The relation between the Gibbs measure for a potential which depends on the first coordinate and Double Variational Principle on XY model

Misato Ogawa (Ochanomizu University/RIKEN AIP)

joint work with
Mao Shinoda (Ochanomizu University)

2024.5.31

Background

A **dynamical system** is
a model which formulates the time development of a phenomena.

- Analysis of the orbit on a system
  enables us to understand phenomena.
- By splitting the state space of a system and giving each area
  a symbol, we can express the orbit as a symbolic sequence.

Symbolic dynamics

We say the system is one of the **symbolic dynamics**
which is obtained by describing a dynamical system with symbols.

- It is possible that
  symbolic dynamics make analysis of phenomena easier.

Background

- $\mathbb{K}$: a compact metric space or a countable set
- $\sigma : \mathbb{K}^{\mathbb{Z}} \to \mathbb{K}^{\mathbb{Z}}$;
  $\sigma((\dots, x_0, x_1, x_2, x_3, \dots)) = (\dots, x_1, x_2, x_3, x_4, \dots)$
- A shift with finite symbol : $(\mathcal{A}^{\mathbb{Z}}, \sigma)$
  $\mathcal{A}$: a finite set

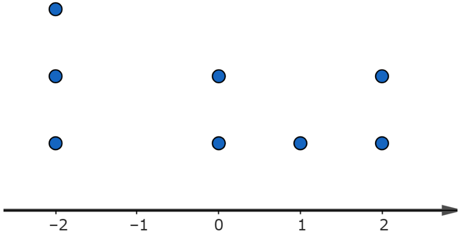

Figure: $\mathcal{A}^{\mathbb{Z}}$

- A shift with countable symbol : $(\mathbb{N}^{\mathbb{Z}}, \sigma)$

Background

- **XY model** : $(M^{\mathbb{Z}}, \sigma)$
  M : a connected and compact manifold

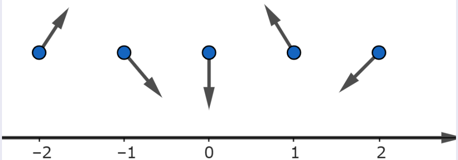

Figure: XY model where M is $\mathbb{S}^1$

- We want to understand the statistical property of XY model.
- See the invariant measure on XY model,
  in particular, "the Gibbs measure".
- Extend conclusions about a shift with finite symbol to XY model.

Background

Relation between Gibbs measures and Double Variational Principle

$\mathbb{K}$	$\mathcal{A}$: a finite set	M : a connected and compact manifold
The Ruelle operator $\mathcal{L}_\varphi f(x)$	$\sum_{\sigma y=x} e^{\varphi(y)} f(y)$	$\int e^{\varphi(ax)} f(ax)\ da$
Gibbs measures for a Hölder continuous function φ satisfy	Variational Principle	Double Variational Principle with potential?

- $\int e^{\varphi(ax)} f(ax)\ da$: [A.Baraviera, et al. 2011]
- Double Variational Principle with potential : [M.Tsukamoto, 2020]

Backgroud

The invariant measure

- $\mathbb{K}$: a compact metric space or a countable set
- $\sigma : \mathbb{K}^{\mathbb{Z}} \to \mathbb{K}^{\mathbb{Z}}$;
  $\sigma((\ldots, x_0, x_1, x_2, x_3, \ldots)) = (\ldots, x_1, x_2, x_3, x_4, \ldots)$
- $\mathscr{M}(\mathbb{K}^{\mathbb{Z}})$: The set of Borel probability measures on $\mathbb{K}^{\mathbb{Z}}$

We say a probability measure $\mu \in \mathscr{M}(\mathbb{K}^{\mathbb{Z}})$ is invariant if

$$\forall E \in \mathcal{B}(\mathbb{K}^{\mathbb{Z}}), \quad \mu(\sigma^{-1}E) = \mu(E).$$

Background

1. The Gibbs measure
2. The equilibrium measure

- If the invariant measure satisfies Variational Principle, we call it the **equilibrium measure**.

	A shift with finite symbol	XY model
	Variational Principle	Double Variational Principle with potential?
①⇒②	Dobruschin' s Theorem[3]	My study
②⇒①	Lanford-Ruelle Theorem[4]	

Main result

- $\mathbb{K} = [0,1]$
- $\sigma : [0,1]^{\mathbb{Z}} \to [0,1]^{\mathbb{Z}}$; $\sigma((x_m)_{m\in\mathbb{Z}}) = (x_{m+1})_{m\in\mathbb{Z}}$: the shift
- $\boldsymbol{d}$: the metric on $[0,1]^{\mathbb{Z}}$;

$$d(x,y) := \sum_{m\in\mathbb{Z}} 2^{-|m|}|x_m - y_m|, \quad (x = (x_m)_{m\in\mathbb{Z}},\ y = (y_m)_{m\in\mathbb{Z}})$$

- $\varphi : [0,1]^{\mathbb{Z}} \to [0,1]$; $\varphi((x_m)_{m\in\mathbb{Z}}) = x_0$: a potential
- $\forall N \in \mathbb{N}$, $\forall A_1, A_2, \dots, A_N \in \mathscr{B}([0,1])$,

$$\boldsymbol{\mu}(\cdots \times [0,1] \times A_1 \times A_2 \times \cdots \times A_N \times [0,1] \times \dots)$$

$$:= \left(\frac{1}{\int_{[0,1]} e^{\varphi(x)}\ dx}\right)^N \int_{A_1} e^{\varphi(x_1)}\ dx_1 \int_{A_2} e^{\varphi(x_2)}\ dx_2 \cdots \int_{A_N} e^{\varphi(x_N)}\ dx_N.$$

Main result

Then, $\boldsymbol{d}$ and $\boldsymbol{\mu}$ satisfy Double Variational Principle with potential.

Background

Variational Principle

- $\mathscr{M}^T(\mathcal{X}) := \{\text{Invariant measures on } (\mathcal{X}, T)\}$ where $\mathcal{X}$ is a compact metric space and T is a continuous map on $\mathcal{X}$ to itself
- h_{top} : the topolpgical entropy
- h_μ : a measure-theoretic entropy
- $\varphi : \mathcal{X} \to \mathbb{R}$: the continuous potential
- $\mathscr{P}_{top}(\cdot)$: Topological pressure

1. $h_{top} = \sup_{\mu \in \mathscr{M}^T(\mathcal{X})} h_\mu,$
2. $\mathscr{P}_{top}(A) = \sup_{\mu \in \mathscr{M}^T(\mathcal{X})}(h_\mu + \int \varphi \, d\mu).$

- Let $(\mathcal{X}, T) = (\mathcal{A}^{\mathbb{Z}}, \sigma)$.

  Then, if φ is Hölder, the Gibbs measure for φ is the unique measure which attains the supremum of ② [R.Bowen, 1975] and we say φ satisfies Variational Principle.

Background

It is known that

$$h_{top}([0,1]^{\mathbb{Z}}, \sigma) = \infty.$$

- It makes no sense to hold Variational Principle on XY model and Double Variational Principle is introduced to overcome this problem.
- Based on the conclusion about a shift with finite symbol, it is natural to suppose Gibbs measures on XY model relate to Double Variational Principle with potential.

Double Variational Principle with potential

Theorem 1 [M.Tsukamoto, 2020]

- $(\mathcal{X}, T)$: a dynamical system which has the marker property
- $\varphi : \mathcal{X} \to \mathbb{R}$: a continuous function
- $\mathscr{M}^T(\mathcal{X})$: the set of T-invariant Borel probability measures on $\mathcal{X}$
- $\mathscr{D}(\mathcal{X})$: the set of distance functions on $\mathcal{X}$

Then,

$$\begin{aligned}\mathrm{mdim}(\mathcal{X}, T, \varphi) &= \min_{d \in \mathscr{D}(\mathcal{X})} \sup_{\mu \in \mathscr{M}^T(\mathcal{X})} \left(\overline{\mathrm{rdim}}(\mathcal{X}, T, d, \mu) + \int_{\mathcal{X}} \varphi \, d\mu \right) \\ &= \min_{d \in \mathscr{D}(\mathcal{X})} \sup_{\mu \in \mathscr{M}^T(\mathcal{X})} \left(\underline{\mathrm{rdim}}(\mathcal{X}, T, d, \mu) + \int_{\mathcal{X}} \varphi \, d\mu \right).\end{aligned}$$

- We call this principle **Double Variational Principle with potential**.

$([0,1]^{\mathbb{Z}}, \sigma)$

- **XY model** $([0,1]^{\mathbb{Z}}, \sigma)$: a dynamical system
  $\sigma : [0,1]^{\mathbb{Z}} \to [0,1]^{\mathbb{Z}}$; $\sigma((x_m)_{m\in\mathbb{Z}}) = (x_{m+1})_{m\in\mathbb{Z}}$: the shift
- $\varphi : [0,1]^{\mathbb{Z}} \to [0,1]$; $\varphi((x_m)_{m\in\mathbb{Z}}) = x_0$: a potential

- $([0,1]^{\mathbb{Z}}, \sigma)$ doesn't have the marker property.
  On the other hand, it is a typical example of XY model.

Calculation of $\mathrm{mdim}([0,1]^{\mathbb{Z}}, \sigma, \varphi)$ [M.Tsukamoto, 2020]

Denote the **mean dimension with potential** by $\mathrm{mdim}(\mathcal{X}, T, \varphi)$.

- $\mathrm{mdim}([0,1]^{\mathbb{Z}}, \sigma, \varphi) = 2$.

- Recall

$$\mathrm{mdim}([0,1]^{\mathbb{Z}}, \sigma, \varphi) = \min_{d} \sup_{\mu} \left(\overline{\mathrm{rdim}}([0,1]^{\mathbb{Z}}, \sigma, d, \varphi) + \int \varphi \, d\mu \right)$$
$$= \min_{d} \sup_{\mu} \left(\underline{\mathrm{rdim}}([0,1]^{\mathbb{Z}}, \sigma, d, \varphi) + \int \varphi \, d\mu \right).$$

Mutual Information

- Fix a probability space $(\Omega, \mathbb{P})$ and
  assume that all random variables are defined on $(\Omega, \mathbb{P})$.
- X, Y : r.v.s taking values in some measurable spaces $\mathcal{X}, \mathcal{Y}$
- $H(X) = -\sum_{x \in \mathcal{X}} \mathbb{P}(X = x) \log \mathbb{P}(X = x)$

Mutual Information of finite sets

- $\mathcal{X}, \mathcal{Y}$: a finite set

Mutual Information $I(X; Y)$;

$$I(X; Y) := H(X) - H(X|Y).$$

Mutual Information

1. Take finite measurable partitions of $\mathcal{X}, \mathcal{Y}$ (not necessarily finite sets) ;
   $\mathcal{P} = \{P_1, .., P_M\} : \bigcup_k P_k = \mathcal{X} \quad M < \infty \quad \mathcal{P}$ is a disjoint set,
   $\mathcal{Q} = \{Q_1, .., Q_N\} : \bigcup_k Q_k = \mathcal{Y} \quad N < \infty \quad \mathcal{Q}$ is a disjoint set.
2. For $x \in \mathcal{X}, y \in \mathcal{Y}$, $\tilde{\mathcal{P}}(x) := P_m, \ \tilde{\mathcal{Q}}(y) := Q_n \quad x \in P_m, \ y \in Q_n$
3. It is possible to consider Mutual Information $I(\mathcal{P} \circ X; \mathcal{Q} \circ Y)$ as the definition of Mutual Information of finite sets.

Mutual Information of any sets

- $\mathcal{X}, \mathcal{Y}$: a set

Mutual Information $I(X;Y)$;

$$I(X;Y) := \sup_{\mathcal{P},\mathcal{Q}} I(\tilde{\mathcal{P}} \circ X; \tilde{\mathcal{Q}} \circ Y)$$

The rate distortion function

Definition of the rate distortion function

- $(\mathcal{X}, T)$: a dynamical system with a T-invariant measure μ
- $\varepsilon \in \mathbb{R}_{>0}$

Then, define the rate distortion function by

$$R(d, \mu, \varepsilon) = \inf_{N,X,Y} \frac{I(X;Y)}{N}, \quad \text{where}$$

- $N \in \mathbb{N}$
- $X, Y = (Y_0, ..., Y_{N-1})$: r.v.s defined on a probability space $(\Omega, \mathbb{P})$, $\forall X, Y_n$, they take values in $\mathcal{X}$ and satisfy

$$X \text{ has the distribution } \mu,$$

$$\mathbb{E}\left(\frac{1}{N}\sum_{n=0}^{N-1} d(T^n X(\omega), Y_n(\omega))\right) < \varepsilon \quad (\omega \in \Omega), \tag{1}$$

and we call (1) the distortion condition.

The rate distortion dimension

Definition of the rate distortion dimension

- $(\mathcal{X}, T)$: a dynamical system with a T-invariant measure μ
- d : a metric on $\mathcal{X}$

Define the upper and the lower rate distortion dimensions by

$$\overline{\mathrm{rdim}}(\mathcal{X}, T, d, \mu) = \limsup_{\varepsilon \to 0} \frac{R(d, \mu, \varepsilon)}{\log(1/\varepsilon)},$$

$$\underline{\mathrm{rdim}}(\mathcal{X}, T, d, \mu) = \liminf_{\varepsilon \to 0} \frac{R(d, \mu, \varepsilon)}{\log(1/\varepsilon)}.$$

- If both of these limits coincide,
  we call the value the **rate distortion dimension** $\mathrm{rdim}(\mathcal{X}, T, d, \mu)$.

Calculation of the rate distortion dimension

Calcurate the rate distortion dimension to evaluate the right-hand side of Double Variational Principle for mean dimension with potential.

Setting

- **XY model** $([0,1]^{\mathbb{Z}}, \sigma)$: a dynamical system
  $\sigma : [0,1]^{\mathbb{Z}} \to [0,1]^{\mathbb{Z}}$; $\sigma((x_m)_{m\in\mathbb{Z}}) = (x_{m+1})_{m\in\mathbb{Z}}$: the shift
- $\boldsymbol{d}$: the metric on $[0,1]^{\mathbb{Z}}$;
  $$d(x,y) := \sum_{m\in\mathbb{Z}} 2^{-|m|}|x_m - y_m|, \quad (x = (x_m)_{m\in\mathbb{Z}},\ y = (y_m)_{m\in\mathbb{Z}})$$
- $\boldsymbol{\varphi} : [0,1]^{\mathbb{Z}} \to [0,1]$; $\varphi((x_m)_{m\in\mathbb{Z}}) = x_0$: a potential

We want to calculate

$\overline{\mathrm{rdim}}([0,1]^{\mathbb{Z}}, \sigma, d, \mu) + \int_{[0,1]^{\mathbb{Z}}} \varphi \, d\mu$ or $\underline{\mathrm{rdim}}([0,1]^{\mathbb{Z}}, \sigma, d, \mu) + \int_{[0,1]^{\mathbb{Z}}} \varphi \, d\mu$

or both.

Calculation of the rate distortion dimension

Under this setting,
we can construct a Gibbs measure on XY model for a potential which depends on the first coordinate.

$\forall N \in \mathbb{N}$, $\forall A_1, A_2, \dots, A_N \in \mathscr{B}([0,1])$,

$$\boldsymbol{\mu}(\cdots \times [0,1] \times A_1 \times A_2 \times \cdots \times A_N \times [0,1] \times \dots)$$
$$:= \left(\frac{1}{\int_{[0,1]} e^{\varphi(x)}\ dx} \right)^N \int_{A_1} e^{\varphi(x_1)}\ dx_1 \int_{A_2} e^{\varphi(x_2)}\ dx_2 \cdots \int_{A_N} e^{\varphi(x_N)}\ dx_N.$$

- $\varphi : [0,1]^{\mathbb{Z}} \to [0,1]$; $\varphi((x_m)_{m\in\mathbb{Z}}) = x_0$
- We call μ the **Gibbs measure** because it is constructed by the eigenvalue and the eigenfunction for the Ruelle operator $\mathcal{L}_\varphi$.

Calculation of the rate distortion dimension

First, calculate $\mathrm{rdim}(\mathcal{X}, \sigma, d, \mu)$.

- Take $\varepsilon \in \mathbb{R}_{>0}$.
- $X, Y = (Y_0, ..., Y_{N-1})$: r.v.s defined on $(\Omega, \mathbb{P})$; $\forall \omega \in \Omega$, $X(\omega) := (X_m(\omega))_{m \in \mathbb{Z}} \in [0,1]^{\mathbb{Z}}$, $Y_k(\omega) := (Y_{k,m}(\omega))_{m \in \mathbb{Z}} \in [0,1]^{\mathbb{Z}}$
- X has the distribution μ.
- Y_k satisfies the distortion condition (1).

- $N \in \mathbb{N}$
- $X, Y = (Y_0, ..., Y_{N-1})$: r.v.s defined on a probability space $(\Omega, \mathbb{P})$, $\forall X, Y_n$, they take values in $\mathcal{X}$ and satisfy

X has the distribution μ,

$$\mathbb{E}\left(\frac{1}{N}\sum_{n=0}^{N-1} d(T^n X(\omega), Y_n(\omega))\right) < \varepsilon \quad (\omega \in \Omega) \tag{1}$$

Calculation of the rate distortion dimension

$$I(X;Y) \geq I((X_0, X_1, \dots, X_{n-1}); (Y_{0,0}, Y_{1,0}, \dots, Y_{n-1,0}))$$

because when we define two maps as following

$$Y : \Omega \to ([0,1]^{\mathbb{Z}})^n \ ; \ Y(\omega) = (Y_0(\omega), Y_1(\omega), \dots, Y_{n-1}(\omega)),$$
$$f : ([0,1]^{\mathbb{Z}})^n \to [0,1]^n \ ; \ f((y_{0,m})_{m\in\mathbb{Z}}, \dots, (y_{n-1,m})_{m\in\mathbb{Z}}) = (y_{0,0}, \dots, y_{n-1,0}),$$

$f \circ Y$ is measurable and we can use data-processing inequality.

Because $X_0, \dots X_{n-1}$ is independent with respect to the distribution μ,

$$I((X_0, X_1, \dots, X_{n-1}); (Y_{0,0}, Y_{1,0}, \dots, Y_{n-1,0})) \geq \sum_{m=0}^{n-1} I(X_m; (Y_{0,0}, Y_{1,0}, \dots, Y_{n-1,0})).$$

Calculation of the rate distortion dimension

By the composision of $f \circ Y$ and the following measurable map for all $m = 0, 1, ..., n-1$ and data processing inequality,

$$\tilde{f}_m : [0,1]^n \to [0,1] \; ; \; (y_{0,0}, \ldots, y_{n-1,0}) \mapsto y_{m,0},$$

$$\sum_{m=0}^{n-1} I(X_m; (Y_{0,0}, Y_{1,0}, \ldots, Y_{n-1,0})) \geq \sum_{m=0}^{n-1} I(X_m; (Y_{m,0})).$$

Hence,

$$I(X;Y) \geq \sum_{m=0}^{n-1} I(X_m; (Y_{m,0})).$$

Moreover, because X and Y hold the distortion condition,

$$\frac{1}{n}\sum_{m=0}^{n-1} \mathbb{E}|X_m - Y_{m,0}| \leq \frac{1}{n}\mathbb{E}\left(\sum_{m=0}^{n-1} d(\sigma^m X, Y_m)\right) < \varepsilon.$$

Calculation of the rate distortion dimension

Define the value $r(\varepsilon)$ by $\boldsymbol{r(\varepsilon)} := \inf_{U,V} I(U; V)$ where

- U, V : the random variables which take values on $[0, 1]$
- U has the distribution μ and V holds $\mathbb{E}|U - V| \le \varepsilon$.

Then, from

$$I(X; Y) \ge \sum_{m=0}^{n-1} I(X_m; Y_{m,0}), \quad \frac{1}{n}\sum_{m=0}^{n-1} \mathbb{E}|X_m - Y_{m,0}| < \varepsilon,$$

$$\boldsymbol{\frac{I(X; Y)}{n}} \ge \frac{1}{n}\sum_{m=0}^{n-1} r(\mathbb{E}|X_m - Y_{m,0}|) \ge r\left(\frac{1}{n}\sum_{m=0}^{n-1} \mathbb{E}|X_m - Y_{m,0}|\right) \ge \boldsymbol{r(\varepsilon)}.$$

Calculation of the rate distortion dimension

Consequently, because $R(d, \mu, \varepsilon) := \inf_{N,X,Y}(I(X;Y)/N)$,

$$R(d, \mu, \varepsilon) \geq r(\varepsilon).$$

On the other hand,

$$r(\varepsilon) \sim |\log \varepsilon| \quad (\varepsilon \to 0).$$

Therefore,

$$\liminf_{\varepsilon \to 0} \frac{R(d, \mu, \varepsilon)}{|\log \varepsilon|} \geq 1.$$

Hence, from the consequence of [Lindenstrauss, et.al 2018],

$$\limsup_{\varepsilon \to 0} \frac{R(d, \mu, \varepsilon)}{|\log \varepsilon|} \leq 1.$$

$$\therefore R(d, \mu, \varepsilon) \sim |\log \varepsilon| \quad (\varepsilon \to 0).$$

Calculation of the rate distortion dimension

Therefore,

$$\limsup_{\varepsilon\to 0} \frac{R(d,\mu,\varepsilon)}{\log(1/\varepsilon)} = \liminf_{\varepsilon\to 0} \frac{R(d,\mu,\varepsilon)}{\log(1/\varepsilon)} = 1.$$

As a result,

$$\mathrm{rdim}([0,1]^{\mathbb{Z}}, \sigma, d, \mu) = 1.$$

Hence, because $\forall (x_m)_{m\in\mathbb{Z}} \in [0,1]^{\mathbb{Z}}$, $\varphi((x_m))_{m\in\mathbb{Z}} \leq 1$ and

the consequences of [M.Tsukamoto, 2020] and [Lindenstrauss, et.al 2018],

$$\mathrm{rdim}([0,1]^{\mathbb{Z}}, \sigma, d, \mu) + \int_{[0,1]^{\mathbb{Z}}} \varphi \; d\mu = 2.$$

The relation between the Gibbs measure and Double Variational Principle

$$\mathrm{mdim}([0,1]^{\mathbb{Z}}, \sigma, \varphi) = \mathrm{rdim}([0,1]^{\mathbb{Z}}, \sigma, d, \mu) + \int_{[0,1]^{\mathbb{Z}}} \varphi \, d\mu.$$

This indicates $\boldsymbol{\mu}$ satisfies
Double Variational Principle for mean dimension with potential.

A Gibbs measure on XY model μ

$\forall N \in \mathbb{N}$, $\forall A_1, A_2, \ldots, A_N \in \mathscr{B}([0,1])$,

$$\boldsymbol{\mu}(\cdots \times [0,1] \times A_1 \times A_2 \times \cdots \times A_N \times [0,1] \times \ldots)$$
$$:= \left(\frac{1}{\int_{[0,1]} e^{\varphi(x)} \, dx} \right)^N \int_{A_1} e^{\varphi(x_1)} \, dx_1 \int_{A_2} e^{\varphi(x_2)} \, dx_2 \cdots \int_{A_N} e^{\varphi(x_N)} \, dx_N.$$

- $\varphi : [0,1]^{\mathbb{Z}} \to [0,1]$; $\varphi((x_m)_{m\in\mathbb{Z}}) = x_0$

Next goal

I try to

calculate both the mean dimension and the rate distortion dimension in the case of defining the measure with a potential which depends on $\boldsymbol{N} \in \mathbb{N}$ coordinates on the XY model.

References I

AT Baraviera, LM Cioletti, Artur O Lopes, Joana Mohr, and Rafael Rigao Souza.
On the general one-dimensional xy model: positive and zero temperature, selection and non-selection.
Reviews in Mathematical Physics, Vol. 23, No. 10, pp. 1063–1113, 2011.

Rufus Bowen.
Equilibrium states and the ergodic theory of anosov diffeomorphisms.
Lecture notes in mathematics, Vol. 470, pp. 11–25, 1975.

PL Dobruschin.
The description of a random field by means of conditional probabilities and conditions of its regularity.
Theory of Probability & Its Applications, Vol. 13, No. 2, pp. 197–224, 1968.

References II

Oscar E Lanford III and David Ruelle.
Observables at infinity and states with short range correlations in statistical mechanics.
Communications in Mathematical Physics, Vol. 13, No. 3, pp. 194–215, 1969.

Elon Lindenstrauss and Masaki Tsukamoto.
From rate distortion theory to metric mean dimension: variational principle.
IEEE Transactions on Information Theory, Vol. 64, No. 5, pp. 3590–3609, 2018.

Masaki Tsukamoto.
Double variational principle for mean dimension with potential.
Advances in Mathematics, Vol. 361, p. 106935, 2020.

References III

Peter Walters.
An introduction to ergodic theory, Vol. 79.
Springer Science & Business Media, 2000.

Appendix

The marker property

We say a dynamical sytem $(\mathcal{X}, T)$ has the **marker property** if $\forall N > 0$, $\exists U \subset \mathcal{X}$: an open set s.t.

$$\mathcal{X} = \bigcup_{n \in \mathbb{Z}} T^{-n} U, \quad U \cap T^{-n} U = \emptyset \quad (1 \leq \forall n \leq N).$$

- $([0,1]^{\mathbb{Z}}, \sigma)$ doesn't have the marker property.