alpha\}, \{n\beta\}\big)_{m,n\in\mathbb{N}}$ dense in $[0,1] \times [0,1]$?

Question 1.8. *If $B \subset A \subset S$ and A is dense in S, B is dense in A, must B be dense in S?*

(Answer: Yes, by the $\epsilon/2$ trick. For any x in S and any $\epsilon > 0$, approximate x in A with error at most $\epsilon/2$. Approximate this approximation again in B with error at most $\epsilon/2$.)

Question 1.9. *Suppose α is irrational and $A \subset \mathbb{N}$ is infinite. Is $\{n\alpha \mid n \in A\}$ necessarily dense in $[0,1]$?*

(Answer: No. See an example of A in Exercise 1.10.)

Exercise 1.10. *Suppose $J = [a, b]$ is a sub-interval of $[0,1]$. Must the set $A = \big\{n : \{n\alpha\} \in J\big\}$ be infinite?*

[Answer: Yes. First show that $\{A\alpha\}$ is dense in J.]

Note: $\{n\alpha \mid n \in A\}$ is not dense in $[0,1]$ for the A in the exercise above.

Question 1.11. *Is $\{n\alpha \mid n \equiv 0 \bmod 17\}$ dense in $[0,1]$?*

(Answer: Yes. Write $n = 17k, k \in \mathbb{N}$. Since 17α is irrational, apply Proposition 1.2 to the irrational $\alpha' = 17\alpha$. Then $S = \{k \cdot (17\alpha) \mid k \in \mathbb{N}\}$ is dense.)

It is natural to ask a general question:

Which integer sequences (a_n) satisfy: For every irrational α, the sequence $(\{a_n\alpha\})_{n\in\mathbb{N}}$ is dense in $[0,1]$?

Question 1.12 (Pico). *Is the set $\left\{n^2\alpha\right\}_{n\in\mathbb{N}}$ dense in $[0,1]$?*

Misha and Aditya ask: What about the sequence of primes? This was settled by a well-known Theorem:

Theorem 1.13 (Vinogradov). *If α is irrational, the sequence $\left\{p_n\alpha\right\}_{n\in\mathbb{N}}$ is dense in $[0,1]$. Here p_n is the n^{th} prime.*

Claim 1.14. *Suppose α is irrational. Then the sequence $\left(n^2\alpha \bmod 1\right)_{n\in\mathbb{N}}$ is dense in $[0,1]$. More generally, if $f(x) \in \mathbf{Z}[x]$ is a nonconstant polynomial, then the sequence $\left(f(n)\alpha \bmod 1\right)_{n\in\mathbb{N}}$ is dense in $[0,1]$.*

The following result is even stronger.

Theorem 1.15 (Weyl)**.** *If $f(t) \in \mathbb{R}[x]$ and at least one coefficient (other than the constant term) is irrational, then the sequence $\left(f(n) \bmod 1\right)_{n\in\mathbb{N}}$ is dense in $[0,1]$.*

Remark. In fact: $\left\{f(p_n)\right\}$ is also dense!

1.2 Next time

Please review Riemann integration of functions $f : [0,1] \to \mathbb{R}$, and review the famous result of Weierstrass:

Theorem 1.16 (Weierstrass Approximation Theorem)**.** *Any continuous function on a closed interval $[a,b]$ can be uniformly approximated by polynomials.*

Next time we will discuss:

Theorem 1.17 (Furstenberg-Sárközy Theorem)**.** *If $S \subset \mathbb{N}$ and no two numbers in S differ by a perfect square, then the asymptotic density of S is zero.*

1.3 Solutions to Exercises

Exercise 1.1.

1. $\mathbb{Q}$ is dense in $\mathbb{R}$.
   As mentioned earlier, if $x \in \mathbb{R}$, truncations of its decimal expression, or its continued fraction approximation, are rational numbers that approximate x to arbitrary precision.

2. This problem only asks about a finite product $\mathbb{Q}^2 = \mathbb{Q} \times \mathbb{Q}$. In fact, this works for arbitrary products.

   Claim 1.18. *If every A_i is dense in B_i, then $\prod\limits_{i\in I} A_i$ is dense in $\prod\limits_{i\in I} B_i$. Here we use the product topology.*

   Proof. Suppose $x \in \prod\limits_{i\in I} B_i$ and $\mathcal{O}$ is an open neighborhood of x. By the definition of the product topology, there exist open $U_i \subseteq B_i$ such that
   $$x \in \prod_{i\in I} U_i \subset \mathcal{O}.$$
   Since U_i is open and A_i is dense, then $U_i \cap A_i \neq \emptyset$, and we may choose $p_i \in U_i \cap A_i$. Let $p \in \prod_{i\in I} B_i$ be the point with i^{th} coordinate p_i. Then $p \in \mathcal{O}$. □

3. if C be the classical middle-thirds Cantor set, the set of endpoints of all the removed intervals is a dense subset of C.

   - Use the base-3 decimal definition, and note that truncations of decimals that have only the digits $0, 2$ still produces a decimal with only digits $0, 2$. Those truncations are the endpoints.
   - From the intersection of nested sets definition of C, if $x \in C$ and $n \in \mathbb{N}$, then in the n^{th} iteration, the segment containing x has length $< \frac{1}{3^n}$. Both endpoints are boundary points of some removed

open intervals. So x is within distance $\frac{1}{3^n}$ of endpoints of removed intervals. For any $\epsilon > 0$, choose n so that $\frac{1}{3^n} < \epsilon$.

Answer. The map $x \mapsto x\alpha$ is a bijection $[0,1] \to [0,\alpha]$, and it preserves the notion of "dense sequence." This allows us to translate statements about $[0,1]$ to statements about $[0,\alpha]$.

Question 1.7. When is the sequence $\big(\{n\alpha\},\{n\beta\}\big)_{n\in\mathbb{N}}$ dense in $[0,1]\times[0,1]$?
(Much easier variant) When is the sequence $\big(\{m\alpha\},\{n\beta\}\big)_{m,n\in\mathbb{N}}$ dense in $[0,1]\times[0,1]$?

The simpler case $(\{m\alpha\},\{n\beta\})$ follows directly from the one variable case: For $\alpha \in \mathbb{R}$, sequence $(\{n\alpha\})_{n\in\mathbb{N}}$ is dense $\iff$ $\alpha \notin \mathbb{Q}$. A stronger result is proved in Theorem 3.18 below.

The sequence $\big(\{n\alpha\},\{n\beta\}\big)_{n\in\mathbb{N}}$ is dense when $1,\alpha,\beta$ are linearly independent over $\mathbb{Q}$. In the answer to Exercise 6.21 below, we analyze when that sequence in uniformly distributed.

Exercise 1.10. If $J = [a,b] \subseteq [0,1]$, must $A = \big\{n : \{n\alpha\} \in J\big\}$ be infinite?

1. *By contradiction.* Suppose the set is finite. Choose $n_0, n_1 \in \mathbb{N}$ so that $\{n_0\alpha\}$ is the smallest element of $A\alpha$ and $\{n_1\alpha\}$ is the second smallest. Since $\{n\alpha | n \in \mathbb{N}\}$ is dense in $[0,1]$, the interval $[\{n_0\alpha\},\{n_1\alpha\}]$ must contain $\{m\alpha\}$ for some $m \in \mathbb{N}$. Contradiction. □

2. *Direct Proof.* By denseness, there is at least one $\{m_1\alpha\} \in [x_1,x_2] := I_1$. Since $x_1 \neq x_2$, $\{m_1\alpha\}$ can't coincide with both endpoints of the interval, and WLOG assume $m_1 \neq x_2$. Take $I_2 = [m_1, x_2]$. Let S be the set of n where this process stops working. WOP on it to show $S = \varnothing$ and you now constructed an infinite sequence $\{m_i\alpha\}_{i=1}^{\infty}$. □

Lecture #2

2.1 Questions from Last Time

Question 2.1 (Lev)**.** *Let $A \in M_{d\times d}(\mathbb{Z})$ be a matrix with integer entries. For a vector $\vec{v} \in [0,1]^d$, its orbit is $\{A^n\vec{v} \bmod 1\}_{n\in\mathbb{N}}$. Are orbits dense?*

Example 2.2. *When $d = 2$ let $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. We will get to this later.*

Note: It could happen that matrix A in the question satisfies that $A^k = \mathrm{Id}$ for some k. Then every orbit is periodic, and no orbit can be dense.

Perhaps we should be asking about orbits that are periodic.

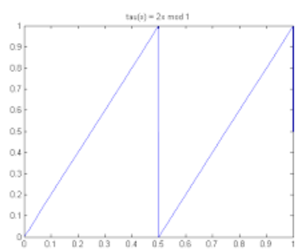
Question 2.3. *For $A, \vec{v}$ as in Question 2.1, when does it happen that the orbit is finite? Define $\vec{v}$ to be a* periodic point *for A if there exists $m > 0$ with $A^m\vec{v} = \vec{v}$.*

Let's work with the one dimensional case first.

Example 2.4. *Suppose $d = 1$ and $A = (2)$, which points are periodic? In other words, for which $x \in [0,1)$ does there exist some $m \in \mathbb{N}$ such that $2^m x \equiv x \pmod 1$?*

Note: (2) is not invertible in the ring of 1×1 integer matrices.

We may investigate this as iterations of the function $f(x) = 2x \bmod 1$ on $[0,1]$.



Question 2.5. *What happens if our matrix A is invertible?*

Question 2.6. *For which x is the sequence $(2^n x \bmod 1)$ dense in $[0,1]$?*

LEV'S SUGGESTION: This should occur if x is normal [in base 2].

Conjecture 2.7 (Pratyush)**.** This should occur if x is irrational.

This turns out to not be true.

Question 2.8. *Is it true that* $(2^n x \bmod 1)_{n\in\mathbb{N}}$ *is dense for uncountably many* x*?*

This one is true. For $x \in [0,1]$, write x in base two: $x = \sum_{i=1}^{\infty} \frac{\beta_i}{2^i} = (0.\beta_1\beta_2\beta_3\ldots)_2$, where $\beta_i \in \{0,1\}$. Then $2x = (\beta_1.\beta_2\beta_3\beta_4...)_2$. Then $(2x \bmod 1) = (0.\beta_2\beta_3\beta_4...)_2$ is just a left-shift. In other words, if we represent a real number by its sequence of its binary digits, then

$$\begin{aligned} x \bmod 1 &\sim (\beta_1, \beta_2, \beta_3, \beta_4, \ldots) \\ 2x \bmod 1 &\sim (\beta_2, \beta_3, \beta_4, \beta_5, \ldots) \end{aligned}$$

and $f(x) = (2x \bmod 1)$ is the left-shift operation.

Question 2.9 (James)**.** *Some numbers don't have a unique decimal expansion. What do we do for those?*

Answer. If a number has two expansions then one is infinite and other finite. To save headaches, always pick the finite representation. Note: This ambiguity occurs for only countably many $x \in [0,1]$. Details are left to the reader.

SRINATH: Is $\{2^n x\}$ dense when the binary expansion of x contains every finite $\{0,1\}$ string?

Answer. Yes!

Theorem 2.10. $(2^n x \bmod 1)$ *is dense in* $[0,1]$ *if and only if every finite length string of* 0*'s and* 1*'s appears in the binary expansion of* x*.*

2.2 Uniform Distribution (U.D.)

Definition 2.11. A sequence $(x_n) \in [0,1]$ is *uniformly distributed* (U.D.) if for any subinterval $[a,b]$ one has

$$\lim_{N\to\infty} \frac{|\{x_n : 1 \le n \le N\} \cap [a,b]|}{N} \to b - a.$$

Here's another description:

Definition 2.12. (x_n) is U.D. if for every subinterval $[a,b]$, the probability that x_n is in $[a,b]$ equals $b-a$.

It's not obvious that such a sequences exist. There are uncountably many intervals $[a,b]$ of a fixed length r and the sequence has to spend the fraction r of its lifetime in each such interval.

Example 2.13. *Consider the sequence*

$$0, 1, 0, \frac{1}{2}, 1, 0, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, 1, \ldots .0, \frac{1}{2^k}, \frac{2}{2^k} \ldots .$$

Our intuition suggests that this is uniformly distributed.

Exercise 2.14 (Mandatory exercise)**.** *Prove: In the definition of U.D., it suffices to verify that limit for intervals* $[a,b]$ *where* a, b *are dyadic rationals. That is,* $a, b \in \left\{ \frac{m}{2^d} \,\middle|\, m, d \in \mathbb{N} \right\}$.

Here's an equivalent definition of uniform distribution. Here $C([0,1])$ is the set of continuous functions $f : [0,1] \to \mathbb{R}$.

Definition 2.15. A sequence (x_n) in $[0,1]$ is uniformly distributed if for every $f \in C([0,1])$:

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} f(x_i) = \int_0^1 f(x)\,dx. \tag{2.1}$$

If we use the indicator function $1_{[a,b]}$ here instead of continuous f, we obtain the other definition of uniform distribution. Here's another equivalent form:

Definition 2.16. Sequence (x_n) in $[0,1]$ is U.D. if (2.1) holds for every Riemann integrable f.

Exercise 2.17. *All three definitions of U.D. mentioned above are equivalent.*

Question 2.18. *If we allow f in Definition 2.15 to be any Lebesgue integrable function, would we have another equivalent version of the definition?*

Answer. NO.

2.3 Weierstrass Approximation Theorem

Theorem 2.19 (polynomial version)**.** *For any $f \in C([0,1])$ and $\epsilon > 0$ there exists a polynomial p such that:*

$$|f(x) - p(x)| < \epsilon$$

for every $x \in [0,1]$.

Theorem 2.20 (trignometric version)**.** *For any $f \in C([0,1])$ and $\epsilon > 0$ there exists a trigonometric polynomial τ with*

$$|f(x) - \tau(x)| < \epsilon,$$

for every $x \in [0,1]$

Here, a trigonometric polynomial is a finite sum of the type $\displaystyle\sum_{k=1}^{n} a_k \sin(kx) + b_k \cos(kx)$.

One final exercise for today.

Exercise 2.21. *Assume that for all $\epsilon > 0$ there exists n such that $(n^2\alpha \bmod 1) < \epsilon$. Derive from this that $(n^2\alpha \bmod 1)$ is dense in $[0,1]$.*

2.4 Preview of Metric Spaces and Cantor Sets.

Metric Spaces. For $p > 1$, and $x, y \in \mathbb{R}^n$, define:

$$\delta_p(x,y) = \sqrt[p]{\sum_{i=1}^{n} |x_i - y_i|^p} \quad \text{for } p \geq 1 \tag{$*$}$$

This is a metric on $\mathbb{R}^n$. (Why does the triangle inequality hold?)

- Check that δ_2 is the usual Euclidean distance.
- When $p = \infty$, $(*)$ is interpreted as $\delta_\infty(x,y) = \max_{1\leq i\leq n}\left\{|x_i - y_i|\right\}$.
  (Explain why this makes sense. Does $\lim\limits_{p\to\infty} \delta_p(x,y) = \delta_\infty(x,y)$,?)

Cantor Sets. The classical Cantor set is homeomorphic to the space $\{0,1\}^{\mathbb{N}}$ under the metric

$$d(\vec{x},\vec{y}) = \sum_{i\in\mathbb{N}} \frac{|x_i - y_i|}{2^i}, \quad \vec{x},\vec{y} \in \{0,1\}^{\mathbb{N}}$$

2.5 Solutions to Exercises

Exercise 2.14. In the definition of U.D., it suffices to verify that limit for intervals $[a,b]$ where a,b are dyadic rationals. That is, $a,b \in \left\{\frac{m}{2^d} \,\middle|\, m,d \in \mathbb{N}\right\}$.

Choose arbitrary $(a,b) \in [0,1]$. For simplicity, let $\epsilon > 0$ with $\epsilon/4 < b-a$. Take dyadic rationals c,d s.t. $c \in (a, a+\epsilon/4)$ and $d \in (b-\epsilon/4, b)$. Then by hypothesis,

$$\lim_{N\to\infty} \frac{|\{x_n : 1 \le n \le N\} \cap (c,d)|}{N} \;\to\; d-c.$$

Now take N so large that the fraction inside the limit is within $\epsilon/2$ of $d-c$ for every $N' \ge N$. Note that

$$\frac{|\{x_n : 1 \le n \le N\} \cap (c,d)|}{n} \le \frac{|\{x_n : 1 \le n \le N\} \cap (a,b)|}{n}$$

Similarly take $a - \frac{\epsilon}{4} < c' < a < b < d' < b + \frac{\epsilon}{4}$ to squeeze the RHS above between $b-a-\epsilon$ and $b-a+\epsilon$ for sufficient large N.

Exercise 2.17. Equivalence of the three definitions of a U.D. sequence in $[0,1]$. Details are left for students to work out.

Lecture #3

3.1 Exploration: U.D. and Denseness

Question 3.1 (Michael and Aditya)**.** *Can we classify functions $f : [0, 1] \to [0, 1]$ such that if $\{a_n\}_{n\in\mathbb{N}}$ is U.D. then $\{f(a_n)\}_{n\in\mathbb{N}}$ is also U.D.?*

KEVIN: Must we have $f' = \pm 1$?

MICHAEL: If you assume f is nice, then $|f'| = 1$ can be shown, so yes.

But what if f is not differentiable?

PICO: How about fractional part of $2x$, why wouldn't that work?

Or the "tent" function: $2x$ on $[0, \frac{1}{2}]$ and $2 - 2x$ on $[\frac{1}{2}, 1]$.

PICO: The requirement that you need isn't very strong, and for Riemann integrable functions all you need is the image of (a, b) to have "measure" $b - a$. Being Riemann integrable is much stronger than we need, but simplifies a lot of the ugly stuff.

Remark. Suppose $f : [0, 1] \to [0, 1]$. If f is surjective and continuous then f preserves denseness. In fact, if $f : [a, b] \to [0, 1]$ is surjective continuous for a subinterval $[a, b]$, then f preserves denseness.

Here is a variation of that question, involving subsequences rather then function values.

Definition 3.2. A sequence $(n_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$ is *universal* if: Whenever (x_n) is uniformly distributed, then (x_{n_k}) is also uniformly distributed.

Question 3.3. *Which sequences $(n_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$ are universal?*

Question 3.4. *Easier version: Which sequences n_k satisfy: if (x_n) is dense, then (x_{n_k}) is dense?*

Proposed Answer. Sequences $(n_k)_{k\in\mathbb{N}}$ that miss at most finitely many elements of $\mathbb{N}$.

Proof. NECESSITY. Suppose $(n_k)_{k\in\mathbb{N}}$ has infinite complement in $\mathbb{N}$, and let $(m_k)_{k\in\mathbb{N}}$ be the complementary sequence. Define $\{x_i\}_{i\in\mathbb{N}}$ by setting:

$$x_{n_k} = 1, \quad \text{and} \quad x_{m_k} = r_k, \quad \forall k \in \mathbb{N}.$$

Here r_k is the k^{th} rational number under some enumeration.

SUFFICIENCY. A dense sequence must have infinitely many elements in any given subinterval $[a, b]$. □

Example 3.5. *If $\alpha \notin \mathbb{Q}$ then $(n\alpha \bmod 1)$ is dense in $[0,1]$, and $(n^2\alpha \bmod 1)$ is also dense.*

Question 3.6. *Which functions preserve denseness?*

Here's a much stronger result:

Theorem 3.7. *If $\alpha \notin \mathbb{Q}$, then both sequences $(n\alpha \bmod 1)$ and $(n^2\alpha \bmod 1)$ are uniformly distributed in $[0,1]$.*

Those two results will be discussed further. For the first, see Theorem 3.18.

Question 3.8. *For which α is the sequence $(2^n\alpha \bmod 1)$ uniformly distributed?*

Review of the fact:

Theorem 3.9. *If every finite 0-1 word occurs in the binary expansion of α, then $(2^n\alpha \bmod 1)$ is dense in $[0,1]$.*

We can generalize this:

Theorem 3.10. *Let k be a positive integer. Then $(k^n\alpha \bmod 1)$ is dense in $[0,1]$ if the base k expansion of α contains all finite words in $\{0,1,2,\dots,k-1\}$.*

But the results above are restricted to integers. We can ask a question about more general rational or real numbers:

Question 3.11. *For what x is $(\pi^n x \bmod 1)$ dense? What about $\left((\frac{3}{2})^n x \bmod 1\right)$?*
(Is this related to a "base π-expansion"? Is there a theory of base β expansions?)

Theorem 3.12 (Open Problem)**.** *Is $\left((\frac{3}{2})^n \bmod 1\right)$ dense in $[0,1]$?*

Exercise 3.13. *Give an example of x such that x^n is dense $\bmod 1$.*

Theorem 3.14 (Furstenberg)**.** *Let α be irrational. Then the set*

$$\{2^n 3^m \alpha \bmod 1 : n, m \in \mathbb{N}\}$$

is dense in $[0,1]$.

Remark. When we deal with denseness, the order of the sequence is immaterial. This is not true for uniform distribution!

For example, $\mathbb{Q} \cap [0,1]$ is dense in $[0,1]$, but it takes work to find an ordering of $]Q$ to produce a U.D. sequence.

Let

$$(a_n)_{n\in\mathbb{N}} := (1,2,3,4,6,8,9,12,16,18,\cdots),$$

the sequence that is a listing of the set $\{2^n 3^m \mid m, n \in \mathbb{N}\}$ in order of size.

Question 3.15 (**B**ig **B**onus **P**roblem.)**.** *For (a_n) defined above, find irrational α such that $(a_n\alpha \bmod 1)$ is not U.D.*

Exercise 3.16. *There are uncountably many irrational α for which $(2^n\alpha \bmod 1)$ is not dense in $[0,1]$.*

3.2 Criterion for Uniform Distribution.

Question 3.17 (Motivating Problem)**.** *Which sequences are uniformly distributed and why?*

Theorem 3.18. *If $\alpha \in \mathbb{R}$ is irrational, then the sequence $(n\alpha)_{n\in\mathbb{N}}$ is uniformly distributed mod 1.*

Proof. Refer to definition 2.15. We will show that for every $f \in C([0,1])$,

$$\frac{1}{n}\sum_{j=0}^{n-1} f(j\alpha) \longrightarrow \int_0^1 f(x)dx.$$

Since f may be approximated by sums of trig polynomials $e^{2\pi ikx}$, it suffices to show this for $f(x) = e^{2\pi ikx}$.

To simplify notations, define $\delta_k = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases}$.

Then

$$\frac{1}{n}\sum_{j=0}^{n-1} e^{2\pi ijk\alpha} = \frac{1}{n}\frac{e^{2\pi ikn\alpha} - 1}{e^{2\pi ik\alpha} - 1} \longrightarrow \delta_k$$

and

$$\int_0^1 e^{2\pi ikx}dx = \delta_k.$$

□

Subtle Point: If f can be approximated by trigonometric polynomials then $f(0) = f(1)$. So the proof above fails when $f(0) \neq f(1)$. How can this be fixed? That fix (left to the reader) is also needed in the next Theorem.

Theorem 3.19 (Weyl's Criterion)**.** *A sequence (x_n) in $[0,1]$ is uniformly distributed mod 1 if and only if for any $h \in \mathbb{Z}$ with $h \neq 0$:*

$$\frac{1}{N}\sum_{n=0}^{N-1} e^{2\pi ihx_n} \longrightarrow 0.$$

Example 3.20. *More examples of U.D. sequences:*

- ($\sqrt{n}$ mod 1).
- ($\log^2 n$ mod 1).
- ($n^2\alpha + \log^2 n$ mod 1).

Conjecture 3.21 (Kevin)**.** If $\lim_{n\to\infty}(a_{n+1} - a_n) \to 0$ and $\lim_{n\to\infty} a_n \to \infty$, must $(a_n)_{n\in\mathbb{N}}$ be U.D.?

Fact 3.22. *The sequence* ($\log n$ mod 1) *is not U.D., but for every $\epsilon > 0$, the sequence* ($\log^{1+\epsilon} n$ mod 1) *is U.D.*

Exercise 3.23. *Kevin's "criterion" does not imply U.D. but is enough for denseness!*

Exercise 3.24 (Pico)**.** *Does the harmonic series not being U.D. (mod 1) follow from* $\log(n)$ *not being U.D. (mod 1)?*

Lecture #4

4.1 Fejér's Theorem.

In connection with Kevin's conjecture 3.21, here is a new conjecture:

Conjecture 4.1. Assume $(a_n) \subset \mathbb{R}$ such that $a_n \to \infty$ monotonically, $a_{n+1} - a_n \to 0$ monotonically, and $n(a_{n+1} - a_n) \to \infty$. Then $(a_n \bmod 1)$ is uniformly distributed in $[0, 1]$.

This is almost true! See [2] Theorem 1.3.4, page 29. Also compare this with the following classical theorem due to Lipót Fejér. (See [3], problem 174, p. 90.)

Theorem 4.2 (Fejér). *Let $f : [0, \infty) \to [0, \infty)$ be a continuously differentiable function such that as $x \to \infty$ we have the limits: $f(x) \to \infty$ monotonically, $f'(x) \to 0$ monotonically, and $xf'(x) \to \infty$. Then $\big(f(n) \bmod 1\big)$ is uniformly distributed in $[0, 1]$.*

Functions satisfying those conditions are called *Fejér functions*.

Question 4.3 (Aditya). *Are there interesting functions $f : [0, 1] \to [0, 1]$ with $(f^n(x_0) \bmod 1)$ uniformly distributed for some x_0? What if we also require f to be continuous?*

Example 4.4. *$f(x) = 2x$ mod 1. This is not continuous, but can replace it by a "tent function." Alternatively, if we use the circle $\mathbb{T}$ instead of $[0, 1]$, then the discontinuity at $\frac{1}{2}$ disappears.*

Example 4.5. *If $0 < c < 1$ is $(n^c \bmod 1)_{n\in\mathbb{N}}$ U.D. (mod 1)? If $c > 1$ is $(\log(n)^c \bmod 1)$ U.D. (mod 1)?*

Example 4.6. *Is $\log(n) \log\log(n)$ dense (mod 1)?*

Exercise 4.7. *How about $(n^c \bmod 1)$ for $c > 1$. What if $c \in \mathbb{N}$?*
How about $(n^c \log^b(n) \bmod 1)$? For which parameters b, c is this sequence dense? Uniformly distributed?

Definition 4.8. A number $x \in [0, 1]$ is *base 2 normal* if every finite 0-1 word w appears in the binary expansion of x with probability $\dfrac{1}{2^{|w|}}$. Here $|w|$ is the length of word w.

Theorem 4.9. *Let $x \in [0, 1]$. Then $(2^n x \bmod 1)$ is uniformly distributed if and only if x is base 2 normal.*

Theorem 4.10. *The set of base 2 normal numbers in $[0, 1]$ has full measure. In other words, the set of non-normal numbers has measure 0. Similarly for base k normal numbers, for any integer base $k > 1$*

Corollary 4.11. *Almost all numbers in $[0, 1]$ are normal in every base.*

Example 4.12. *Champernowne's constant: Concatenation of all integers written in base ten produces a decimal:*

$0.12345678910111213\ldots$ *This is a base* 10 *normal number. Concatenation of squares also produces a base* 10 *normal number:* $0.1491625\ldots$ *So does concatenation of primes:* $0.2357111317\ldots$

There are rather general theorems of the type: if $f : \mathbb{N} \to \mathbb{N}$ is "nice", then the decimal $0.f(1)f(2)f(3)\ldots$ is base 10 normal.

Question 4.13. *Let p, q be positive with $p + q = 1$. Define p-q normality for a binary sequence to mean:* 0 *appears with probability p and* 1 *appears with probability q. Provide a more precise definition.*

Proposition 4.14. *There is a natural bijection between the collection of subsets of $\mathbb{N}$ and $\{0, 1\}^{\mathbb{N}}$.*
This allows us to discuss when a binary sequence (elements of $\{0, 1\}^{\mathbb{N}}$) is normal.

Proof. Left to the reader. □

A subset S of $\mathbb{N}$ is *normal* if its indicator function 1_S is a normal binary sequence.

Definition 4.15. The density $d(S)$ of a subset $S \subset \mathbb{N}$ is:

$$d(S) = \lim_{N\to\infty} \frac{|S \cap \{1, 2, \ldots, N\}|}{N} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} 1_S(n).$$

Are there subsets for which that limit does not exist and the density is not defined?

Exercise 4.16. *Suppose S is normal.*

- *$N \setminus S$ is also normal and both S and $N \setminus S$ have density $\frac{1}{2}$.*
- *For every $n \neq 0$: $d\big(S \cap (S - n)\big) = \frac{1}{4}$*
- *For distinct nonzero $n_1, n_2, \ldots, n_k$ in $\mathbb{N}$, $d\Big(S \cap (S - n_1) \cap (S - n_2) \cap \cdots \cap (S - n_k)\Big) = \frac{1}{2^{k+1}}$.*

Exercise 4.17. *In the equation above, if we replace some of the sets with their complements, the formula is still valid.*

Example 4.18. *Example of U.D. sequences based on Fejèr's theorem.*

1. $(\sqrt{n} \bmod 1)$. *More generally,* $(n^c \bmod 1)$ *for* $0 < c < 1$.
2. $(\log^c n \bmod 1)$, *for* $c > 1$.
3. *Candidates to consider:*
   - $\log n \cdot \log\log n$
   - $\log n \cdot \log\log n \cdot \log\log\log n$
   - $\log n \cdot \log^c \log n$
   - $\log n \cdot \log\log^c \log n$
   - *How about* $(n^c \bmod 1)$, *when* $c \notin \mathbb{N}$ *and* $c > 0$?
   - *How about* $n^c \log^b n$ *for parameter* $c \notin \mathbb{N}, b > 1$. *For which* b, c *is the sequence dense? U.D.?*

Exercise 4.19. $(\sin n \bmod 1)$ *is dense. Is that sequence U.D.?*

Lecture #5

5.1 Uniform Distribution, Normal Numbers, and van der Corput's Trick

Exercise 5.1 (In response to Pico)**.** *After appropriately defining U.D.* mod 2 *(your exercise), show that* $\sqrt{n}$ *is U.D.* mod 2*.*

How about U.D. mod $\sqrt{2}$*?*

Is $n\sqrt{3}$ *or* $\frac{n}{\sqrt{3}}$ *U.D.* mod $\sqrt{2}$*?* mod 2*?* mod 3*?*

Conjecture 5.2 (Michael and Aditya)**.** Every dense sequence has a U.D. subsequence.

Theorem 5.3. *Any dense sequence has a U.D. rearrangement.*

BERGELSON: This was proved in the 20th century, by a famous guy who was youngish at the time but is now dead. Exercise: Guess who it is. [Answer. See [4].]

Proof outline. Choose a sequence $(x_n)_{n\in\mathbb{N}}$ that is U.D. in $[0,1]$. Suppose Y is a countable dense subset of $[0,1]$. For every n choose $y_n \in Y$ such that $|x_n - y_n| < \frac{1}{n}$. Show that $(y_n)_{n\in\mathbb{N}}$ is U.D. in $[0,1]$.

This sequence (y_n) might not include all elements of Y. To finish the proof we want to insert more terms into (y_n) while preserving the U.D. property. This can be done as follows:

Exercise. Given sequences (y_n) and (c_n), create a new sequence as follows: Choose a zero-density subset $D \subset \mathbb{N}$ and let E be the complement of D. Create a new sequence (z_n) by writing $y_1, y_2, \ldots$ at the indices in E, and write $c_1, c_2, \ldots$ at the indices in D. If (y_n) is U.D., show that (z_n) is also U.D. □

Definition 5.4. We will say that a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ is U.D. if $(x_n \bmod 1)$ is U.D.

When convenient, we identify the unit interval with the 1 dimensional torus $\mathbb{T} = [0,1)$. (Note that $\mathbb{T}$ a circle of circumference 1.)

More generally, the n-dimensional torus is $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$.

Recall our equivalent definitions for a sequence (x_n) to be U.D.:

1. The frequency that (x_n) hits $[a,b]$ is proportional to the length of $[a,b]$, as $n \to \infty$.
2. The average value of any continuous $f \in C[0,1]$ on the first n terms of the sequence converges to its integral on $[0,1]$, as we take $n \to \infty$.

3. As $n \to \infty$, the average value of a Riemann integrable function on the first n terms of the sequence converges to its integral on $[0,1]$.

We now add another equivalent statement, called *Weyl's Criterion*:

4. For any nonzero integer h,
$$\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i h x_n} \to 0.$$

5. The criterion using continuous functions is still valid if we restrict to the subset of all $f \in C[0,1]$ such that $f(0) = f(1)$.

Here are statements of two theorems that we will not prove in these lectures.

Theorem 5.5 (Weyl)**.** *For any sequence (n_k) in $\mathbb{N}$ that goes to infinity, the set of $x \in \mathbb{R}$ for which $(n_k x)$ is uniformly distributed has full measure.*

Theorem 5.6 (Borel)**.** *A sequence $(2^n x \bmod 1)$ is uniformly distributed if and only if x is base 2 normal.*

A number is *absolutely normal* if it is base b normal for every integer $b > 1$. These two theorems imply that the set of absolutely normal numbers has full measure. Both theorems involve properties of "measure 0" rather than the full theory of measure.

For future reference, we mention several versions of "convergence" of a sequence of functions $f_n : [0,1] \to \mathbb{R}$.

Definition 5.7. (f_n) is *almost everywhere convergent* if $\lim_{n\to\infty} f_n(x)$ exists for a set of x having full measure in $[0,1]$.

Definition 5.8. (f_n) *converges uniformly* to f if for every $\epsilon > 0$, there exists N_0 such that for all $n \geq N_0$,
$$\max_{x\in[0,1]} \left|f_n(x) - f(x)\right| < \epsilon.$$

Definition 5.9. (f_n) is L^1-convergent to f if: $\displaystyle\lim_{n\to\infty} \int_0^1 \left|f_n(x) - f(x)\right| dx = 0.$

Exercise 5.10. *No two of those definitions are equivalent.*

In order to prove that the sequence $(n^2\alpha)$ is uniformly distributed, we state the van der Corput trick:

Lemma 5.11 (vdC Trick)**.** *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. For any $h \in \mathbb{N}$, suppose that the sequence $\big((x_{n+h} - x_n) \bmod 1\big)_{n\in\mathbb{N}}$ is uniformly distributed. Then $(x_n)_{n\in\mathbb{N}}$ is uniformly distributed mod 1.*

Proof. See [2] Theorem 1.3.1. □

Now suppose α is irrational, and let $x_n = n^2\alpha$. Then $x_{n+h} - x_n = h^2\alpha + 2nh\alpha$. That sequence (mod 1) is uniformly distributed because it is a shift of $n\alpha$ by a constant. Then the vdC Trick implies that $(n^2\alpha \bmod 1)$ is U.D. Consequently, the set $\{(n^2\alpha \bmod 1) : n \in \mathbb{N}\}$ is dense in $[0,1]$.

Definition 5.12. A set $E \subset \mathbb{N}$ is called a *van der Corput set* if in order to apply van der Corput's Trick you need to check the criterion only for $h \in E$.

Exercise 5.13. *Show that every vdC set is infinite.*

Observation (Misha): For $k \in \mathbb{N}$, the set $k\mathbf{Z}$ is a vdC set. This holds because the subsequences sorted by remainder $\bmod k$ are each U.D. and merging U.D. sets produces a U.D. set.

Example 5.14. *Here are some surprising examples of van der Corput sets:*
$\{n^2\}$, $\{17n\}$, $\cancel{n^2+1}$, n^2-1, $P-1$, $P+1$, $\cancel{P+17}$, $\cancel{P-17}$, *where P is the set of primes.*
The only prime shifts that are vdC sets are $P-1$ and $P+1$. This seems mysterious.

Theorem 5.15. *For any sequence (a_n) in $\mathbb{N}$ that goes to infinity, the set of differences $\{a_n - a_m : n > m\}$ is a van der Corput set.*

Question 5.16 (Jessie)**.** *Can we define a sense of U.D.* $\bmod 1$ *for f in $C[0,1]$?*

PICO: We can start by considering ~~all~~ rational coefficient polynomials on $[0,1]$. We don't want uncountably many things.

Definition 5.17. How would you define (and provide interesting examples) a notion of uniform distribution for $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$?

Exercise 5.18. *Is there a version of van der Corput for denseness?*

Exercise 5.19. *Prove Weyl's theorem on uniform distribution of polynomials by using the vdC trick.*

Lecture #6

6.1 Uniform distribution in higher dimensions.

Question 6.1 (Kevin's Question)**.** *For which vectors $x \in \mathbb{R}^d$ is $(nx \bmod 1)$ dense in $[0,1]^d$ or in $\mathbb{T}^d$?*

Question 6.2. *If x is normal in base 2, is it true that $2x$ is normal in base 2?*

Answer: Yes, by Borel's Theorem.

Exercise 6.3. *The set of α for which $(2^n\alpha \bmod 1)$ is not dense in $[0,1]$ is uncountable.*

Suggestion: For a given binary word w, consider the set all sequences that do not include w as a subword. Show that this set is uncountable.

Exercise 6.4. *The interval $[0,1]$ does not have measure 0.*

Note. The concept of "measure zero" is more elementary than the definition of Lebesgue measure. In this exercise we hope for a direct proof from the definition of "measure zero." Don't just quote a theorem from measure theory.

Proposition 6.5. *If a set E has full measure, then it is uncountable.*

Exercise 6.6. *Given that normal numbers exist, prove that there are uncountably many normal numbers.*

Exercise 6.7. *If $c \notin \mathbb{N}$ prove that $(n^c \bmod 1)$ is uniformly distributed in $[0,1]$.*

Hint: For $c \in (1,2)$, consider the function $x \mapsto (x+h)^c - x^c$. This is a Fejér function (see Theorem (4.2).).

Exercise 6.8. *For $c \notin \mathbb{N}$, can you find a more elementary proof that $(n^c \bmod 1)$ is dense? (i.e. An argument that does not first prove the sequence is U.D.)*

Question 6.9. *How do you define uniformly distributed in $[0,1]^2$?*

Here's an idea due to Misha:
Sequence $(x_n)_{n\in\mathbb{N}}$ in $[0,1]^2$ is U.D. if: for any rectangle $J = [a,b] \times [c,d] \subseteq [0,1]^2$ we have:

$$\lim_{N\to\infty} \frac{\left|\{n : 1 \leq n \leq N \text{ and } x_n \in J\}\right|}{N} = area(J).$$

Note that $area(J) = (b-a)(d-c)$. We may also restrict to cases when a, b, c, d are dyadic rationals.

Another version is:

Sequence $(x_n)_{n\in\mathbb{N}}$ in $[0,1]^2$ is U.D. if: for every continuous (or Riemann integrable) function f on $[0,1]^2$,

$$\frac{1}{N}\sum_{i=1}^{N} f(x_i) \longrightarrow \int_{[0,1]^2} f\, dA.$$

Theorem 6.10 (Weyl)**.** *Sequence (x_n) in $[0,1]^2$ is uniformly distributed if and only if for every $h \in \mathbb{Z}^2$ with $h \neq (0,0)$:*

$$\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i(x_n\cdot h)} \to 0.$$

Notation: If $x_n = (y_n, z_n)$ and $h = (h_1, h_2)$, then $x_n \cdot h = y_n h_1 + z_n h_2$ is the dot product.

Exercise 6.11. *Show that all the above versions of the definition of uniform distribution are equivalent.*

Exercise 6.12. *For the set $(\mathbb{Q}\times\mathbb{Q}) \cap [0,1]^2$, suggest an ordering that produces a sequence uniformly distributed in $[0,1]^2$. More generally, prove: Any dense sequence in $[0,1]^2$ can be rearranged to become uniformly distributed.*

Question 6.13. *For $\alpha \notin \mathbb{Q}$, is $(n\alpha, 2n\alpha)$ uniformly distributed mod 1?*

No. In fact, $(n\alpha, 2n\alpha)$ is not even dense mod 1.

Exercise 6.14. *For which (α,β) is the orbit $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ dense in $\mathbb{T}^2$? U.D.?*

Exercise 6.15 (subexercise)**.** *Find a closed form for $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^n$. [Hint: Compute the first few powers of $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.]*

What if we take a general 2×2 matrix with integer entries?

Question 6.16 (Zhuowen)**.** *Suppose α and β are irrational.*
(i) Is the sequence $(n\alpha, n^2\beta)_{n\in\mathbb{N}}$ mod 1 *U.D. in $[0,1]^2$?*

(ii) Is the sequence $(p_n\alpha, p_n^2\beta)$ mod 1 *dense in $[0,1]^2$? Is it U.D.?*
Here $(p_n)_{n\in\mathbb{N}}$ is the sequence of prime numbers.

(Answer: Those sequences are U.D., but the proof is difficult.)

Exercise 6.17 (Shapiro, to Misha)**.** *Prove that the set of diagonalizable $n\times n$ matrices over $\mathbb{C}$ is dense in the set of all $n\times n$ matrices.*

Exercise 6.18. *Suppose f is a Fejér function, as in Theorem (4.2). Is $\big(f(n)\sqrt{2},\ f(n)\sqrt{3}\big)$ dense mod 1?*

Question 6.19. *Suppose $a, b \in \mathbb{R}$ and $ab \in \mathbb{N}$. Does there exist α such that $(a^n\alpha, b^n\alpha)$ is uniformly distributed in $\mathbb{T}^2$?*

Exercise 6.20. *What about $(a^n\alpha, b^n\beta)$ in $\mathbb{T}^2$ for some β?*

Exercise 6.21. *For which α, β is $(n\alpha, n\beta)_{n\in\mathbb{N}}$ U.D. in $\mathbb{T}^2$?*

Answer to Exercise 6.21. A necessary and sufficient condition is: $1, \alpha, \beta$ are linearly independent over $\mathbb{Q}$. That is, $r\alpha + s\beta$ is irrational for any nonzero $r, s \in \mathbb{Q}$.

Proof. By the two dimensional version of Weyl's criterion, the sequence $(n\alpha, n\beta)$ is U.D. in $\mathbb{T}^2$ if and only if for every $h \in \mathbb{Z}^2$ with $h \neq (0,0)$:

$$\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i (n\alpha, n\beta)\cdot h} \longrightarrow 0 \tag{*}$$

With $\gamma = (\alpha, \beta)\cdot h = h_1\alpha + h_2\beta$, the sum in (*) equals $\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i n\gamma}$.

Then our sequence is U.D. in $\mathbb{T}^2$ if and only if for every $h_1, h_2 \in \mathbf{Z}$, not both zero, and $\gamma = h_1\alpha + h_2\beta$ we have: $\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i n\gamma} \longrightarrow 0$. The classical Weyl criterion implies that the limit is correct if and only if γ is irrational. Then, the necessary and sufficient condition is that no nonzero integer combination of α, β is rational. Equivalently, $1, \alpha, \beta$ are linearly independent over $\mathbb{Q}$. $\square$

Lecture #7

7.1 Topological Dynamical Systems

Conjecture 7.1 (Kevin)**.** Suppose $(x_n), (y_n)$ are sequences in $\mathbb{R}$ and for every $a, b \in \mathbf{Z}$ with $(a, b) \neq (0, 0)$ suppose that $(ax_n + by_n \bmod 1)_{n\in\mathbb{N}}$ is U.D. in $[0, 1]$. Then (x_n, y_n) must be U.D. in $[0, 1]^2$.

Question 7.2 (Pico)**.** *Assume that we have sequences $(x_n), (y_n), (z_n)$ that are pairwise U.D. in the squares. Then must (x_n, y_n, z_n) be U.D. in $[0, 1]^3$?*

KEVIN: What if $x_n = \sqrt{2}n, y_n = \sqrt{3}n, z_n = (1 - \sqrt{2} - \sqrt{3})n$?

By comparison, if A, B, C are subsets of a probability space that are pairwise independent, e.g.. $P(A \cap B) = P(A) \cdot P(B)$. Then does it follow that

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)\ ?$$

Answer. NO.

Fact 7.3. *Two facts that are harder to prove:*

- $(n \log n \bmod 1)$ *is U.D. in* $[0, 1]$.
- $(\frac{n}{\log n} \bmod 1)$ *is U.D. in* $[0, 1]$.

Exercise 7.4. *Prove that* $(\log n)_{n\geq 2}$ *is not U.D. mod 1.*

Juxtapose this with the fact that for every $\epsilon > 0$*, the sequence* $(\log^{1+\epsilon} n)_{n\geq 2}$ *is U.D. mod* 1.

Consider the transformation $(x, y) \mapsto (x + \alpha, y + 2x + \alpha)$.
Check that $T(0, 0) = (\alpha, \alpha),\ T^2(0, 0) = (2\alpha, 4\alpha),\ T^3(0, 0) = (3\alpha, 9\alpha), \ldots$ Show that for any $n \in \mathbb{N}$:

$$T^n(0, 0) = (n\alpha, n^2\alpha)$$

This observation can lead us to a proof that $(n^2\alpha \bmod 1)$ is U.D.

Definition 7.5. Suppose (X, d) be a compact metric space and $T : X \to X$ a homeomorphism (a bijection such that both T and T^{-1} are continuous). Then the triple (X, T, d) is called a *topological dynamical system*.

We are interested in the iterations T^n acting on space X.

Definition 7.6. For $x \in X$, the *orbit* of x is the sequence $(T^n(x))_{n\in\mathbb{N}}$. Sometimes we use negative powers as well, and say that the orbit is $(T^n(x))_{n\in\mathbf{Z}}$, Here are typical questions asked about such systems:

Question 7.7. *Is there $x \in X$ such that the orbit of x is dense in X?*

Conjecture 7.8 (Aditya)**.** Let (X, T, d) be a topological dynamical system. Assume T^n is never the identity map, for $n \in \mathbb{N}$. Then most points $x \in X$ have dense orbits.

Answer by Misha. Consider $X = \mathbb{T}^2$, the two dimensional torus, and T be a horizontal shift by some constant. There are no dense orbits. Each orbit traces out a circle on the torus. □

Suggestion by Jonathan Lin. Consider the map $T : \mathbb{T} \to \mathbb{T}$ defined by $T(x) = 2x$.

Almost every x has a dense orbit by Borel's Theorem.

Similar to Misha's suggestion, but on unit disk $\mathcal{D}$. Consider the map $T : \mathcal{D} \to \mathcal{D}$ defined by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Every orbit lies on a circle. No orbit is dense.

One more simple example. $X = [0, 1]$, $T(x) = x^2$. Do there exist points with dense orbits?

Is it possible that some (X, T) is so nice that every point has a dense orbit? YES!

Definition 7.9. A topological dynamical system (X, T, d) is *minimal* if every orbit is dense.

Question 7.10. *Consider the map $T : (x, y) \mapsto (x + \alpha, y + mx + \beta)$ on $X = [0, 1]^2$. For which m, α, β is this system minimal?*

Question 7.11. *Show that*

$$(x, y, z) \mapsto (x + \alpha, y + 2x + \alpha, z + 3y + 3x + \alpha)$$

is minimal.

Exercise 7.12. *Given polynomial $p(n)$, can you suggest $T : \mathbb{T}^d \to \mathbb{T}^d$ such that $T^n(0, \ldots, 0) = \big(\cdots, p(n)\big)$? (As a first step, try to get $\alpha n + \beta n^2$).*

Definition 7.13. Systems like the ones in (7.11) and (7.12) are known as **skew products**.

Definition 7.14. A topological dynamical system (X, T, d) is called **topologically transitive** if there exists $x_0 \in X$ such that $T^n x_0$ is dense in X.

Question 7.15. *Is it true that in any nontrivial transitive system, there are many points of dense orbit?*

Example 7.16. *Let T be the map $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ on the torus $\mathbb{T}^2$. Do uncountably many points have dense orbits?*

Exercise 7.17. *Describe all periodic points for this map. (A point is periodic if its orbit is finite).*

Lecture #8

8.1 Probability Measures and Cantor Sets

Definition 8.1. A measure space is a pair $(X, \mathcal{B}, \mu)$.

- X is a set, typically $[0,1]$ or $[0,1]^2$.
- $\mathcal{B}$ is a family of "nice" subsets of X.
- $\mu : \mathcal{B} \to \mathbb{R} \cup \{\infty\}$ is a function that is "σ-additive":

$$\text{If sets } B_i \in \mathcal{B} \text{ are pairwise disjoint, then } \quad \mu\left(\bigcup_{i=0}^{\infty} B_i\right) = \sum_{i=0}^{\infty} \mu(B_i).$$

We will usually deal with *probability* measures. Those are measures normalized as follows:

- $\mu(X) = 1$.

Measures on $X = [0,1]^d$ that we are interested in will usually have the property that the measure of every rectangular box in X coincides with its natural d-dimensional Euclidean volume.

Here's a reminder of a previous notion that is more elementary:

Definition 8.2. A set in $\mathbb{R}^d$ has measure 0 if it can be covered with boxes $(I_i)_{i=1}^{\infty}$ such that the sum of the volumes of the I_i is arbitrarily small.

This definition implies: Every countable set has measure zero.

BERGELSON: Is there a uncountable subset of $[0,1]$ of measure zero? Prove that your example has measure zero without using the notion of a general measure.

MISHA: The classical Cantor set. At step n, the remaining intervals have total length $(\frac{2}{3})^n$.

Exercise 8.3. *Suppose set $A \subset X$ has measure 0 in the sense of (8.2). Does it follow that $\mu(A) = 0$, for any measure space $(\mathbb{R}^d, \mathcal{B}, \mu)$ as defined in (8.1)?*
(If all boxes are in $\mathcal{B}$ and μ matches the Euclidean volume on boxes, then does it follow that $\mu(A) = 0$?)

Example 8.4. *Find an uncountable set $E \subset [0,1]$ such that $\mu(E) > 0$ but E contains no intervals.*

1. *(TRIVIAL) The complement of $\mathbb{Q}$ in $[0,1]$.*
2. *(EASY USING PREVIOUS CONTENT) The set of base 2 normal numbers.*

What if we want $\mu(E) = \frac{1}{2}$?

1. *(Aditya)* Trivial: $(\mathbb{R} \setminus \mathbb{Q}) \cap [0, \frac{1}{2}]$ *union with* $\mathbb{Q} \cap (\frac{1}{2}, 1]$
2. *(Jet) A Cantor-like set built by removing intervals of shorter total length. Say, a middle fifths Cantor set.*

Exercise 8.5. *What is the measure of the Cantor set that removes intervals of length $\frac{1}{5^n}$ in the n^{th} iteration?*

Proof. Find the total length of the removed intervals, and subtract this from 1. As the removed intervals are disjoint, their total length is

$$\frac{1}{5} + \frac{2}{25} + \frac{4}{125} + \cdots = \frac{\frac{1}{5}}{1 - \frac{2}{5}} = \frac{1}{5-2} = \frac{1}{3}.$$

Thus, the measure of the middle-fifths Cantor set is $\frac{2}{3}$. □

Remark. The middle-fourths Cantor set is an uncountable $E \subset [0, 1]$ such that $\mu(E) = \frac{1}{2}$ and E contains no intervals.

A more general Cantor-like construction: Suppose (a_n) is a sequence with $a_n > 0$ for every n and $\sum_{n \in \mathbb{N}} a_n < 1$. Starting with $[0, 1]$, at the n^{th} step, remove the middle interval of length $\frac{a_n}{2^n}$ from each of the 2^n remaining intervals. What remains is a Cantor-like set C_a of positive measure!

Exercise 8.6. *Prove C_a (the generalized Cantor set) is homeomorphic to the classical Cantor set.*
Sub-exercise: C_a is uncountable.

Proof. To see that C_a is uncountable, note that as we are removing intervals of total length less than 1, what remains has positive measure. Since any countable set has measure zero, C_a must be uncountable. To find a homeomorphism, match the infinite binary address strings over the Cantor set to infinite binary address strings over C_a. □

Remark. The exercise above shows that homeomorphisms can map a set of measure zero to one of positive measure. Therefore, the property that a subset has positive measure is not stable under homeomorphisms.

[James.] When people were first investigating fat Cantor sets, I think the original goal was to find an example of a set whose indicator function is Riemann integrable and is homeomorphic to a set whose indicator function is not Riemann integrable.

Free Thinking.

- Both C and C_a are topologically small.

Exercise 8.7. *Is there a dense G_δ set in $[0, 1]$ which has measure zero?*

Definition 8.8 (Continuity of measure)**.** Does μ satisfy the following property?

$$\text{If } A_n \to A \text{ then } \mu(A_n) \to \mu(A).$$

That limit of sets means: $\mu(A_n \bigtriangleup A) \to 0$, where $\bigtriangleup$ denotes the symmetric difference.

8.2 Borel Sets.

We describe a recursive way to construct a collection of "nice" subsets of $[0,1]$.

Step 1. Countable or finite disjoint unions of intervals and their complements.

Step 2. Countable or finite disjoint unions of sets from Step 1 and their complements.

Step 3. Repeat forever. Then repeat further.

The *Borel σ-algebra $\mathcal{B}$* is the collection of sets obtained by that process.

Exercise 8.9. *Prove that $\mathcal{B}$ is countably generated, i.e. there is a countable family of subintervals in $[0,1]$ which "produces" all the sets in $\mathcal{B}$ by performing Steps 1, 2, 3, $\cdots$.*

Hint. *Consider the dyadic intervals.*

Exercise 8.10. *Give an example of a set obtained in Step 2 that is new (i.e. was not created in step 1).*

Question 8.11. *Does every subset of the classical Cantor set C belong to $\mathcal{B}$?*

NO! The cardinality of $\mathcal{B}$ equals the cardinality of $\mathbb{R}$. The collection of subsets of C has larger cardinality.

Definition 8.12 (Completion of measure)**.** From now on, we declare all subsets of measure zero sets to have measure zero.

8.3 Measure Preserving Transformations.

Definition 8.13. Let $(X, \mathcal{B}, \mu)$ be a probability measure space. (That is a measure space with $\mu(X) = 1$.) Suppose $T : X \to X$.

T is *measurable* if for every $A \in \mathcal{B}$, the preimage $T^{-1}(A)$ is in $\mathcal{B}$.

T is *measure preserving* if T is measurable and for every $A \in \mathcal{B}$, $\mu\big(T^{-1}(A)\big) = \mu(A)$.

Free Thinking. Why do we prefer preimage to image in the definition above?

Exercise 8.14. *Show that the following maps are measure preserving. Generalize as much as you can (say to $\mathbb{T}^d$).*

(1) $T : x \mapsto x + \alpha$ *on* $\mathbb{T}$.

(2) $T : x \mapsto 3x$ *on* $\mathbb{T}$. *This is a 3-to-1 map.*

(3) $T : (x, y) \mapsto (x + \alpha, y + \beta)$ *on* $\mathbb{T}^2$.

(4) $T : (x, y) \mapsto (x + \alpha, y + x)$ *on* $\mathbb{T}^2$. *Also* $T(x, y) \mapsto (x + \alpha, y + 5x + \beta)$.

(5) $T : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. *Also* $T : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 5 & 19 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$.

Fact 8.15. *To check whether a map is measure preserving, it's enough to check it on a generating set of $\mathcal{B}$.*

Theorem 8.16 (A version of Ergodic Theorem)**.** *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system. Then for any nice function $f : X \to \mathbb{R}$, there is a function $f^* : X \to \mathbb{R}$ such that:*

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \xrightarrow{a.e.} f^*(x) \quad \textit{and} \;\; f^*(Tx) = f^*(x) \textit{ for a.e. } x.$$

If T is ergodic (defined below) then for almost every $x \in X$:

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \longrightarrow \int f d\mu.$$

This says that the "time averages" converge to the "space average."

Lecture #9

Question 9.1 (Kevin)**.** ***Bonus exercise!*** *Let $X = [0,1]$ and let us assume that*

1. *$T : [0,1] \longrightarrow [0,1]$ is continuous.*
2. *$\exists x_0 \in [0,1]$ such that $(T^n x_0)$ is U.D.*

Then must T be (Lebesgue) measure-preserving?

Remark. Examples when $X = [0,1]$ and μ = Lebesgue measure:

- $T : x \longrightarrow x + a$ (Identify $[0,1]$ with $\mathbb{T}$ by gluing endpoints.)
- For $k \in \mathbb{N}$, let $Tx = (kx \bmod 1)$.
- (New kind of examples) Interval Exchange Transformations (IET).

Suggestion: Consider $Tx = 4x(x-1)$.

9.1 Leading Digits of 2-powers.

Problem 9.2 (Famous Problem. You can solve it!)**.** *In the sequence of 2-powers:* $2, 4, 8, 16, 32, 64, \ldots$*, list the first (leading) digits:* $1, 2, 4, 8, 1, 3, 6, 1, 2, 5, \ldots$*. Does the digit 7 appear more frequently than the digit 8?*

What does "frequency" mean here? Let a_n be the first digit of 2^n. If $k \in \{1, 2, \ldots, 9\}$, we are looking at fractions:

$$V_7 = \lim_{N\to\infty} \frac{\left|\{n : 1 \le n \le N \text{ and } a_n = 7\}\right|}{N}$$

$$V_8 = \lim_{N\to\infty} \frac{\left|\{n : 1 \le n \le N \text{ and } a_n = 8\}\right|}{N}$$

PICO: To have a leading 7, you need to be in between $\log_{10}(7)$ and $\log_{10}(8)$.
Since $\log_{10}(2)$ is irrational, $n \cdot \log_{10}(2)$ is U.D., so $V_7 > V_8$.

MISHA: Do we know that those limits exist?

(Answer: Yes, the limits exist. This follows from uniform distribution.)

Exercise 9.3.

(a) *Prove that $\log_{10} 2$ is irrational.*

(b) *Use logarithms to find explicit expressions for V_k, for $k = 1, 2, \cdots, 9$.*

(c) *Formulate a similar question about the first digit sequence of a^n in base b.*

(d) *Does 12345 appear as the leading five digits of some 2^n? If so, what's the frequency?*

(e) *Suppose a_n is the first digit of 2^n and b_n is the first digit of 3^n. What can you say about the behavior of the pairs (a_n, b_n)?*

9.2 Gauss Map and Continued Fractions.

Example 9.4. *Define the Gauss map $T : [0,1] \to [0,1]$ by:* $T(x) = \begin{cases} \{1/x\} & x \neq 0, \\ 0 & x = 0. \end{cases}$

This is not Lebesgue-measure preserving but it does preserve an absolutely continuous measure. See below.

Theorem 9.5 (Ergodic Theorem for the Gauss Map)**.** *For almost every x in $[0,1]$:*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f\big(T^n(x)\big) \;=\; \frac{1}{\log(2)} \int_0^1 \frac{f(x)}{1+x} dx.$$

This leads us to consider the measure μ defined by:

$$\mu([a,b]) = \frac{1}{\log(2)} \int_a^b \frac{dx}{1+x}.$$

Exercise 9.6. *Prove that the Gauss map preserves the measure μ defined above.*

Fact 9.7. *Any number $x \in (0,1)$ has a unique simple continued fraction expansion, defined recursively by:*

$$\theta_1 = \frac{1}{x}, \text{ and for every } n \geq 1: \; a_n = \lfloor \theta_n \rfloor, \quad \theta_{n+1} = \frac{1}{\theta_n - a_n}.$$

It turns out that the sequence (a_n) is eventually periodic if and only if f x is a quadratic irrational.

Well, the continued fraction expansion is unique when x is irrational. For $x \in \mathbb{Q}$ the sequence (a_n) is finite and has two versions. For instance

$$\frac{4}{3} = 1 + \frac{1}{3} = 1 + \frac{1}{2+\frac{1}{1}}.$$

The continued fraction process provides a bijection between irrational $x \in (0,1)$ and infinite sequences (a_n) of positive integers. We sometimes write $a_n(x)$ to emphasize their dependence on x.

The following result about continued fraction entries is really amazing.

Proposition 9.8. *For almost every $x \in [0,1]$:*

$$\frac{a_1(x) + a_2(x) + \cdots + a_n(x)}{n} \longrightarrow \infty$$

$$\big(a_1(x)a_2(x)\cdots a_n(x)\big)^{1/n} \longrightarrow k.$$

Here k is a finite constant independent of x.

That number $k \approx 2.685452\ldots$ is called Khinchin's (or Хинчин's) constant.

Each $x \in [0,1]$ has a binary expansion in $\{0,1\}^{\mathbb{N}}$. The map $x \mapsto (2x \bmod 1)$ corresponds to the shift map on binary sequences. Analogously, each $x \in (0,1]$ has a continued fraction expansion $(a_1(x), a_2(x), a_3(x), \ldots) \in \mathbb{N}^{\mathbb{N}}$. In that continued fraction space, the Gauss map corresponds to the shift .

Exercise 9.9. *Prove the statement above.*

Lecture #10

10.1 Ergodic Transformations.

Definition 10.1. In a measure space $(X, \mathcal{B}, \mu)$, two sets $A, B \subseteq X$ are *equal almost everywhere* if their symmetric difference $A \,\Delta\, B$ has measure zero. That is, $\mu(A \,\Delta\, B) = 0$.

Definition 10.2. Suppose $(X, \mathcal{B}, \mu, T)$ is a dynamical system. That system is *ergodic* if for $A \in \mathcal{B}$,

$$T^{-1}(A) = A \text{ almost everywhere implies } \mu(A) = 0 \text{ or } \mu(A) = 1.$$

In other words, an ergodic system admits only "trivial" T-invariant subsets.

Theorem 10.3 (Birkhoff's Pointwise Ergodic Theorem, 1931)**.** *If $(X, \mathcal{B}, \mu, T)$ is an ergodic system then for almost every $x \in X$ and for every "nice" $f : X \to \mathbb{R}$, one has almost everywhere:*

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \longrightarrow \int_X f\, d\mu$$

- *Here "nice" might mean $f \in L^1(X)$, $L^2(X)$, $L^\infty(X)$, etc.*

For comparison, a sequence $(x_n) \in [0,1]$ is U.D. if for all $f \in \mathcal{R}[0,1]$ (Riemann-integrable),

$$\frac{1}{N}\sum_{n=1}^{N} f(x_n) \to \int_0^1 f dx$$

Question 10.4 (Kevin)**.** *Is there a minimal homeomorphism $T : X \to X$ when $X = [0,1] \times \mathbb{T}$?*

Answer. NO! This is non-trivial and we omit details.

Question 10.5.

1. *How about the analogue of the previous question, when X is the infinite cylinder $\mathbb{R} \times \mathbb{T}$?*
2. *Generally, which subsets of $\mathbb{R}^d$ "support" a minimal homeomorphism?*

Theorem 10.6. *A "typical" pair of rotations of the sphere generates a free group. This typical action is minimal.*

- *Is the orbit of a typical point U.D. ? (Of course we should first define the notion of uniform distribution on a sphere.)*

10.2 Ergodic Theorems.

Theorem 10.7 (von Neumann, 1932)**.** *Suppose $(X, \mathcal{B}, \mu, T)$ is an ergodic dynamical system. For all $f \in L^1(X)$:*

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} f\big(T^n(x)\big) = \int_X f\,d\mu, \qquad \textit{in the } L^1 \textit{norm.}$$

Theorem 10.8. *Suppose $(X, \mathcal{B}, \mu, T)$ is an ergodic dynamical system. The following are equivalent:*

1. *(von Neumann's Ergodic Theorem) For all $f \in L^1(X)$:*

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} f\big(T^n(x)\big) = \int_X f\,d\mu, \qquad \textit{in the } L^1 \textit{norm.}$$

2. *(Birkhoff's Pointwise Ergodic Theorem) For all $f \in L^1(X)$ and for almost all $x \in X$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f\big(T^n(x)\big) = \int_X f\,d\mu.$$

3. *For all $A, B \in \mathcal{B}$,* $\quad \dfrac{1}{N} \displaystyle\sum_{n=0}^{N-1} \mu(A \cap T^{-n}B) \longrightarrow \mu(A)\mu(B)$

4. *For all $A \in \mathcal{B}$,* $\quad \dfrac{1}{N} \displaystyle\sum_{n=0}^{N-1} \mu(A \cap T^{-n}A) \longrightarrow \mu(A)^2.$

Remark. [Ergodic Principle] In statistical mechanics, those limits are interpreted as:

$$\boxed{\text{Time Average} = \text{Space Average.}}$$

Remark. Here's an idea to help think of ergodicity:

If $\mu(A \cap T^{-n}B) = \mu(A)\mu(T^{-n}B) = \mu(A)\mu(B)$, then the images of B are independent from A. The formula in the fourth bullet above can be interpreted as asymptotic independence of images of B from A, *on average*.

Here is an outline of a proof of item 3 in Theorem 10.8.

Substitute 1_A for f in Birkhoff's Theorem above to get

$$\frac{1}{N} \sum_{n=0}^{N-1} 1_A\big(T^n(x)\big) \xrightarrow{a.e} \int 1_A(x)\,d\mu.$$

Now multiply by 1_B:

$$\frac{1}{N} \sum_{n=0}^{N-1} 1_{T^{-n}A}(x) 1_B(x) \to 1_B(x) \int 1_A(x)\,dx.$$

Integrate both sides and compute the integrals, to find:

$$\frac{1}{N} \sum_{n=0}^{N-1} \mu(T^{-n}A \cap B) \to \mu(A)\mu(B). \qquad \square$$

Claim 10.9. *If* $\frac{1}{N}\sum_{n=0}^{N-1}\mu(A \cap T^{-n}A) \to \mu(A)^2$ *for every* $A \in \mathcal{B}$*, then* T *is ergodic.*

Proof. Assume that $T^{-1}A = A$ a.e. Then every $T^{-n}A = A$ a.e. and therefore $\mu(A \cap T^{-n}A) = \mu(A)$. Then
$\mu(A)^2 = \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(A\cap T^{-n}A) = \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(A) = \mu(A)$. Therefore $\mu(A) = 0$ or 1. □

10.3 Hilbert Spaces.

A Hilbert space is an infinite dimensional version of our friend $\mathbb{R}^n$ (or $\mathbb{C}^n$).

Definition 10.10. A *Hilbert space* $\mathcal{H}$ is a vector space (over $\mathbb{R}$ or $\mathbb{C}$) with a positive definite inner product. We write $\langle x, y\rangle$ for the inner product of $x, y \in \mathcal{H}$.

Example 10.11. *The following are Hilbert spaces.*

1. $l^2 = \left\{(x_1, x_2, \cdots), x_i \in \mathbb{R} : \sum |x_i|^2 < \infty\right\}$. *Alternatively, we may use* l^2 *over* $\mathbb{C}$.
2. $L^2([0,1],\lambda)$*, the set of square integrable functions w.r.t Lebesgue measure* λ.
3. $L^2([0,1],\mu)$ *where* μ *is the Gauss measure:* $\mu(A) = \frac{1}{\log 2}\int_A \frac{dx}{1+x}$.
4. $L^2([0,1]^d,\lambda)$ *where* λ *is the Lebesgue measure on* $[0,1]^d$.

A Hilbert space is called *separable* if it has a countable orthonormal basis.

Fact 10.12 (Surprise!)**.** *All separable Hilbert spaces are isomorphic! For instance,* $L^2\left([0,1]^2\right) \cong l^2$.

10.4 Unitary operators.

Definition 10.13. For a Hilbert space $\mathcal{H}$, an invertible linear operator $U : \mathcal{H} \to \mathcal{H}$ is called *unitary* if it preserves the inner product. That is, if:

$$\langle Ux, Uy\rangle = \langle x, y\rangle \quad \text{for every } x, y \in \mathcal{H}.$$

Check that this is equivalent to requiring: $\|Ux\| = \|x\|$ for every $x \in \mathcal{H}$.

Exercise 10.14. *Let* $T : X \to X$ *be an invertible measure preserving transformation on probability space* $(X, \mathcal{B}, \mu)$. *For* $f \in L^2(X,\mathcal{B},\mu)$*, define the function* Uf *by* $(Uf)(x) = f(Tx)$*. Prove that U is a unitary operator on* $L^2(X,\mathcal{B},\mu)$.

A key point: That operator U is linear, even if f is not linear. This follows quickly from definitions:

$$U(f+g) = (f+g)\circ T(x) = f\circ T(x) + g\circ T(x) = Uf + Ug, \quad U(c\cdot f) = c\cdot Uf$$

This definition also gives us a sense that a U-orbit of an L^2 function is closely related to a T-orbit of a point. This follows because $f(T^n x) = (U^n f)(x)$, for every n. (Verify that formula!) Then:

the T-orbit of x is $(T^n x)_{n\in\mathbb{N}}$ while the U-orbit of f is $(U^n f)_{n\in\mathbb{N}}$. That is:
$(U^n f)(x) = f(T^n x)$.

Definition 10.15. For $\mathcal{H} = L^2(X, \mathcal{B}, \mu)$ as above, define $\mathcal{H}_{inv}$ to be the subspace of U-invariant elements. That is,

$$\mathcal{H}_{inv} = \{f \in \mathcal{H} : Uf = f\}.$$

Let $P_{inv} : \mathcal{H} \to \mathcal{H}_{inv}$ be the orthogonal projection to that subspace.

As above, suppose T is a bijective measure preserving map on X, and define operator $U : \mathcal{H} \to \mathcal{H}$ by $f \mapsto f \circ T$.

Theorem 10.16 (Abstract form of von Neumann's Ergodic Theorem.)**.** *For $(X, \mathcal{B}, \mu, T)$, $\mathcal{H}$ and U as above, then:*

$$\frac{1}{N}\sum_{n=0}^{N-1} U^n f \longrightarrow Pf \quad \forall f \in \mathcal{H}.$$

We will pursue these idea next time.

Lecture #11

11.1 von Neumann's Ergodic Theorem and Uniform Distribution.

Both of these themes deal with the paradigm: Time Average = Space Average.

Example 11.1. *Suppose $\lambda \in \mathbb{C}$ with $|\lambda| = 1$. What is the limit of* $\displaystyle\frac{1}{N}\sum_{n=1}^{N}\lambda^n$ *?*

$$\frac{1}{N}\sum_{n=1}^{N}\lambda^n \quad \xrightarrow{N\to\infty} \quad \begin{cases} 0, & \lambda \neq 1 \\ 1, & \lambda = 1 \end{cases} \tag{11.1}$$

Here is a useful variation of the limit in (11.1):

$$\frac{1}{N-M}\sum_{n=M+1}^{N}\lambda^n \xrightarrow{N-M\to\infty} \begin{cases} 0, & \lambda \neq 1 \\ 1, & \lambda = 1 \end{cases} \tag{11.2}$$

Instead of taking averages over long initial segments in that series, we take averages of long "windows" of consecutive terms in the series.

The following result was proved by John von Neumann in October 1931 and published January 1932. G. D. Birkhoff learned about this theorem from von Neumann and hurried to publish a related result in December 1931. Further discussion of this von Neumann-Birkhoff conflict appears in [1], Section 1.

Theorem 11.2. *(von Neumann's Ergodic Theorem) Let $U : \mathcal{H} \to \mathcal{H}$ be a unitary operator as above. Then for every $f \in \mathcal{H}$:*

$$\lim_{N-M\to\infty} \frac{1}{N-M}\sum_{n=M}^{N-1} U^n f = P_{inv} f.$$

Here P_{inv} is the orthogonal projection onto the subspace $\mathcal{H}_{inv} = \{g \in \mathcal{H} \mid Ug = g\}$.

11.2 Unitary Operators and Measure Preserving Transformations

Let $\mathcal{H}$ be an L^2-space and suppose $\varphi \in \mathcal{H}$ with $|\varphi| = 1$ a.e. Define the multiplication operator $M_\varphi : \mathcal{H} \to \mathcal{H}$ by $M_\varphi(f) = \varphi f$. Then M_φ is a unitary operator on $\mathcal{H}$.

Theorem 11.3 (A version of Spectral Theorem for unitary operators.). *If $(\mathcal{H}, U)$ is a Hilbert space with unitary operator, then it has a "model" $\big(L^2(X,\mu), M_\varphi\big)$ for some measure space $(X, \mathcal{B}, \mu)$ and some measurable $\varphi \in L^2(X,\mu)$ with $|\varphi(x)| = 1$ a.e.*

Example 11.4. *For $\alpha \in \mathbb{T}$, define $\tau_\alpha : \mathbb{T} \to \mathbb{T}$ by $x \mapsto x + \alpha$. Define an operator U on $L^2(\mathbb{T})$ by $Uf(x) = f(\tau_\alpha x)$. Check that this operator is unitary.*

Remark. This operator has the following "multiplicative property": for every (say, bounded) $f, g \in L^2$,

$$U\big(f(x) \cdot g(x)\big) = Uf \cdot Ug.$$

The operator m_φ is not usually multiplicative:

$$M_\varphi(f \cdot g) = \varphi \cdot f \cdot g \quad \text{does not always equal} \quad M_\varphi(f) M_\varphi(g) = \varphi \cdot f \cdot \varphi \cdot g.$$

Definition 11.5. The pairs $(\mathcal{H}_1, U_1)$, $(\mathcal{H}_2, U_2)$ are *isomorphic* if there exists a norm-preserving bijection $\sigma : \mathcal{H}_1 \to \mathcal{H}_2$ such that the following diagram is commutative:

$$\begin{tikzcd}
\mathcal{H}_1 \arrow[r, "\sigma"] \arrow[d, "U_1"] & \mathcal{H}_2 \arrow[d, "U_2"] \\
\mathcal{H}_1 \arrow[r, "\sigma"] & \mathcal{H}_2
\end{tikzcd}$$

The meaning of "commutative" here is: $U_2 \circ \sigma = \sigma \circ U_1$. That is, $\sigma\big(U_1(f)\big) = U_2\big(\sigma(f)\big)$ for every $f \in \mathcal{H}_1$.

That commutative property says that U_1 and U_1 are "conjugate": $\boxed{U_1 = \sigma^{-1} \circ U_2 \circ \sigma}$.

11.3 Well-Distribution (W.D.)

A refined form of uniform distribution (U.D.) is well-distribution (W.D.).

Definition 11.6. A sequence $(x_n) \in [0,1]$ is W.D. if for every subinterval $(a,b) \subseteq [0,1]$ one has

$$\lim_{N-M \to \infty} \frac{\big|\{n : M \le n < N \text{ and } x_n \in (a,b)\}\big|}{N - M} \;=\; b - a.$$

Remark. Not every U.D. sequence is W.D.

Exercise 11.7. *(i) Let $x_n = (\sqrt{n} \bmod 1)$. Show that (x_n) is U.D. but is not W.D.*

(ii) For $c > 0, c \notin \mathbb{N}$, the sequence $(n^c \bmod 1)_{n\in\mathbb{N}}$ is U.D. in $[0,1]$, but is not W.D.

Any polynomial with at least one irrational coefficient other than the constant term is $W.D$. This result follows from the following fact:

A W.D. version of Van der Corput's trick: For every $h \in \mathbb{N}$, assume that $(x_{n+h} - x_n)_{n\in\mathbb{N}}$ is W.D. mod 1. Then $(x_n)_{n\in\mathbb{N}}$ is W.D. mod 1.

Exercise 11.8. *(i) (The base case for an induction) Prove: If $\alpha \notin \mathbb{Q}$ then $(n\alpha \bmod 1)_{n\in\mathbb{N}}$ is W.D.*

(ii) Generalize to polynomials.

11.4 Proof of von Neumann's Theorem

Proof. First, note that for a unitary operator $U : \mathcal{H} \to \mathcal{H}$, we have a (rather trivial) orthogonal splitting:

$$\mathcal{H} = \mathcal{H}_{inv} \oplus \mathcal{H}_{erg}, \quad \text{where } \mathcal{H}_{inv} = \{g \in \mathcal{H} \mid Ug = g\}, \text{ and } \mathcal{H}_{erg} = \mathcal{H}_{inv}^{\perp}.$$

Then any $f \in \mathcal{H}$ is expressed as $f = f_{inv} + f_{erg}$, and:

$$U^n f = U^n(f_{inv} + f_{erg}) = U^n f_{inv} + U^n f_{erg} = f_{inv} + U^n f_{erg}$$

Note that $f_{inv} = P_{inv}(\mathcal{H})$, the projection.

Using an isomorphism between this Hilbert space to a "model" $\left(L^2(X,\mu), M_\varphi\right)$ provided by the Spectral Theorem above, we get:

$$\frac{1}{N-M}\sum_{n=M}^{N-1} U^n f \;=\; \frac{1}{N-M}\sum_{n=M}^{N-1} M_\varphi^n f \;=\; \left(\frac{1}{N-M}\sum_{n=M}^{N-1}\varphi^n\right)\cdot f$$

Because $\varphi \in L^2(X,\mu)$ and $|\varphi(x)| = 1$ a.e., the well-distribution of the sequence $\varphi^n(x)$ on $\mathbb{T}$ when $\varphi(x) \neq 1$ implies:

$$\frac{1}{N-M}\sum_{n=M}^{N-1} U^n f(x) \;=\; f^*(x) \;=\; \begin{cases} 1 & \text{when } \varphi(x) = 1 \\ 0 & \text{when } \varphi(x) \neq 1 \end{cases}$$

□

Exercise 11.9. *Check that f^* is the projection of f onto M_φ-invariant functions. In particular, if we took $f \in \mathcal{H}_{inv}$, then $f^* = 0$.*

11.4.1 Application

Let (X, B, μ) be a measure space with μ-preserving map $T : X \to X$. Assume T is ergodic.

Exercise 11.10. *Assume $f \in L^2$ and $f(Tx) = f(x)$ a.e. Prove: f = constant a.e.*

Theorem 11.11 (Exercise: Application of von Neumann's theorem. Compare item 3 in Theorem 10.8.)**.**

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

(Hint: Let $\mathcal{H} = L^2(X, \mathcal{B}, \mu), Uf = f(Tx)$, and apply von Neumann to $f = \mathbf{1}_A$.)

11.5 Unique Ergodicity

Definition 11.12. If that $T : X \to X$ then T is *uniquely ergodic* if there exists a unique T-invariant measure on X.

Fact 11.13. *If $\alpha \notin \mathbb{Q}$ then $T : x \mapsto x + \alpha$ on the circle $\mathbb{T}$ is uniquely ergodic.*

Exercise 11.14. *$T : x \mapsto 2x$ on $\mathbb{T}$ is not uniquely ergodic. In fact, there are many, many, many T-invariant measures.*

Answer to Exercise 11.14. Both δ_0, the point mass at 0 and the Lebesgue measure μ are T-invariant, and so is any of their linear combinations $a\delta_0+(1-a)\mu$. Actually there are many many many more T-invariant measures. ◻

Theorem 11.15. *For a compact metric space X, if $(X, \mathcal{B}, \mu, T)$ is uniquely ergodic, then for any continuous function $f : X \to \mathbb{R}$ (or $f : X \to \mathbb{C}$) and for every $x \in X$:*

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} f(T^n x) = \int_X f\, d\mu$$

Exercise 11.16. *Examples of uniquely ergodic transformations:*

- *$(x_1, \dots, x_d) \mapsto (x_1 + \alpha_1, \dots, x_d + \alpha_d)$ on $\mathbb{T}^d$ where α_i are $\mathbb{Q}$-independent.*
- *$(x, y) \mapsto (x + \alpha, y + 2x + \alpha)$ on $\mathbb{T}^2$, for $\alpha \notin \mathbb{Q}$. Observe that $T^n(0, 0) = (n\alpha, n^2\alpha)$.*

  Suppose $f(x, y) = F(y) \in C(\mathbb{T}, \mathbb{R})$ depends only on the second coordinate. Then when $(x, y) = (0, 0)$ unique ergodicity gives us

$$\frac{1}{N-M} \sum_{n=M}^{N-1} T^n f(x, y) = \frac{1}{N-M} \sum_{n=M}^{N-1} f(n\alpha, n^2\alpha) = \frac{1}{N-M} \sum_{n=M}^{N-1} F(n^2\alpha) \to \int_{\mathbb{T}} F(y)dy.$$

  The last convergence for every F shows that $(n^2\alpha \bmod 1)$ is W.D. in $[0, 1]$.

For a general $f \in C(\mathbb{T}^2, \mathbb{R})$, the convergence

$$\frac{1}{N-M} \sum_{n=M}^{N-1} f(n\alpha, n^2\alpha) \to \int_{\mathbb{T}} \int_{\mathbb{T}} f(x, y)\, dx dy$$

shows $(n\alpha, n^2\alpha)$ is W.D. in $\mathbb{T}^2$.

Exercise 11.17. *Define a notion of W.D. for $\mathbb{T}^2$ (or $\mathbb{T}^d$).*

Lecture #12

Question 12.1 (Kevin)**.** *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T_1, T_2 : X \to X$ be invertible μ-preserving transformations, and suppose $A, B, C \in \mathcal{B}$. Consider*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(A \cap T_1^{-n} B \cap T_2^{-n} C). \tag{12.1}$$

(i) When does (12.1) converge?
(ii) If it does converge, must its limit equal $\mu(A)\mu(B)\mu(C)$?

Answer. (i) This works for any $(X, \mathcal{B}, \mu)$ and any A, B, C, provided T_1 and T_2 commute.
(ii) Require that T_1, T_2 commute and also $T_1 \times T_2$ on $X \times X$ is ergodic, and $T_1 T_2^{-1}$ is ergodic.

Special case of interest is when $T_2 = T_1^2$. In that case, we are asking when

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(A \cap T^{-n} B \cap T^{-2n} C) \;=\; \mu(A)\mu(B)\mu(C). \tag{12.2}$$

Answer. The limit in (12.2) holds if and only if T is "weakly mixing," or equivalently that $T \times T$ acting on $X \times X$ is ergodic.

Fact 12.2. *From basic linear algebra: Any finite dimensional unitary transformation of $\mathbb{C}^n$ is diagonalizable. That is, if $U : \mathbb{C}^n \to \mathbb{C}^n$ is unitary then there exists an orthonormal basis consisting of eigenvectors of U. In other words, there exists a transformation ($n \times n$ matrix) T such that $T^{-1}UT$ is a diagonal matrix of the type $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ where $|\lambda_k| = 1$ for every k.*

The matrix theorem above is an elementary version of the Spectral Theorem. $U : \mathcal{H} \to \mathcal{H}$ has a model φ such that $Uf = \varphi f$ and $|\varphi| = 1$ a.e. For an eigenvector x we know $Ux = \lambda x$ and $|\lambda| = 1$. Then $\|\lambda x\| = \|x\|$.

12.1 Sárközy Theorem

Definition 12.3. For a set $E \subset \mathbb{N}$, define the *upper density* $\overline{d}(E)$ as follows:

$$\overline{d}(E) := \limsup_{N\to\infty} \frac{|E \cap \{1, 2, \ldots, N\}|}{N}$$

When the limit above actually exists (not just limsup), we write $d(E)$, the *natural density* of E.

Exercise 12.4. *For $\alpha > 1$ let $E_\alpha = \{\lfloor n\alpha \rfloor : n \in \mathbb{N}\}$. Show that $d(E_\alpha)$ exists and equals $\frac{1}{\alpha}$.*

Question 12.5. *Do there exist $E_1, E_2 \subset \mathbb{N}$ such that $E_1 \cap E_2 = \varnothing$ but $\overline{d}(E_1) = \overline{d}(E_2) = 1$?*

Answer. This can be done by constructing "intermittent sets", like

$$E_1 = \mathbf{Z} \cap \left(\bigcup_{n=0}^{\infty} [10^{(2n)!}, 10^{(2n+1)!}) \right), \quad E_2 = \mathbf{Z} \cap \left(\bigcup_{n=0}^{\infty} [10^{(2n+1)!}, 10^{(2n+2)!}) \right).$$

Question 12.6. *Can we do a similar construction with n sets E_i? With infinitely many E_i?*

Theorem 12.7 (Sárközy's Theorem)**.** *For any $E \subset \mathbb{N}$ with $\overline{d}(E) > 0$, there exist $x, y \in E$ and $n \in \mathbb{N}$ such that $x - y = n^2$.*

This Theorem is difficult to prove. Below, we point out results that will lead to a proof.

It is interesting to ask whether other expressions might work here, instead of n^2. Let us say that an integer-valued polynomial $g(x)$ is *Sárközy-good* if the Theorem holds true with $g(n)$ in place of n^2.

Investigating which polynomials work provide some intriguing patterns. Here some examples:

$$\boxed{n^k}^{\text{OK}}, \cancel{n^2+1}, \boxed{n^2-1}^{\text{OK}}, \cancel{n^2+17}, \boxed{(n^2-2)(n^2-17)(n^2-34)}^{\text{OK}}$$

Proposition 12.8. *An integer-valued polynomial $g(x)$ is Sárközy-good if and only if it is* intersective. *That means, for every $m \in \mathbb{N}$, $g(\mathbf{Z}) \cap m\mathbf{Z} \neq \varnothing$. Equivalently: $g(x)$ has a root mod m, for every $m \in \mathbb{N}$.*

Remark. By the Chinese Remainder Theorem, to check whether $g(x) \in \mathbf{Z}[x]$ is intersective, it suffices to check that it has a root (mod p^m) for every prime power p^m.
If p is prime and $g(x)$ has a simple root (mod p), then Hensel's Lemma implies that $g(x)$ has a root (mod p^m) for every m.

Exercise 12.9. *Let $h(x) = (x^2 - 2)(x^2 - 17)(x^2 - 34)$.*

(a) If p is odd, show that $h(x)$ has a simple root (mod p).
[Hint: If p is prime and $a, b \in \mathbf{Z}$ then at least one of a, b, ab is a square (mod p).]

(b) $h(x)$ has roots (mod 2) but they are not simple.
Prove that $x^2 - 17$ has a root (mod 2^m) for every m.

(c) Deduce that $h(x)$ is intersective.

Some other sets (not generated by polynomials) will also work for Sárközy's Theorem. For example, letting p run over the set of primes, we obtain a similar list:

$$\cancel{p}, \boxed{p-1}^{\text{OK}}, \boxed{p+1}^{\text{OK}}, \cancel{p+17}, \cancel{p-17}, \boxed{p^2-1}^{\text{OK}}$$

Theorem 12.10 (Furstenberg's "Ergodic Sárközy Theorem")**.** *For any measure preserving system $(X, T, \mu, \mathcal{B})$ and and $A \in \mathcal{B}$ with $\mu(A) > 0$, there exists an integer n such that $\mu\big(T^{-n^2} A \cap A\big) > 0$.*

Surprisingly, Theorem 12.10 is equivalent to Sárközy's theorem!

Corollary 12.11. *For almost every $x \in A$, there exists $n \in \mathbb{N}$ such that $T^{n^2} x \in A$.*

Assume that the system $(X, \mathcal{B}, \mu, T)$ is *totally ergodic* (meaning that T^n is ergodic for every $n \in \mathbb{N}$).

Then for every $A, B \in \mathcal{B}$,

$$\frac{1}{N-M} \sum_{n=M}^{N-1} \mu(A \cap T^{-n^2} B) \xrightarrow{N-M\to\infty} \mu(A)\mu(B).$$

In physics, one deals with measure-preserving flows $(T_t)_{t\in\mathbb{R}}$. If $\big(X, \mathcal{B}, \mu, (T_t)_{t\in\mathbb{R}}\big)$ is an ergodic $\mathbb{R}$-system, then for all but countably many $t_0 \in \mathbb{R}$, T_{t_0} is a totally ergodic transformation. For inspiring reading, look at the treatment of these ideas in [1].

Several notions:

$$T \text{ is ergodic means: } f \circ T = f \text{ a.e. } \Rightarrow f = \text{constant a.e.}$$

$$T \text{ is weakly mixing means: } \text{ if } \exists\, f \text{ and } \lambda \text{ such that } f \circ T = \lambda f \text{ a.e. then } f = \text{constant a.e.}$$

Defining $Uf = f(Tx)$, then weakly mixing means: $Uf = \lambda f$ a.e. $\Longrightarrow$ f = constant a.e.

Theorem 12.12 (Hilbertian version of van der Corput trick)**.** *Assume that (x_n) is a bounded sequence in $\mathcal{H}$. Also assume that for every $h \in \mathbb{N}$: $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_{n+h}, x_n\rangle = 0$. Then* $\left\| \frac{1}{N} \sum_{n=1}^{N} x_n \right\|_{\mathcal{H}} \to 0.$

In particular, take $x_n = U^n f \in L^2$ and consider $\int f(T^{n+h}x) f(T^n x)$.

This implies the classical van der Corput trick!

Here's an application of the Hilbertian van der Corput trick:

Theorem 12.13 (von Neumann)**.** *Suppose U is an ergodic operator on a Hilbert space $\mathcal{H}$. Then*

$$\frac{1}{N} \sum_{n=1}^{N} U^n f \to 0$$

with respect to the Hilbert space norm.

Theorem 12.14. *Assume that U is totally ergodic, meaning that U^n is ergodic for every n. Then*

$$\frac{1}{N} \sum_{n=1}^{N} U^{n^2} f \longrightarrow 0.$$

Proof. Let $x_n = U^{n^2} f$. Then

$$\langle x_{n+h}, x_n\rangle = \langle U^{n^2+2nh+h^2} f,\, U^{n^2} f\rangle = \langle U^{2nh+h^2} f,\, f\rangle = \langle U^{2nh} f,\, U^{-h^2} f\rangle.$$

Then by total ergodicity, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} \langle x_{n+h},\, x_n\rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} \langle U^{2hn} f,\, U^{-h^2} f\rangle$$

which by the von Neumann Ergodic Theorem converges to 0, since strong L^2 convergence implies weak convergence. Hence, by van der Corput, we have $\frac{1}{N} \sum_{n=1}^{N} U^{n^2} f \to 0$, completing the proof. ☐

This Theorem implies that if T is totally ergodic, then

$$\frac{1}{N}\sum_{n}\mu\left(T^{-n^2}A\cap A\right)\longrightarrow\mu(A)^2$$

Bibliography

[1] V. Bergelson, Some historical comments and modern questions around the Ergodic Theorem, *Internat. Math. Nachrichten* Nr. 205 (2007) 1-10

[2] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*. J. Wiley & Sons, 1974.

[3] G. Pólya and G. Szegő , *Problems and Theorems in Analysis, vol 1.*, Springer Verlag, 1978.

[4] J. Von Neumann, Gleichmässig dichte Zahlenfolgen (Hungarian; German summary), *Mat. Fiz. Lapok* **32** (1925), 32-40. Also in Collected Works, Vol. I, Pergamon Press, Oxford, 1961, pp. 58-66.