Unpredictability of initial points

Kenshi Miyabe（宮部賢志）

JSPS Research fellow, University of Tokyo

RIMS Workshop: Dynamical Systems and Computation
15 Jan 2014 RIMS

Introduction

- Kenshi Miyabe

- algorithmic randomness
  = computability theory + information theory + etc.

- computable analysis
  = the study of computability of real functions.

- Study the notion of probability via computability theory. (Probability in statistics, learning theory, randomized algorithm and dynamical systems.)

- Chaos theory studies the behavior of dynamical systems that are highly sensitive to initial conditions. *snip* In other words, the deterministic nature of these systems does not make them predictable.
“Wikipedia: Chaos theory”

The orbit is “unpredictable” when the initial point is “unpredictable” or “random”.

Theorem (Birkhoff 1932)

Let (X, μ) be a measure space, $T : X \to X$ be a measure-preserving transformation and $f : X \to \mathbb{R}$ be an integrable function. The value

$$\frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x))$$

converges for all x outside a null set. Furthermore, if T is ergodic, the limit is equal to

$$\int f \; d\mu.$$

In other words, "time average" is equal to "space average".

Theorem (Bienvenu-Day-Hoyrup-Mezhirov-Shen 2012, Franklin-Greenberg-Miller-Ng 2012)

Let (X, μ) be a computable probability space, $T : X \to X$ be a computable measure-preserving ergodic transformation. Then the following are equivalent for $x \in X$:

1. x is Martin-Löf random.
2. We have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x)) = \int f \, d\mu$$

for every integrable lower semicomputable function $f : X \to [0, \infty]$.

Computability theory and algorithmic randomness

What is random?

- Typicalness (Martin-Löf 1966)
- Unpredictability
  (von Mises 1919, Ville 1939, Schnorr 1971)
- Incompressibility
  (Kolmogorov, Levin-Schnorr 1973)

Unpredictability

- The initial capital is 1 yen.
- One predicts the next bit given a finite binary sequence.
- One bets his money to 0 or 1; the money doubles if correct and one loses the money if wrong.
- If the capital is bounded, we say the sequence is unpredictable.

Martingale

Definition

A martingale is a function $d : 2^{<\omega} \to \overline{\mathbb{R}}^{+}$ such that

$$d(\sigma) = \frac{d(\sigma 0) + d(\sigma 1)}{2}.$$

Theorem (Schnorr 1971)

A sequence A is ML-random if and only if $d(A \upharpoonright n) < O(1)$ for every c.e. martingale d.

- Roughly speaking, non-ML-random sequences are the ones one can find some regularity infinitely often.

- Addition, deletion or change of finite bits preserve ML-randomness.

- The set of all ML-random sequences has measure 1.

- Every ML-random sequence satisfies with SLLN.

Kučera's theorem

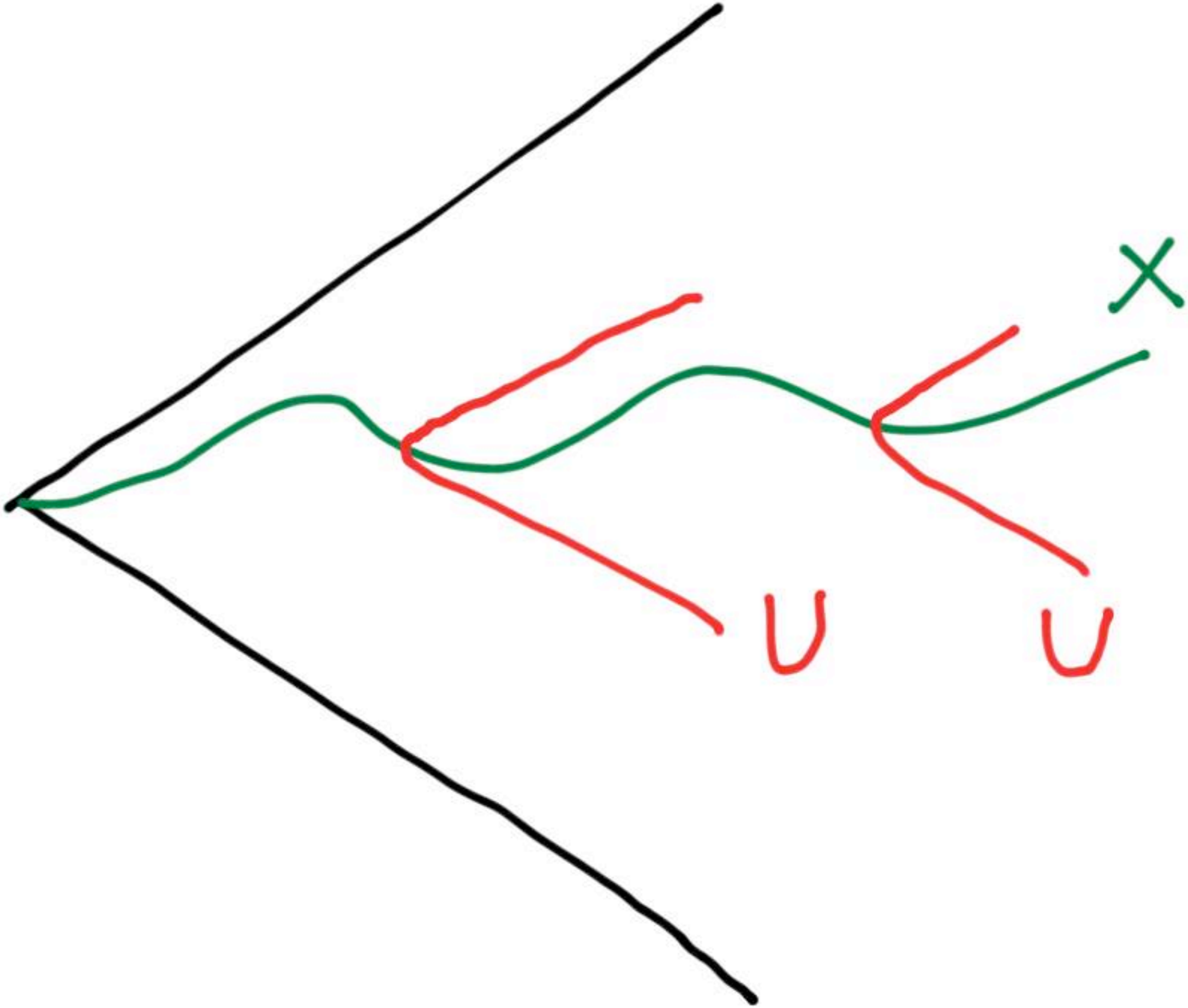
Theorem (Kučera 1985)

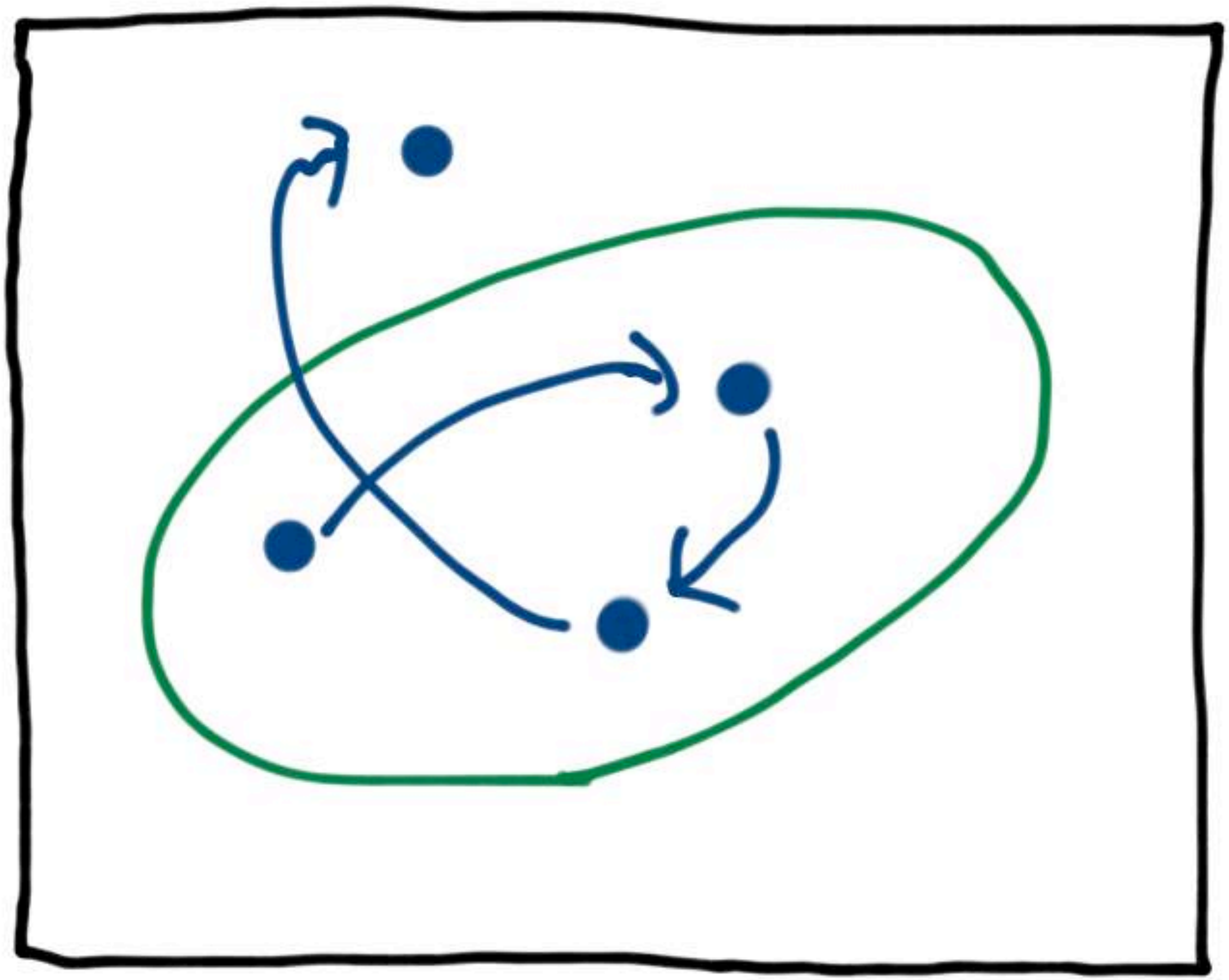
The following are equivalent for $x \in 2^\omega$:

1. x is not Martin-Löf random
2. All tails of x in a c.e. open set U with $\mu(U) < 1$.

Remark

This can be seen as an effectivization of Poincaré reccurrence theorem.


x
U
U



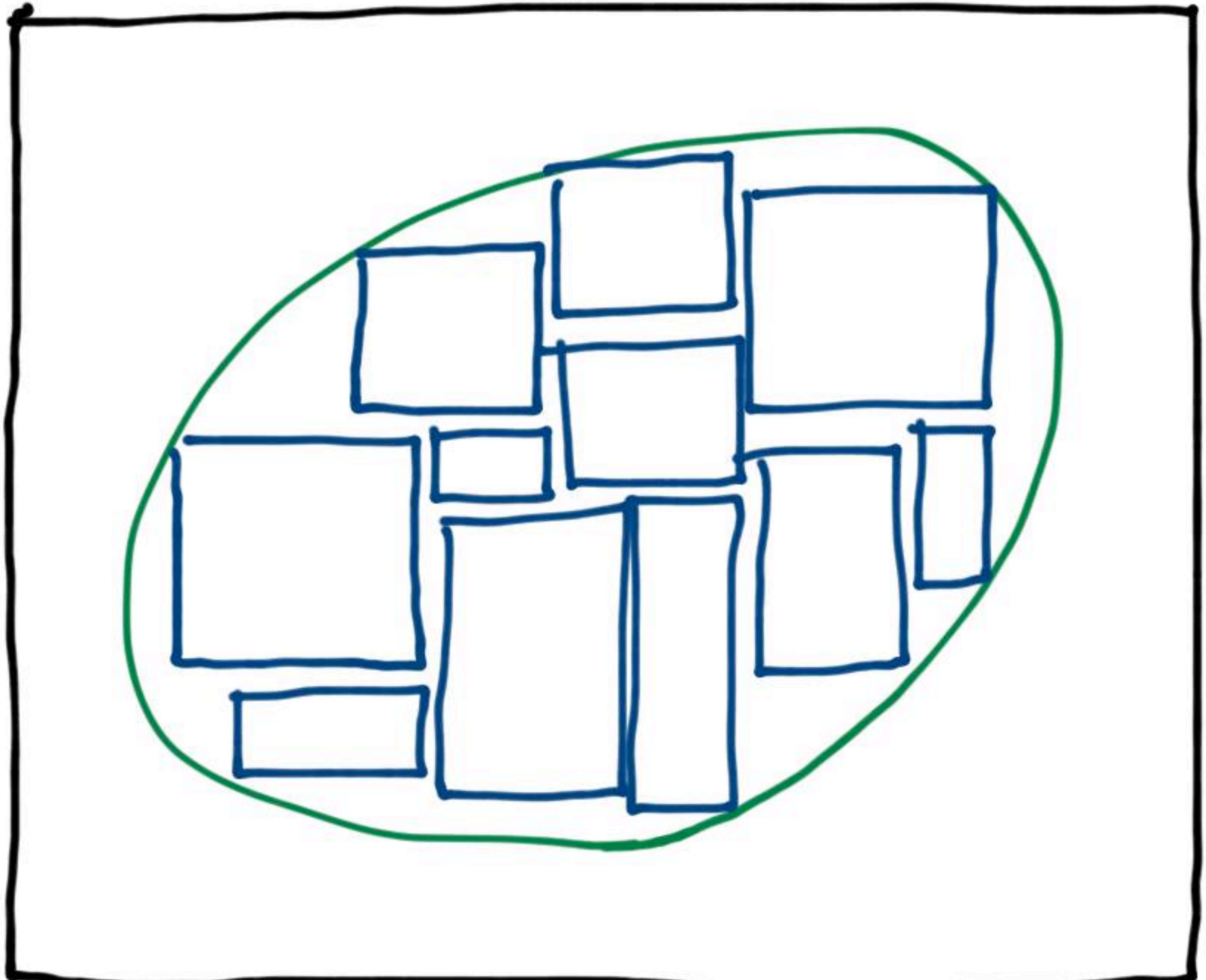
On Cantor space 2^ω, for $\sigma \in \{0,1\}^*$,

$$[\sigma] = \{X \ : \ \sigma \prec X\}$$

forms a base. Thus, every open set G can be written as a union of these base sets:

$$G = \bigcup_{n \in \mathbb{N}} [\sigma_n].$$

If $n \mapsto \sigma_n$ is computable, we say that G is a c.e. open set.



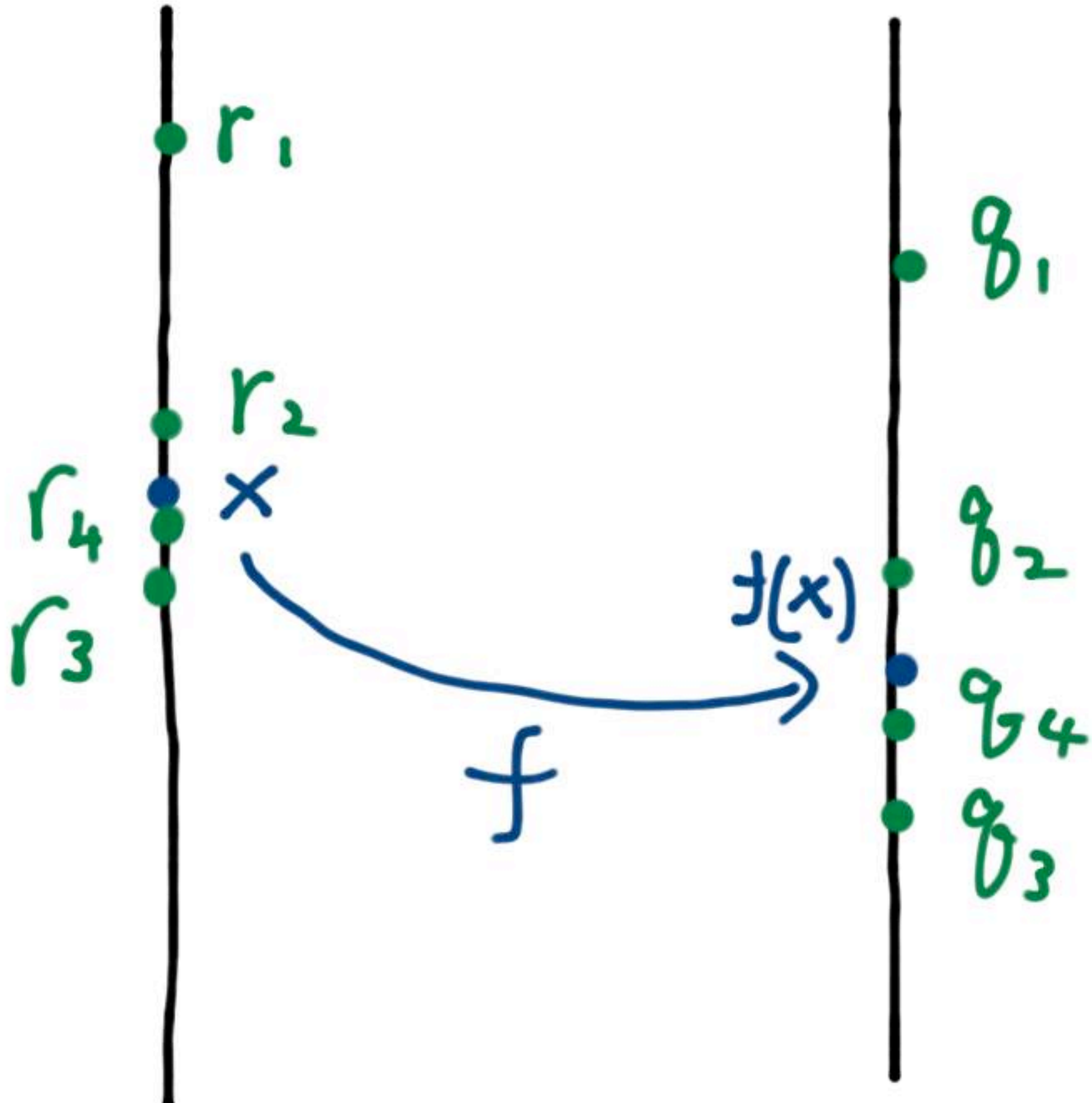
Computability of real functions

A sequence $\{r_n\}$ of rationals is called a fast Cauchy sequence if

$$r_n - r_{n-1} \leq 2^{-n}$$

for every n. Every fast Cauchy sequence $\{r_n\}$ represents the real $x = \lim_n r_n$.

A function $f : \mathbb{R} \to \mathbb{R}$ is called computable if there is a Turing machine M such that, for every $x \in \mathbb{R}$, M maps from every representation of x to a representation of $f(x)$.

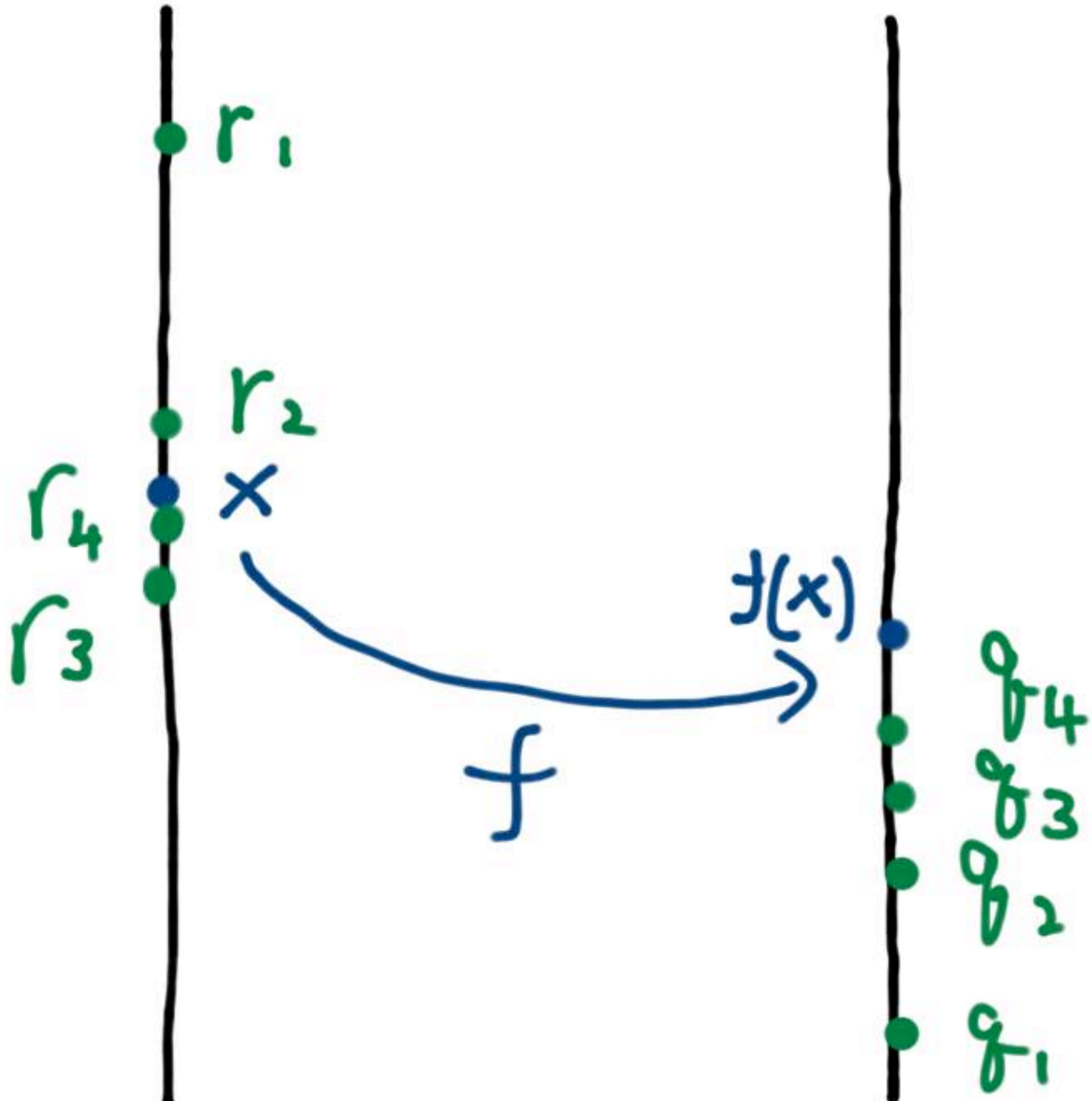

r_1
r_2
r_4
x
r_3
$f(x)$
f
q_1
q_2
q_4
q_3

$x \in \mathbb{R}$ is called left-c.e. or lower semicomputable if there is a computable non-decreasing sequence $\{r_n\}$ such that

$$x = \lim_n r_n.$$

The sequence $\{r_n\}$ is called lower representation of x.

A function $f : \mathbb{R} \to \mathbb{R}$ is called lower semicomputable if there is a Turing machine M such that, for every $x \in \mathbb{R}$, M maps from every representation of x to a lower representation of $f(x)$.



r_1
r_2
r_4
x
r_3
$f(x)$
f
q_4
q_3
q_2
q_1

Theorem

The following are equivalent for $x \in [0,1]$:

(i) x is ML-random.

(ii) $f(x) < \infty$ for every integrable lower semicomputable function f.

Theorem (Bienvenu-Day-Hoyrup-Mezhirov-Shen 2012, Franklin-Greenberg-Miller-Ng 2012)

Let (X, μ) be a computable probability space, $T : X \to X$ be a computable measure-preserving ergodic transformation. Then the following are equivalent for $x \in X$:

1. x is Martin-Löf random.
2. We have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x)) = \int f \; d\mu$$

for every integrable lower semicomputable function $f : X \to [0, \infty]$.

Summary of former half

- The orbit is unpredictable because the initial point is unpredictable.

Computable initial points and orbits

Random initial points?

- One can calculate space average by computing time average with a “random” initial point.
- “One can take a rational numver (or an algebraic number) randomly. It will be correct almost surely.” This is wrong.
- Can we take a computable initial point with a desired property?

Problem

Let (X, μ) be a computable probability space, $T : X \to X$ be a computable measure-preserving ergodic transformation. Assume an integrable lower semicomputable function $f : X \to [0, \infty]$ is given. Does there exist a computable point x such that

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

Can we compute such an x uniformly in f? Can we compute the rate of the convergence?

Theorem (Towsner)

Let (X, μ) be a computable probability space, $T : X \to X$ be a computable measure-preserving ergodic transformation. Suppose that f be $L^1(\mu)$-computable function with $f \in L^2(\mu)$. Let

$$\overline{f}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x)).$$

Then the following are equivalent:

(i) $||\overline{f}||_2$ is computable.

(ii) The pointwise rate of convergence of f is computable.

(iii) $\overline{f}$ is L^1-computable.

Corollary

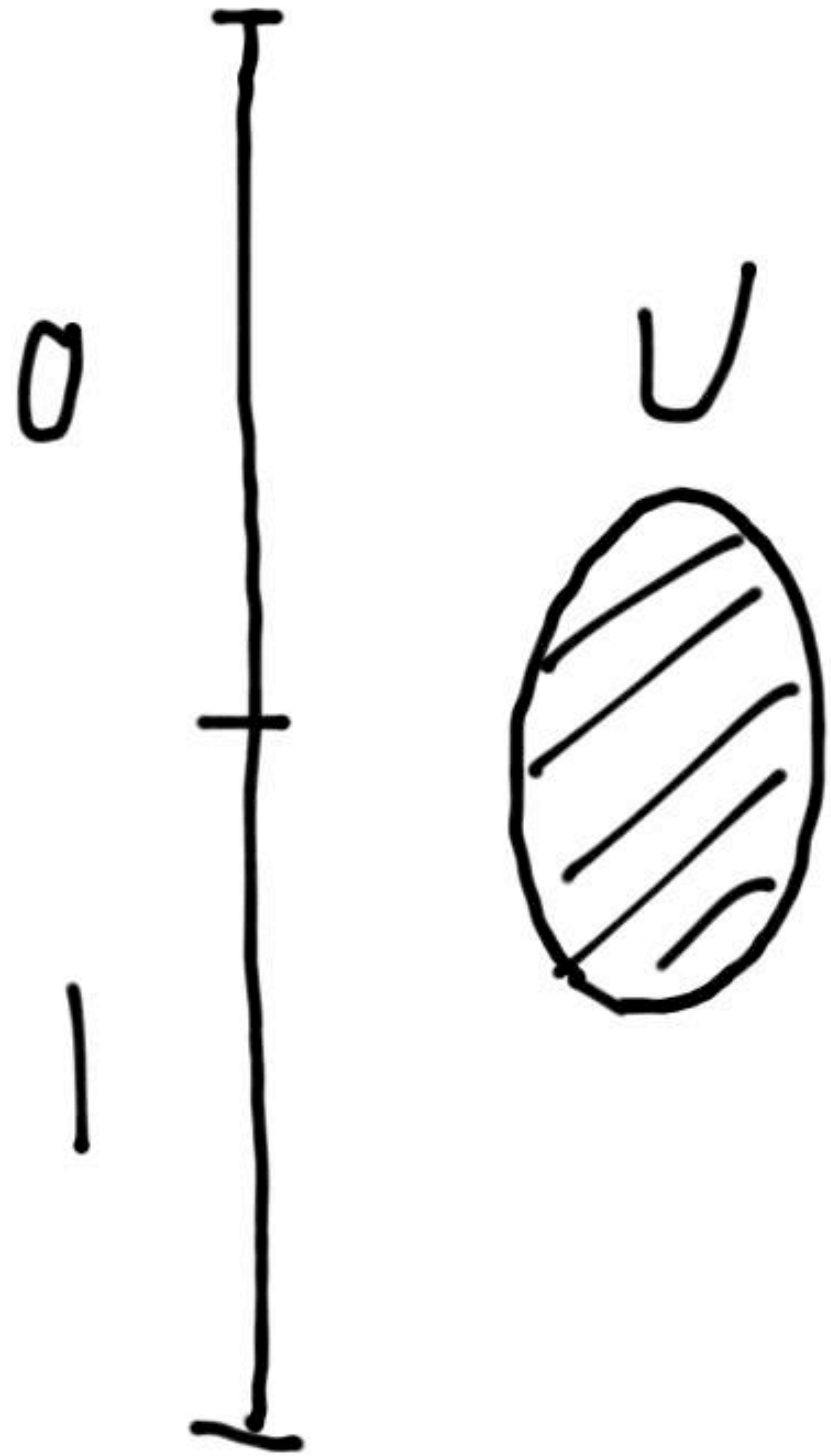
If $||\overline{f}||_2$ is computable, there exists a computable point x such that $\overline{f}(x) = \int f d\mu$.

The complement of a c.e. open set is called a co-c.e. closed set.

Proposition

(i) There exists a co-c.e. closed set that does not contain a computable point.

(ii) Every co-c.e. closed set with a computable measure contains a computable point.


0
1
U

Schnorr randomness

Theorem (Bienvenu-Miller 2012)

The following are equivalent for $x \in 2^\omega$:

1. x is not Schnorr random
2. All tails of x in a c.e. open set U with a computable measure strictly less than 1.

For every such U, the set of all sequences whose all tails in U is a null set. If a set is contained in the null set, such set is called Schnorr null set.

L^1-computability

Definition (Pour-El-Richard 1989)

A function $f : [0,1] \to \mathbb{R}$ is L^1-computable if there exists a computable sequence of rational polygons $\{p_n\}$ such that

$$||f - p_n||_1 \leq 2^{-n}$$

where $||g||_1 = \int |g| d\mu$.

Definition (Pathak-Simpson-Rojas 2014)

Let $\hat{f}(x) = \lim_n p_n(x)$. The value does not depend on $\{p_n\}$ for every Schnorr random point x.

L^1-computability

Observation

Let f, g be effective L^1-computable functions with $||f - g||_1 = 0$. Then, the set

$$\{x \ : \ f \neq g\}$$

is Schnorr null. Thus, there exists a computable point x such that $f(x) = g(x)$.

Convergence speed

Note that L^1-computability means that the convergence speed is computable in L^1-norm.

Question

Can we compute the convergence speed when given a Schnorr point?

Answer

It depends on the randomness deficiency.

Randomness deficiency

(i) $0^\omega = 0000000000000000\cdots$ is not Schnorr random $\Rightarrow$ may not converge.

(ii) $0^{10000000000}101011101011011011010\cdots$ is Schnorr random with large randomness deficiency $\Rightarrow$ The convergence is slow.

(iii) 10101010010010100011101010101 is Schnorr random with low randomness deficientcy $\Rightarrow$ The convergence is fast.

Summary

- For a well-behaved function, one can find a computable point with the desired property.
- Convergence rate depends on the randomness deficiency.

Related work

- The values of entropies of subshifts of finite type over Z^d (d is larger than 1) are exactly right-c.e. reals. (Hochman and Meyerovitch 2010)

- The complexity of the orbits of random points equals to the Kolmogorov-Sinaï entropy. (An effectivization of Brudno's theorem.) The supremum of the complexity of orbits equals the topological entropy. (Galatola-Hoyrup-Rojas 2010)

Future work

- How about polynomial-time computable points?