Application of Ergodic Theory to Number Theory

Carllos Eduardo de Holanda

Instituto Superior Tecnico

October 9, 2018

Introduction: the ideia of Ergodic Theory

The general scenario:

- A space X;
- σ-algebra $\mathcal{A}$ in X;
- An **invariant measure** μ and
- A transformation $T : X \to X$, (the **dynamical system**.)

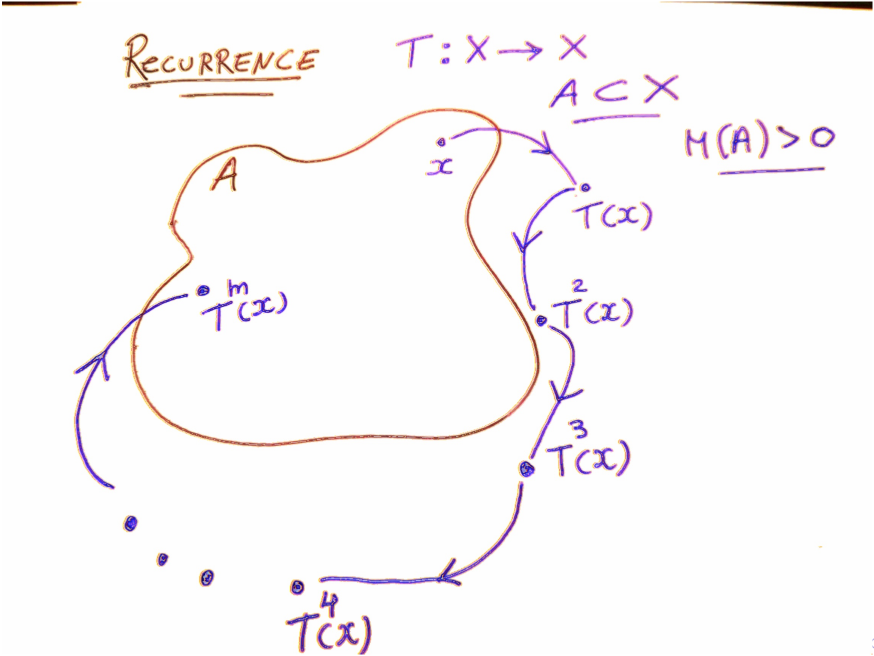
A measure μ is said to be T-invariant if $\mu(T^{-1}(A)) = \mu(A)$ for every $A \in \mathcal{A}$.

Given any point $x \in X$ we have its orbit:

$$x, T(x), T^2(x), ..., T^n(x)...$$

Some Interesting Questions

1. Let $A \subset X$, $A \in \mathcal{A}$ and $\mu(A) > 0$. Is there any point $x \in A$ such that $T^n(x) \in A$ for some $n \geq 1$?
2. If they exist, how many "returning" points are there? and more...
3. How many times they are coming back?



RECURRENCE
T : X → X
A ⊂ X
μ(A) > 0
A
x
T(x)
T²(x)
T³(x)
T⁴(x)
Tᵐ(x)

Answering the questions, we have the classical **Poincare's Recurrence Theorem:**

Theorem 1.

Let $T : X \to X$ *a measurable transformation and let* μ *be a finite* T*-invariant measure in* X*. If* $A \subset X$ *is measurable, then the set*

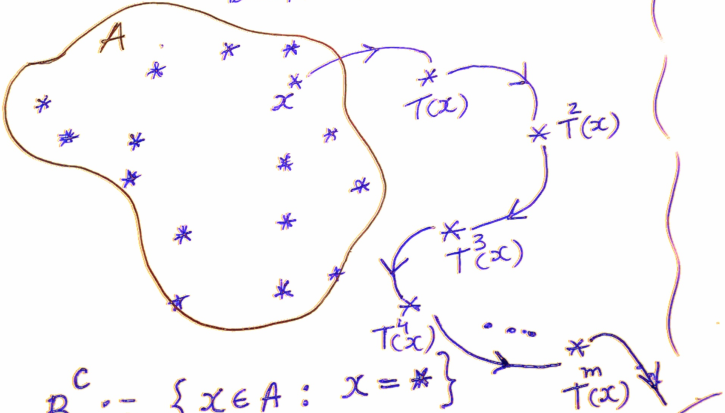
$$B = \{x \in A : T^n x \in A \textit{ for infinitely many } n \in \mathbb{N}\}$$

has measure $\mu(B) = \mu(A)$*.*

RECURRENCE

$T : X \to X$

$B \subset A \subset X$



$$B^c := \{ x \in A : x = * \}$$

$$\mu(B^c) = 0$$

Recurrence

Proof: We have

$$B = A \cap \left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} T^{-k}(A) \right) = A \setminus \bigcup_{n=1}^{\infty} \left(A \setminus \bigcup_{k=n}^{\infty} T^{-k}(A) \right). \quad (1)$$

Moreover,

$$A \setminus \bigcup_{k=n}^{\infty} T^{-k}(A) \subset \bigcup_{k=0}^{\infty} T^{-k}(A) \setminus T^{-n} \left(\bigcup_{k=0}^{\infty} T^{-k}(A) \right). \quad (2)$$

On the other hand, since μ is T-invariant we have

$$\mu \left(\bigcup_{k=0}^{\infty} T^{-k}(A) \right) = \mu \left(T^{-n} \left(\bigcup_{k=0}^{\infty} T^{-k}(A) \right) \right) \quad (3)$$

Moreover, since

$$\bigcup_{k=0}^{\infty} T^{-k}(A) \supset \bigcup_{k=n}^{\infty} T^{-k}(A) = T^{-n}\left(\bigcup_{k=0}^{\infty} T^{-k}(A)\right)$$

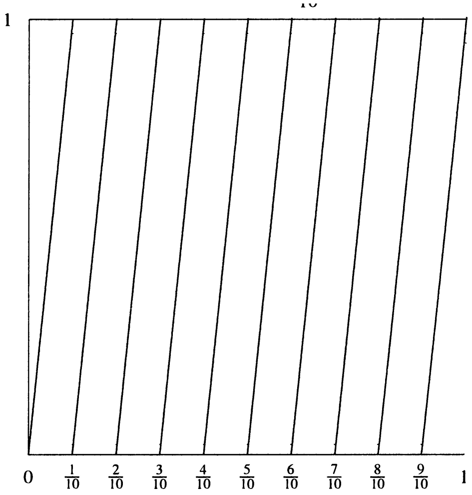
and μ is finite, it follows from (2) and (3) that

$$\mu\left(A \setminus \bigcup_{k=n}^{\infty} T^{-k}(A)\right) \leq \mu\left(\bigcup_{k=0}^{\infty} T^{-k}(A)\right) - \mu\left(T^{-n}\bigcup_{k=0}^{\infty} T^{-k}(A)\right) = 0$$

and from (1) we conclude that $\mu(B) = \mu(A)$. □

Recurrence: The Decimal Expansion

$T : [0,1] \to [0,1]$, $T(x) = 10x - [10x]$, where $[10x]$ represents the greatest integer less or equal to $10x$, that is, $T(x)$ is the fractional part of $10x$.



Decimal Expansion Transformation

This transformation is related to the decimal expansion algorithm: if x is given by

$$x = 0.a_0 a_1 a_2 a_3 ..., \quad a_j \in \{0, 1, 2, ..., 9\},$$

then its image is given by

$$T(x) = 0.a_1 a_2 a_3 a_4 ...$$

and for every $n \geq 1$ we have that

$$T^n(x) = 0.a_n a_{n+1} a_{n+2} a_{n+3} ...$$

We also have an important thing: Lebesgue measure m in $[0, 1]$ is T-invariant.

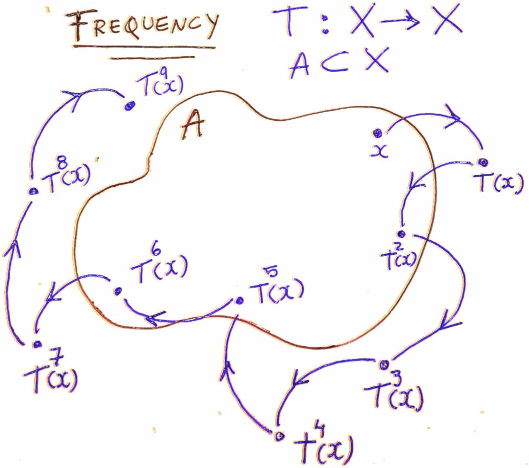
Now let E_7 be the set of all $x \in [0,1]$ which decimal expansion starts with the digit 7, that is, $a_0 = 7$.

By the Poincare's recurrence theorem, there are infinite numbers $n \geq 1$ such that $T^n(x) \in E_7$, that is, $a_n = 7$ for infinite numbers $n \in \mathbb{N}$.

Almost every number x which decimal expansion starts with 7 has infinite digits equal to 7.

□

What can we say about the frequency with which an orbit visits a given set?



Frequency

Let $T : X \to X$ and $A \subset X$. Given $x \in X$ and $n \in \mathbb{N}$, we define

$$\tau_n(A, x) := \mathrm{card}\{k \in \{0, 1, ..., n-1\} : T^k(x) \in A\}.$$

This is the same as writing

$$\tau_n(A, x) = \sum_{k=0}^{n-1} \chi_A(T^k(x)),$$

where χ_A is the characteristic function of the set A. When the limit

$$\tau(A, x) := \lim_{n\to\infty} \frac{\tau_n(A, x)}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_A(T^k(x))$$

exists, it gives the **frequency** with which the orbit of x visits the set A.

Frequency

By Poincare's recurrence theorem, $\tau_n(A, x) \to \infty$ when $n \to \infty$, for μ-almost every $x \in A$, but it gives no information about the frequency.

So,

In what conditions the frequency exists?

Boltzmann Ergodic Hypothesis

This question started when the physicist **Ludwig Boltzmann** (1844-1906) was developing the **Kinetic Theory of Gases**.

Boltzmann needed of an important condition, which is known today as the **Ergodic Hypothesis**. Mathematically, the hypothesis is stated as follows:

Ergodic Hypothesis: For systems describing the movement of particles, the frequency $\tau(A, x)$ for any measurable set A exists and is proportional to the measure of A, for almost every point x.

In modern **Ergodic Theory**, this is the subject considered by **Birkhoff's Ergodic Theorem**.

Birkhoff's Ergodic Theorem

Before the statement of the theorem, some important definitions:

Definition 2.

Given a transformation $T : X \to X$, we say that:

1. A set $A \subset X$ is T-invariant if $T^{-1}(A) = A$.
2. A function $\phi : X \to \mathbb{R}$ is T-invariant if $\phi(T(x)) = \phi(x)$ for every $x \in X$.

And now, the formal statement of the Theorem:

Theorem 3.

*(**Birkhoff's Ergodic Theorem**) Let $T : X \to X$ be a measurable transformation and let μ be a finite T-invariant measure in X. If $\phi \in L^1(X, \mu)$, then the limit*

$$\phi_T(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \phi(T^k(x))$$

exists for μ-almost every $x \in X$ and

1. *ϕ_T is T-invariant almost everywhere.*
2. *$\phi_T \in L^1(X, \mu)$ and*

$$\int_X \phi_T d\mu = \int_X \phi d\mu.$$

In particular, when $\phi = \chi_A$, we obtain the result for the frequency:

$$\tau(A,x) := \lim_{n\to\infty} \frac{\tau_n(A,x)}{n} = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \chi_A(T^k(x))$$

with $\tau(A,x) \in L^1(X,\mu)$ and $\int_X \tau(A,x)d\mu = \mu(A)$.

Since,

$$\frac{1}{n}\sum_{k=0}^{n-1} \phi(T^{k+1}(x)) = \frac{n+1}{n} \cdot \frac{1}{n+1}\sum_{k=0}^{n} \phi(T^k(x)) - \frac{\phi(x)}{n},$$

$\phi_T(T(x))$ is well defined $\iff$ $\phi_T(x)$ is well defined and

$$\phi_T(T(x)) = \phi_T(x).$$

Here we introduce an important concept in ergodic theory: the notion of **ergodic measure**.

Definition 4.

Let $T : X \to X$ be a transformation and μ a measure in X. μ is called *ergodic* with respect to T if the measure of any T-invariant subset $A \subset X$ satisfies $\mu(A) = 0$ or $\mu(X \setminus A) = 0$. In this case we also say that T is ergodic with respect to μ or that the system (T, μ) is ergodic.

We note that an ergodic measure is not necessarily invariant.

Birkhoff's Theorem for Invariant Ergodic Measures

Now we consider the particular case of ergodic invariant measures.

Theorem 5.

Let $T : X \to X$ be a measurable transformation and let μ be T-invariant ergodic measure in X with $\mu(X) < \infty$. If $\phi \in L^1(X, \mu)$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \phi(T^k(x)) = \frac{1}{\mu(X)} \int_X \phi d\mu$$

for μ-almost every $x \in X$.

Taking in particular $\phi = \chi_A$ for some subset $A \subset X$ we obtain the result asked for Boltzmann in his *Ergodic Hypothesis:*

$$\tau(A, x) = \lim_{n\to\infty} \frac{\tau_n(A, x)}{n} = \frac{\mu(A)}{\mu(X)} \quad \text{for } \mu\text{-almost every } x \in X.$$

In most cases, to show that a measure is ergodic is not a simple task!

In general, there is no algorithm to determine if a measure is ergodic.

For example, in some cases we can use a little bit of **Fourier analysis** with some criterion.

Some simple applications

Now we use some of these tools from **Ergodic Theory** to obtain some interesting results of **Number Theory**.

These are just simple cases and there are a lot of more sophisticated applications.

Fractional Parts of Polynomials

Consider the polynomial $P(t) = a_0 t^r + a_1 t^{r-1} + \cdots + a_r$, where $r \in \mathbb{N}$ and $a_0, a_1, ..., a_r \in \mathbb{R}$.

Definition 6.

The numbers $P_n = P(n) \mod 1 \in [0, 1)$, for $n \in \mathbb{N}$, are called the fractional parts of the polynomial P.

Now we introduce the notion of **uniformly distributed** sequence.

Definition 7.

We say that a sequence $(y_k)_{k \in \mathbb{N}} \subset [0, 1]$ is *uniformly distributed* if

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n} \phi(y_k) = \int_0^1 \phi(x) dx$$

for every continuous function $\phi : [0, 1] \to \mathbb{R}$.

Degree 1: Given $r = 1$, $a_0 = \alpha \in \mathbb{R} \setminus \mathbb{Q}$, and $a_1 = \beta \in \mathbb{R}$, we have $P_n = T^n(\beta)$, where $T(x) = x + \alpha \mod 1$ (**Irrational Traslation**).

- T is Ergodic w.r.t. m (using the criterion and Fourier coefficients);
- The Lebesgue measure m is T-invariant.

By B.E.T for ergodic invariant measures, for any continuous function $\phi : [0,1] \to \mathbb{R}$, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} \phi(P_k) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} \phi(T^k(\beta)) = \int_0^1 \phi(x)dx$$

for Lebesgue-almost every $\beta \in \mathbb{R}$. Then, $(P_n)_{n\in\mathbb{N}}$ with $P(t) = \alpha t + \beta$ is uniformly distributed for Lebesgue-almost every $\beta \in \mathbb{R}$. □

Now we consider polynomials of **degree 2**.

Proposition 8.

Given $\alpha \in \mathbb{R} \setminus \mathbb{Q}$*, the sequence* $(P_n)_{n\in\mathbb{N}}$ *of fractional parts of the polynomial* $P(t) = \alpha t^2 + \beta t + \gamma$ *is uniformly distributed for Lebesgue-almost every* $(\beta, \gamma) \in \mathbb{R}^2$*.*

Proof: Let $T : \mathbb{T}^2 \to \mathbb{T}^2$ be defined by

$$T(x, y) = (x + \alpha, y + 2x + \alpha) \mod 1$$

$$T^n(x, y) = (x + n\alpha, \alpha n^2 + 2nx + y) \mod 1 = (x + n\alpha, P_n) \mod 1$$

for every $n \in \mathbb{N}$, with $\beta = 2x$ and $\gamma = y$.

We have that

- m is T-invariant;
- m is ergodic with respect to T.

Given $\psi : [0,1] \to \mathbb{R}$, we can define a continuous function in $\mathbb{T}^2$ by $\phi(x,y) = \psi(y)$. By the B. E. T. for ergodic invariant measures:

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \phi(P_k) &= \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \phi(T^k(x,y)) \\ &= \int_{\mathbb{T}^2} \phi dm \\ &= \int_0^1 \psi(\tau) d\tau\end{aligned}$$

for m-almost every $(x,y) \in \mathbb{T}^2$, and thus for Lebesgue almost every $(\beta,\gamma) \in \mathbb{R}^2$, as desired. $\square$

And one more application: Continued Fractions

Definition 9.

We define the *Gauss transformation* $T : [0,1) \to [0,1)$ by

$$T(x) = \begin{cases} \frac{1}{x} \mod 1 & \text{if } x \neq 0, \\ 0 & \text{if } x = 0; \end{cases}$$

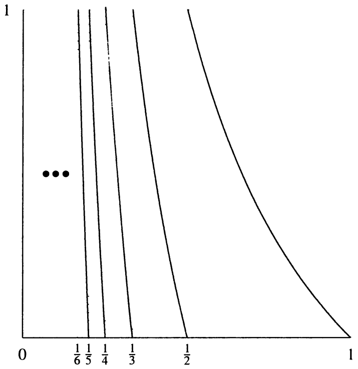


Figure 1.3 The continued fraction map T

Given an irrational number $x \in (0,1)$, we can define positive integers

$$n_j(x) = \left\lfloor \frac{1}{T^{j-1}(x)} \right\rfloor \tag{4}$$

for each $j \in \mathbb{N}$.
Since

$$T^j(x) = T(T^{j-1}(x)) = \frac{1}{T^{j-1}(x)} - n_j(x),$$

we obtain

$$T^{j-1}(x) = \frac{1}{n_j(x) + T^j(x)}$$

for each $j \in \mathbb{N}$.

Therefore,

$$x = \frac{1}{n_1(x) + T(x)} = \frac{1}{n_1(x) + \frac{1}{n_2(x) + T^2(x)}}$$

$$= \frac{1}{n_1(x) + \frac{1}{n_2(x) + \frac{1}{n_3(x) + T^3(x)}}}$$

and so on. Furthermore, we can prove that the sequence

$$\frac{1}{n_1(x)}, \quad \frac{1}{n_1(x) + \frac{1}{n_2(x)}}, \quad \frac{1}{n_1(x) + \frac{1}{n_2(x) + \frac{1}{n_3(x)}}}, \cdots$$

converges to x. So we simply write

$$x = \frac{1}{n_1(x) + \frac{1}{n_2(x) + \cdots}} \tag{5}$$

What is the frequency with which an integer $k \in \mathbb{N}$ occurs in the sequence $n_1(x), n_2(x), \dots$?

Translation to the language of Ergodic Theory: It follows that

$$n_j(x) = k \iff T^{j-1}(x) \in \left(\frac{1}{k+1}, \frac{1}{k} \right].$$

Therefore, given an irrational number $x \in (0,1)$, the frequency with which $k \in \mathbb{N}$ appears in the sequence $n_1(x), n_2(x), \dots$ is given by

$$\begin{aligned} \eta_k(x) &:= \lim_{n \to \infty} \frac{1}{n} \mathrm{card}\{j \in \{1, \dots, n\} : n_j(x) = k\} \\ &= \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n} \chi_{\left(\frac{1}{k+1}, \frac{1}{k}\right]}(T^{j-1}(x)), \end{aligned}$$

whenever this limit exists.

We want to show that $\eta_k(x)$ is well defined for Lebesgue almost every $x \in (0,1)$. We start by defining an appropriate measure.

Definition 10.

The *Gauss measure* is the probability measure μ in $[0,1)$ defined by

$$\mu(A) = \frac{1}{\log 2} \int_A \frac{dx}{1+x}$$

for each measurable set $A \subset [0,1)$.

We note that

$$\frac{m(A)}{2\log 2} \leq \mu(A) \leq \frac{m(A)}{\log 2}, \quad \text{where } m \text{ is the Lebesgue measure.}$$

μ is absolutely continuous with respect to m and vice-versa. That is, the Gauss measure is equivalent to Lebesgue measure.

We have the following properties:

Proposition 11.

1. *The Gauss transformation preserves the Gauss measure;*
2. *The Gauss measure is ergodic with respect to Gauss transformation.*

Now, one more time we can apply the B. E. T. for invariant ergodic measures.

Then, it follows that

$$\begin{aligned}\eta_k(x) &= \lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n}\chi_{\left(\frac{1}{k+1},\frac{1}{k}\right]}(T^{j-1}(x))\\ &= \int_{[0,1)}\chi_{\left(\frac{1}{k+1},\frac{1}{k}\right]}d\mu\\ &= \frac{1}{\log 2}\int_{1\backslash(k+1)}^{1\backslash k}\frac{dx}{1+x}\\ &= \frac{1}{\log 2}\log\frac{(k+1)^2}{k(k+2)}\end{aligned}$$

for μ-almost every $x \in [0,1)$. So, we actually have a formula for the frequency!!!

For example, the frequency with which $k = 1$ occurs is given by $\eta_1(x) = \frac{1}{\log 2}\log\frac{4}{3} \approx 0.415$.

Since $\frac{d}{dx}\left(\log\frac{(x+1)^2}{x(x+2)}\right) = -\frac{1}{x(x+1)(x+2)} < 0$ for $x > 0$, all frequencies $\eta_k(x)$ are distinct and

$$\eta_1(x) > \eta_2(x) > \eta_3(x) > \cdots$$

for μ-almost every $x \in [0, 1)$.

References

1. L. Barreira. Ergodic Theory, Hyperbolic Dynamics and Dimension Theory, Springer, 2012.
2. K. Dajani, C. Kraaikamp. Ergodic Theory of Numbers, The Mathematical Association of America, 2002.
3. K. Oliveira, M. Viana. Fundamentos da Teoria Ergódica, Sociedade Brasileira de Matemática, 2014.
4. H. Weyl, Uber die Gleichverteilung von Zahlen mod. Eins, Math. Ann. 77 (1916), 313-352.