§11. More about entropy

§11.1. The variational principle

Assume that X is a compact metric space and $\varphi : X \to X$ is continuous.

Here we show

Thm $h_{top}(\varphi) = \sup\{ h_\mu(\varphi) \mid \mu \text{ a } \varphi\text{-invariant probability measure} \}$

Part I of the proof is to show that

$$\boxed{h_\mu(\varphi) \le h_{top}(\varphi) \text{ for every } \mu.}$$

That is, $\forall$ finite partition ξ

$$\boxed{h_\mu(\varphi, \xi) \le h_{top}(\varphi).}$$

Step I. 1 Let $\xi = (A_j)_{j=1}^m$.

By the regularity property of Borel measures, $\forall \varepsilon > 0$ $\exists$ compact sets $B_j \subset A_j$ such that $\mu(A_j \setminus B_j) \le \varepsilon$.

Since A_j is a partition, we have $B_j \cap B_k = \emptyset$ for $j \ne k$

Define $B_0 := X \setminus \bigcup_{j=1}^{m} B_j$.

Then $\boxed{\eta := \{B_j\}_{j=0}^{m}}$ is a partition.

We estimate $H(\xi|\eta)$:

$$H(\xi|\eta) = -\sum_{j=0}^{m} \mu(B_j) \sum_{\ell=1}^{m} \mu(A_\ell | B_j) \log \mu(A_\ell|B_j).$$

We only have $\mu(A_\ell|B_j) \neq 0$ in 2 cases:

(a) $j = \ell \Rightarrow$ then $\mu(A_j|B_j) = 1$ as $B_j \subset A_j$,
still gives 0 contribution to $H(\xi|\eta)$

(b) $j = 0 \Rightarrow$ the contribution of this is
$\leq \mu(B_0) \cdot \log m$

And $\mu(B_0) = \sum_{j=1}^{m} \mu(A_j \setminus B_j) \leq m\varepsilon$.

So $H(\xi|\eta) \leq \varepsilon m \log m$.

Choose ε small enough so that

$$\boxed{H(\xi|\eta) \leq 1.}$$

Then we have (by one of the lemmas in §10.3)

$$\boxed{h_\mu(\varphi, \xi) \leq h_\mu(\varphi, \eta) + 1.}$$

Step I. 2

Fix $\varepsilon_0 > 0$ such that $\forall j \neq k$, $j,k \geq 1$, the distance between B_j and B_k is $> 2\varepsilon_0$.

Then $\forall x \in X$, the metric ball $B_d(x, \varepsilon_0)$ intersects ≤ 2 sets in $\eta = \{B_0, B_1, \dots, B_m\}$.

Indeed, it cannot intersect more than one of the sets $B_1, \dots, B_m$.

Recall the Bowen distance: for $n \geq 1$

$$d_n(x,y) = \max_{\ell=0}^{n-1} d(\varphi^\ell(x), \varphi^\ell(y)).$$

<u>Claim:</u> $\forall x \in X$, the Bowen metric ball $B_{d_n}(x, \varepsilon_0)$ intersects $\underline{\leq 2^n \text{ sets}}$ in the partition $\eta^{(n)} := \eta \vee \varphi^{-1}(\eta) \vee \dots \vee \varphi^{-(n-1)}(\eta)$.

Indeed, assume that $y \in B_{d_n}(x, \varepsilon_0)$. Then $\forall \ell = 0, \dots, n-1$, we have

$$\varphi^\ell(y) \in B_d(\varphi^\ell(x), \varepsilon_0)$$

So $\varphi^\ell(y) \in$ at most 2 sets in η (which ones only depends on x)

So we have ≤ 2 options at each of n steps $\Rightarrow \leq 2^n$ options total.

Recalling the notation of §9, let $C_1, \dots, C_N$, $N = D_\varphi(\varepsilon_0, n)$ be a collection of sets s.t. $X = \bigcup_{r=1}^{N} C_r$ and $\mathrm{diam}_{d_n}(C_r) \leq \varepsilon_0$ $\forall r$.

Then each C_r is contained in $B_{d_n}(x, \varepsilon_0)$ for any $x \in C_r$, and thus intersects $\leq 2^n$ sets in $\eta^{(n)}$.

Each nonempty set in $\eta^{(n)}$ has to intersect C_r for some r.

Thus the number $\#(\eta^{(n)})$ of nonempty sets in $\eta^{(n)}$ satisfies

$$\#(\eta^{(n)}) \leq 2^n \cdot N = 2^n \cdot D_\varphi(\varepsilon_0, n).$$

We have (by the Question in §10.1)

$$H_\mu(\eta^{(n)}) \leq \log \#(\eta^{(n)}) \leq n \log 2 + \log D_\varphi(\varepsilon_0, n)$$

Thus $h_\mu(\varphi, \eta) := \lim_{n\to\infty} \frac{1}{n} H_\mu(\eta^{(n)}) \leq$

$$\leq \log 2 + \lim_{n\to\infty} \frac{1}{n} \log D_\varphi(\varepsilon_0, n) \leq \log 2 + h_{top}(\varphi).$$

Step I. 3

Combining steps I. 1 & I. 2, we get

$$h_\mu(\varphi,\xi) \le h_\mu(\varphi,\eta)+1$$
$$\le h_{top}(\varphi)+1+\log 2.$$

Therefore $\boxed{h_\mu(\varphi) \le h_{top}(\varphi)+1+\log 2}$

This is true $\forall$ continuous μ-preserving φ.

We apply this to the iterate φ^k where k is large.

We have $h_\mu(\varphi^k) = k\, h_\mu(\varphi)$ (see Pset 5)

and $h_{top}(\varphi^k) = k\, h_{top}(\varphi)$ (see §9).

So $h_\mu(\varphi) \le h_{top}(\varphi) + \dfrac{1+\log 2}{k}$.

Letting $k \to \infty$, we get

$$h_\mu(\varphi) \le h_{top}(\varphi)$$ as needed.

Part II of the proof is

to show that $\forall \varepsilon > 0$

$\exists$ φ-invariant prob. measure μ such that

$$h_\mu(\varphi) \geq h_\varepsilon(\varphi) \leftarrow \text{ defined in §9.}$$

Step II.1

Recall that $h_\varepsilon(\varphi) = \lim\limits_{n\to\infty} \frac{1}{n} \log D_\varphi(\varepsilon, n)$

where $D_\varphi(\varepsilon, n) =$ minimal size of a cover of X by sets of d_n-diameter $\leq \varepsilon$

From Lemma 1 in §9.1 we have

$D_\varphi(\varepsilon, n) \leq N_\varphi(\frac{\varepsilon}{2}, n)$ where

$N_\varphi(\varepsilon, n) =$ maximal number of points in X which are ε-separated w.r.t. the Bowen metric d_n.

Thus, replacing ε with $\frac{\varepsilon}{2}$, it suffices to show:

$\forall \varepsilon > 0$ $\exists$ φ-invariant probability measure μ such that

$$\boxed{\liminf_{n\to\infty} \frac{\log N_\varphi(\varepsilon, n)}{n} \leq h_\mu(\varphi).}$$

We first construct the measure μ and then the partition ξ so that $\liminf \ldots \leq h_\mu(\varphi, \xi)$.

<u>Step II.2</u> : Construction of the measure μ

Denote by δ_x the δ-measure centered at $x \in X$.

Take $n \geq 1$ & let $\boxed{E_n \subset X}$ be a set with $\#(E_n) = N_\varphi(\varepsilon, n)$ which is ε-separated w.r.t. d_n, i.e. $\forall x, y \in E_n$, if $x \neq y$ then $\forall j = 0, \dots, n-1$

$$d(\varphi^j(x), \varphi^j(y)) > \varepsilon.$$

Define the measures $\boxed{\nu_n := \frac{1}{N_\varphi(\varepsilon,n)} \sum_{x \in E_n} \delta_x}$

and $\mu_n := \frac{1}{n} \sum_{j=0}^{n-1} \varphi^j_* \nu_n$, i.e. $\forall f$

$$\int_X f \, d\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} \int_X f \circ \varphi^j \, d\nu_n$$

$$= \frac{1}{n \cdot N_\varphi(\varepsilon,n)} \sum_{j=0}^{n-1} \sum_{x \in E_n} f(\varphi^j(x)).$$

By Compactness Thm for probability measures, $\exists$ subsequence $\boxed{\mu_{n_k} \to \text{some } \mu}$ weakly.

Similarly to the proof of Krylov-Bogolyubov Thm. in §1 we see that <u>μ is φ-invariant.</u>

Step II.3: construction of the partition ξ

We claim that $\exists$ a finite partition ξ such that: each $A \in \xi$ satisfies

$$\mathrm{diam}_d(A) < \varepsilon,$$

$$\mu(\partial A) = 0$$

where $\partial A := \overline{A} \setminus \mathrm{interior}(A)$.

Indeed, $\forall x \in X$ $\exists$ $\varepsilon_x < \varepsilon/2$ such that

$$\mu(\partial B(x, \varepsilon_x)) = 0$$

(since $\partial B(x, \varepsilon_x) \subset \{y \in X \mid d(x,y) = \varepsilon_x\}$ are disjoint for different ε_x and their uncountable union over $\varepsilon_x \in (0, \varepsilon)$ has finite measure)

Since X is compact, we can pick finitely many points $x_1, \dots, x_m \in X$ such that

$$X \subset \bigcup_{j=1}^{m} B(x_j, \varepsilon_{x_j}).$$

Now let $\xi = \{A_1, \dots, A_m\}$ where

$$x \in A_j \iff \begin{array}{l} x \in B(x_j, \varepsilon_{x_j}) \text{ and} \\ x \notin B(x_\ell, \varepsilon_{x_\ell}) \quad \forall \ell < j. \end{array}$$

By a standard property of weak convergence, since $\mu_{n_k} \xrightarrow[k\to\infty]{} \mu$ weakly, we have $\mu_{n_k}(A) \xrightarrow[k\to\infty]{} \mu(A)$ for all $A \in \xi$.

Since μ is φ-invariant, we in fact have $\forall n \geq 0$, $\forall A \in \xi$

$$\mu(\partial \varphi^{-n}(A)) = 0.$$

This shows the convergence statement

$\forall \ell$ $\forall A \in \xi^{(\ell)}$ $(= \xi \vee \varphi^{-1}(\xi) \vee \dots \vee \varphi^{-(\ell-1)}(\xi))$

we have $\boxed{\mu_{n_k}(A) \xrightarrow[k\to\infty]{} \mu(A)}$.

This implies the <u>entropy convergence</u>: $\forall \ell$

$$H_{\mu_{n_k}}(\xi^{(\ell)}) \xrightarrow[k\to\infty]{} H_\mu(\xi^{(\ell)}). \quad (*)$$

<u>Step II.4</u>: We are trying to prove that for μ in <u>Step II.2</u> and ξ in <u>Step I.3</u>

$$h_\mu(\varphi, \xi) \geq \liminf_{n\to\infty} \frac{\log N_\varphi(\varepsilon, n)}{n}.$$

Recall that

$$h_\mu(\varphi, \xi) = \lim_{\ell\to\infty} \frac{H_\mu(\xi^{(\ell)})}{\ell}.$$

So we need a lower bound on $H_\mu(\xi^{(\ell)})$ and for that we will bound below

$H_{\mu_n}(\xi^{(\ell)})$ (here ℓ fixed, $n\to\infty$).

Recall also that

$\mu_n = \frac{1}{n}\sum_{j=0}^{n-1} \varphi_*^j \nu_n$ where

$$\boxed{\nu_n = \frac{1}{N_\varphi(\varepsilon,n)} \sum_{x\in E_n} \delta_x.}$$

Here we do some preparation and give 2 facts:

<u>Fact 1</u>: $H_{\nu_n}(\xi^{(n)}) = \log N_\varphi(\varepsilon, n)$.

Indeed, each element of $\xi^{(n)}$ has d_n-diameter $< \varepsilon$, so it can contain <u>at most one</u> point in E_n. (as E_n was ε-separated)

So $\xi^{(n)}$ has $N_\varphi(\varepsilon,n)$ sets of $\nu_n = \frac{1}{N_\varphi(\varepsilon,n)}$ each.

Fact 2: For any partition η,

$$H_{\mu_n}(\eta) \geq \frac{1}{n} \sum_{j=0}^{n-1} H_{\varphi_*^j \nu_n}(\eta).$$

Proof: denoting $\Phi(x) := x \log x$ which is **convex**,

we have

$$H_{\mu_n}(\eta) = -\sum_{B \in \eta} \Phi(\mu_n(B)) = -\sum_{B \in \eta} \Phi\left(\frac{1}{n}\sum_{j=0}^{n-1} \varphi_*^j \nu_n(B)\right)$$

$$\geq -\sum_{B \in \eta} \frac{1}{n}\sum_{j=0}^{n-1} \Phi(\varphi_*^j \nu_n(B))$$

$$= -\sum_{j=0}^{n-1} \frac{1}{n} \sum_{B \in \eta} \Phi(\varphi_*^j \nu_n(B)) = \frac{1}{n}\sum_{j=0}^{n-1} H_{\varphi_*^j \nu_n}(\eta).$$

Step II.5: We now come back to $\log N_\varphi(\varepsilon, n) = H_{\nu_n}(\xi^{(n)})$ and try to get an upper bound using various inequalities for the entropy.

Our goal is to get $H_{\mu_n}(\xi^{(\ell)})$ in the RHS.

First, let us write
$\boxed{n = q\ell + r}$ where $\boxed{0 \le r < \ell.}$

Then $\xi^{(n)} = \bigvee_{m=0}^{q-1} \varphi^{-m\ell}(\xi^{(\ell)}) \vee \left(\bigvee_{j\in S} \varphi^{-j}(\xi)\right)$

where $S = \{q\ell, q\ell+1, \dots, \underbrace{q\ell+r-1}_{= n-1}\}$.

Using that $H_\nu(\eta \vee \zeta) \le H_\nu(\eta) + H_\nu(\zeta)$,
(for any ν, η, ζ)
we get

$$H_{\nu_n}(\xi^{(n)}) \le \sum_{m=0}^{q-1} H_{\nu_n}(\varphi^{-m\ell}(\xi^{(\ell)})) + \sum_{j\in S} H_{\nu_n}(\varphi^{-j}(\xi)).$$

Now, $H_{\nu_n}(\varphi^{-j}(\xi)) \le \log \#(\xi)$
(by the Question in §10.1),
and $H_{\nu_n}(\varphi^{-m\ell}(\xi^{(\ell)})) = H_{\varphi_*^{m\ell}\nu_n}(\xi^{(\ell)})$.

So $\log N_\varphi(\varepsilon, n) = H_{\nu_n}(\xi^{(n)})$
$\le \sum_{m=0}^{q-1} H_{\varphi_*^{m\ell}\nu_n}(\xi^{(\ell)}) + \log \#(\xi).$

This captures some elements of
the sum $\sum_{j=0}^{n-1} H_{\varphi_*^j\nu_n}(\xi^{(\ell)}) \le n H_{\mu_n}(\xi^{(\ell)})$ (by Fact 2)

To capture all the elements, we need a slightly refined argument.

Fix ℓ, assume $n > \ell$.

Take p integer s.t. $0 \le p < \ell$

(our previous argument corresponded to $p=0$)

Define $q_p := \left\lfloor \frac{n-p}{\ell} \right\rfloor$. Then

$$\xi^{(n)} = \bigvee_{m=0}^{q_p - 1} \varphi^{-m\ell - p}\left(\xi^{(\ell)}\right) \vee \left(\bigvee_{j \in S_\ell} \varphi^{-j}(\xi) \right)$$

where $S_\ell = \{0, 1, \dots, p-1\} \cup$

$\cup \{q_p \ell + p, \dots, n-1\}$.

Note that $|S_\ell| \le 2\ell$.

E.g. $n = 12$, $\ell = 3$, $p = 1$: $q_p = 3$,

$$0 \quad \underbrace{\boxed{1 \quad 2 \quad 3}}_{m=0} \quad \underbrace{\boxed{4 \quad 5 \quad 6}}_{m=1} \quad \underbrace{\boxed{7 \quad 8 \quad 9}}_{m=3} \quad 10 \quad 11$$

and $0, 10, 11 \in S$.

18.118
11-14

Now $\log N_\varphi(\varepsilon,n) = H_{\nu_n}(\xi^{(n)}) \le$

$$\le \sum_{m=0}^{q_p-1} H_{\nu_n}(\varphi^{-ml-p}(\xi^{(l)})) + \sum_{j\in S} H_{\nu_n}(\varphi^{-j}(\xi))$$

$$\le \sum_{m=0}^{q_p-1} H_{\varphi_*^{ml+p}\nu_n}(\xi^{(l)}) + 2l\cdot\log\#(\xi)$$

This is true $\forall p = 0,1,\dots,l$.

Summing over p, we get

$$l\cdot\log N_\varphi(\varepsilon,n) \le \sum_{p=0}^{l-1}\sum_{m=0}^{q_p-1} H_{\varphi_*^{ml+p}\nu_n}(\xi^{(l)}) + 2l^2\log\#(\xi)$$

$$= \sum_{j=0}^{n-1} H_{\varphi_*^j\nu_n}(\xi^{(l)}) + 2l^2\log\#(\xi)$$

(Fact 2) $\eta := \xi^{(l)}$ $\le n H_{\mu_n}(\xi^{(l)}) + 2l^2\log\#(\xi)$.

Thus

$$\frac{\log N_\varphi(\varepsilon,n)}{n} \le \frac{H_{\mu_n}(\xi^{(l)})}{l} + \frac{2l}{n}\log\#(\xi).$$

Step II.6: We now fix ℓ and let $n = n_k$, $k \to \infty$ in the last inequality:

$$\liminf_{n\to\infty} \frac{\log N_\varphi(\varepsilon, n)}{n} \le \liminf_{k\to\infty} \frac{\log N_\varphi(\varepsilon, n_k)}{n_k}$$

$$\le \lim_{k\to\infty} \frac{H_{\mu_{n_k}}(\xi^{(\ell)})}{\ell} \underset{\text{(by Step II.3)}}{=} \frac{H_\mu(\xi^{(\ell)})}{\ell}.$$

Since this holds $\forall \ell \geq 1$, we now take $\lim_{\ell\to\infty}$ to get

$$\liminf_{n\to\infty} \frac{\log N_\varphi(\varepsilon, n)}{n} \le \lim_{\ell\to\infty} \frac{H_\mu(\xi^{(\ell)})}{\ell} = h_\mu(\varphi, \xi),$$

finishing the proof of Part II of the variational principle. $\square$

Example:

$X = S^1 = \mathbb{R}/\mathbb{Z}$, $\varphi(x) = 2x \mod \mathbb{Z}$.

Let's take as E_n the set

$$E_n = \left\{ \frac{j}{N} \;\middle|\; 0 \leq j < N \right\}$$

where N is odd and

$$2^n \leq N \leq 2^{n+1}.$$

This is not quite a maximal d_n-separated set but is close enough.

Now, $\nu_n = \frac{1}{N} \sum_{j=0}^{N-1} \delta_j$.

And actually, ν_n is φ-invariant: indeed, $\varphi : E_n \circlearrowleft$ is a permutation since N is odd.

So $\mu_n = \frac{1}{n} \sum_{\ell=0}^{n-1} \varphi_*^\ell \nu_n = \nu_n$ and

$\mu_n \xrightarrow[n\to\infty]{}$ Lebesgue measure weakly.

§11.2. A bit of symbolic dynamics

Assume $\varphi : X \circlearrowleft$ is a map.

Take some partition $\xi = (A_\ell)_{\ell=1}^{m}$ of X:

the elements of ξ can have "measure 0" overlap (let's not worry about it here).

Define the alphabet

$$\mathcal{A} = \{1, \dots, m\}$$

and consider the set of sequences

$$\mathcal{A}^{\mathbb{N}_0} = \{\alpha_0 \alpha_1 \alpha_2 \dots \mid \alpha_j \in \mathcal{A}\}.$$

For $x \in X$, define $\Phi(x) \in \mathcal{A}^{\mathbb{N}_0}$ as follows: $\Phi(x) = \alpha_0 \alpha_1 \dots$

where $\boxed{\varphi^j(x) \in A_{\alpha_j}}$

Define the shift operator $\sigma : \mathcal{A}^{\mathbb{N}_0} \circlearrowleft$ by

$$\sigma(\alpha_0 \alpha_1 \dots) = \alpha_1 \alpha_2 \dots$$

Then we have the commutative diagram

$$\begin{tikzcd} X \arrow[r, "\varphi"] \arrow[d, "\Phi"'] & X \arrow[d, "\Phi"] \\ \mathcal{A}^{\mathbb{N}_0} \arrow[r, "\sigma"] & \mathcal{A}^{\mathbb{N}_0} \end{tikzcd}$$

This can sometimes be used to reduce questions about the dynamics of φ to those for σ, which is a "combinatorial object"

<u>Basic example</u>: $X = \mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$, $\varphi(x) = 2x \bmod \mathbb{Z}$

$\xi = \{A_0, A_1\}$, $A_0 = [0, \frac{1}{2}]$, $A_1 = [\frac{1}{2}, 1]$.

$\Phi : x \mapsto \alpha_0 \alpha_1 \ldots$ s.t.

$$x = \sum_{j=0}^{\infty} \alpha_j 2^{-j-1} = 0.\alpha_0\alpha_1\ldots$$
in binary

<u>Note</u>: on the side of $\{0,1\}^{\mathbb{N}_0}$, the Lebesgue measure on X corresponds to taking α_j i.i.d. with

$$P(\alpha_j = 0) = P(\alpha_j = 1) = \frac{1}{2}.$$

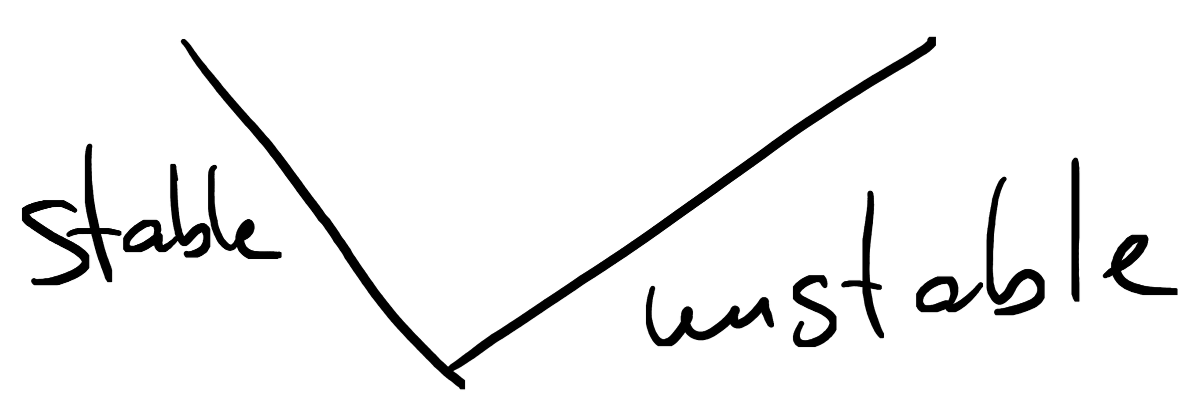
More complicated example:

$X = \mathbb{T}^2 = \mathbb{R}^2 / \mathbb{Z}^2$,

$\varphi(x) = Ax \mod \mathbb{Z}^2$,

$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ A has eigenvalues $\lambda = \frac{3+\sqrt{5}}{2}$ and λ^{-1}.

In the pictures below we use the



directions for A, given by eigenvectors with eigenvalues λ, λ^{-1}.

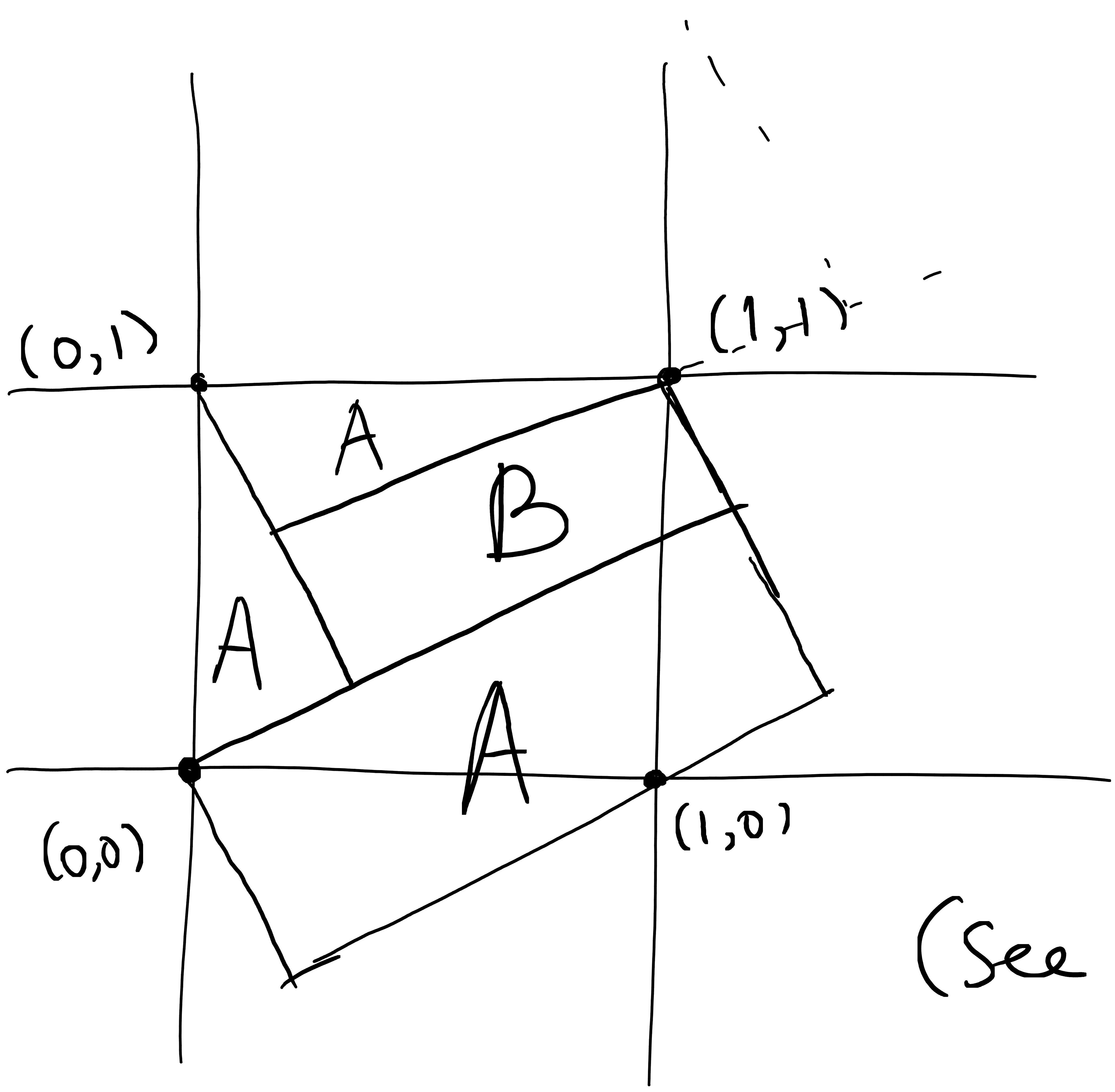
We will define a <u>Markov partition</u>

$\xi = \{ \Delta_1, \Delta_2, \Delta_3, \Delta_4, \Delta_5 \}$

where each Δ_j is a rectangle with stable/unstable sides.

18.118
11-20

We first partition $X = \mathbb{T}^2$ into A and B which are also rectangles: we use the stable/unstable manifolds passing through $(0,0) \in \mathbb{T}^2$ which corresponds to $\mathbb{Z}^2 \subset \mathbb{R}^2$ under the projection $\mathbb{R}^2 \to \mathbb{T}^2$:

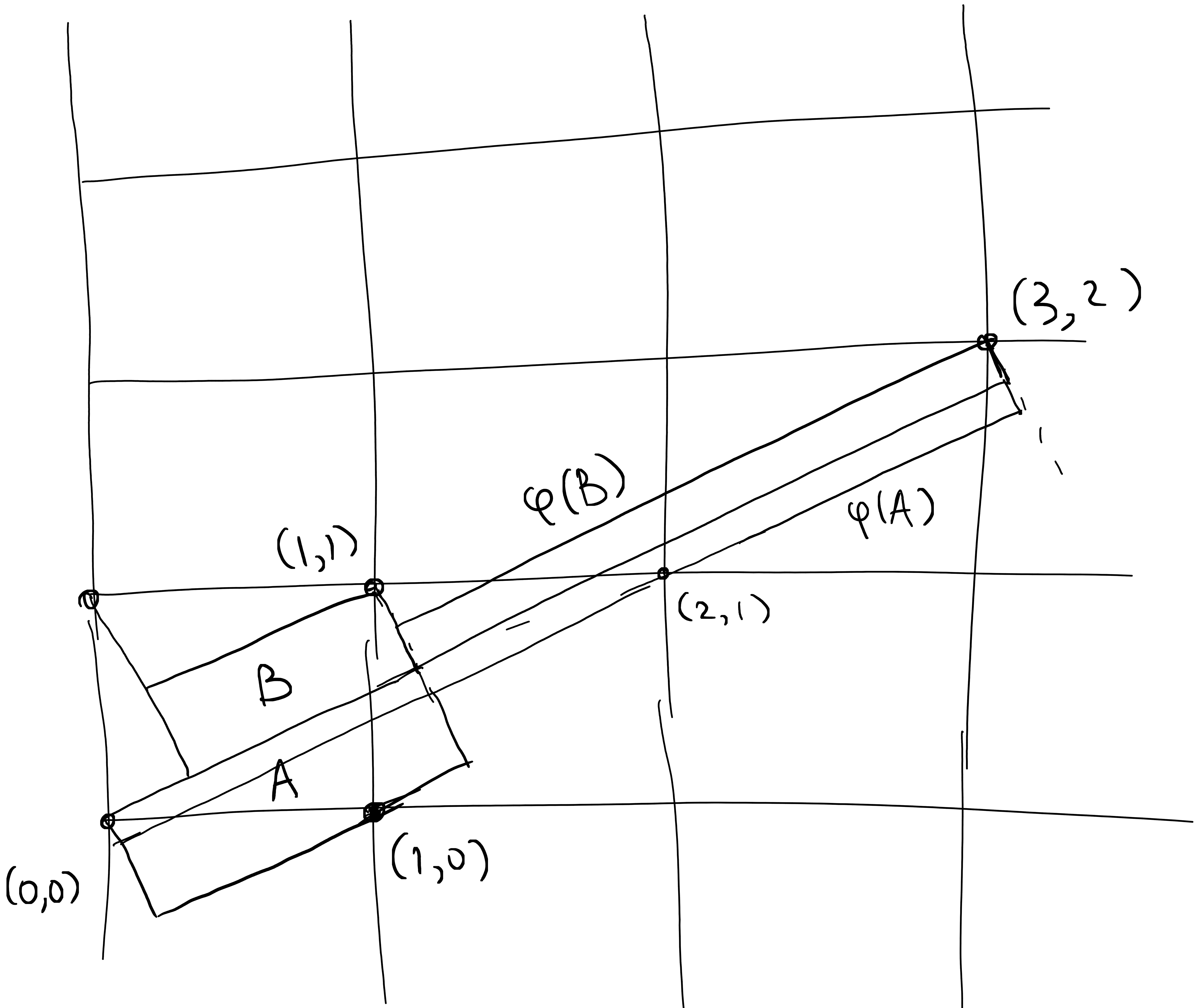


(See the slides...)

18.118
11-21

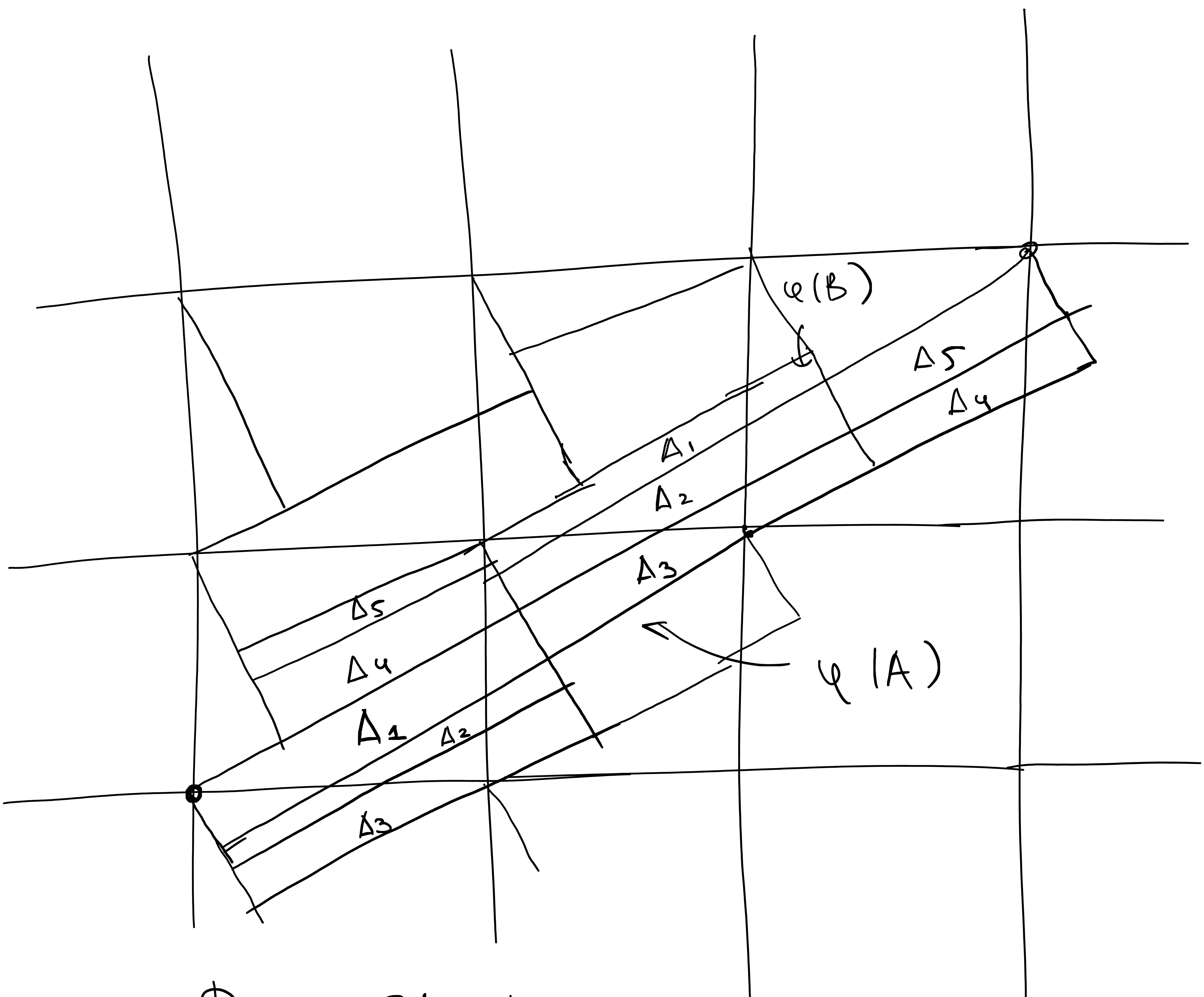
Now let us draw $\varphi(A), \varphi(B)$

(So we are effectively studying dynamics of φ^{-1}: $x \in \varphi(A) \Leftrightarrow \varphi^{-1}(x) \in A \ldots$)



$$\varphi: (1,1) \mapsto (3,2)$$
$$(0,0) \mapsto (0,0)$$
$$(1,0) \mapsto (2,1)$$
$$(0,1) \mapsto (1,1)$$

We see that the intersections of A, B with $\varphi(A), \varphi(B)$ consist of 5 rectangles. Those are $\Delta_1, \dots, \Delta_5$:

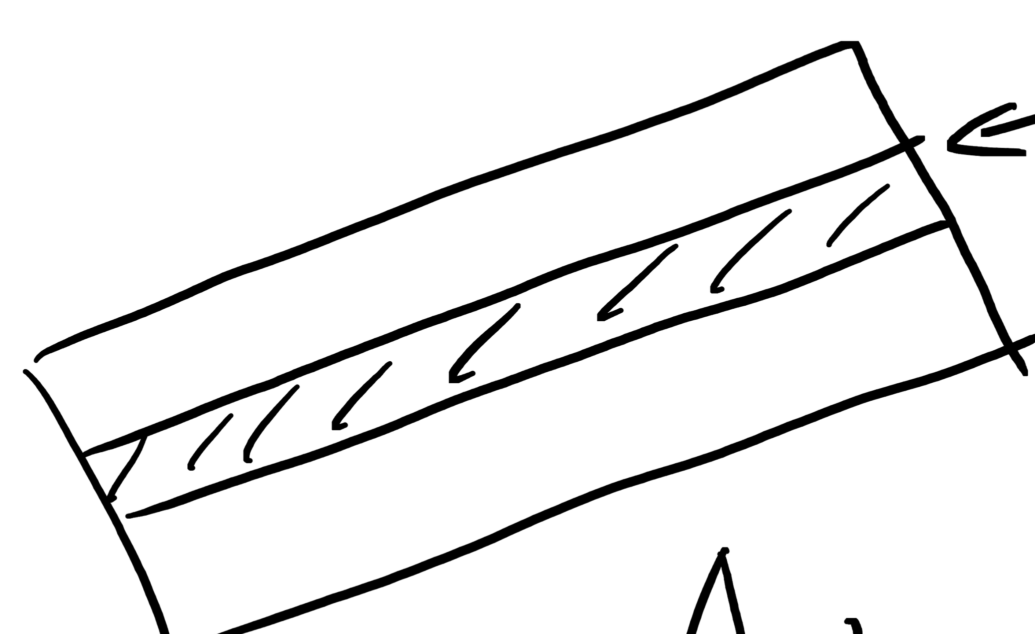


(see also Brin-Stuck, p. 136 for a better picture)

Now, note that each $\varphi(\Delta_j)$ is again going to be a rectangle and we can cut it into several rectangles, each of which lies entirely in one of Δ_j's:
contracting the stable direction,
expanding the unstable direction,
and then chopping into pieces across the stable boundaries of the original A, B.

For all $\alpha, \beta \in \{1, \dots, 5\}$, either

① $\varphi(\Delta_\beta) \cap \Delta_\alpha$ is a rectangle which spans the entire unstable length of Δ_α:



$\varphi(\Delta_\beta)$,

Δ_α

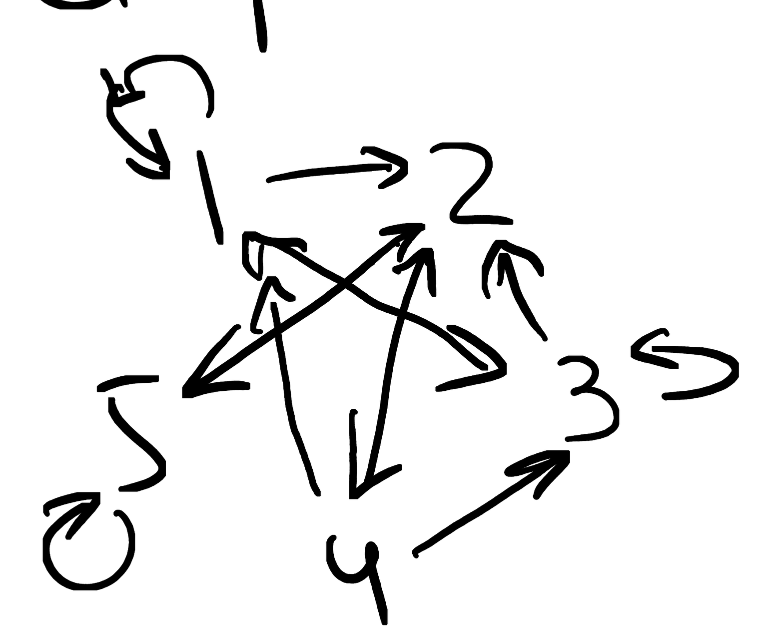
OR

② $\varphi(\Delta_\beta)$ does not overlap Δ_α.

In fact, case (1) holds iff $M_{\alpha\beta} = 1$ where M is the adjacency matrix

$$M = \begin{array}{c|ccccc} \alpha \backslash \beta & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 1 \\ 3 & 1 & 1 & 1 & 0 & 0 \\ 4 & 1 & 1 & 1 & 0 & 0 \\ 5 & 0 & 0 & 0 & 1 & 1 \end{array}$$

Graph:



<u>Note:</u> the largest (in absolute value) eigenvalue of M is exactly $\lambda = \frac{3+\sqrt{5}}{2}$

This is not a coincidence.

Let us now look at the refined partition $\xi_{\varphi^{-1}}^{(n)}$ where $\xi = \{\Delta_1, \dots, \Delta_5\}$

It consists of sets

$$\Delta_{\alpha_0 \alpha_1 \dots \alpha_{n-1}} = \Delta_{\alpha_0} \cap \varphi(\Delta_{\alpha_1}) \cap \dots \cap \varphi^{n-1}(\Delta_{\alpha_{n-1}}).$$

18.118 11-25

For each given $\alpha_0 \dots \alpha_{n-1} \in \{1, \dots, 5\}^n$, we have 2 cases:

① $\alpha_0 \dots \alpha_{n-1}$ is a path in the graph, i.e. $\forall\, j = 0, \dots, n-2$, we have $A_{\alpha_j \alpha_{j+1}} = 1$.

Then $\Delta_{\alpha_0 \dots \alpha_{n-1}}$ is a rectangle spanning the entire unstable length of Δ_{α_0}.
This can be seen by induction on n:

$$\Delta_{\alpha_0 \dots \alpha_{n-1}} = \Delta_{\alpha_0} \cap \varphi(\Delta_{\alpha_1 \dots \alpha_{n-1}})$$

② If $\alpha_0 \dots \alpha_{n-1}$ is not a path on the graph, then $\Delta_{\alpha_0 \dots \alpha_{n-1}}$ is negligible (measure 0, empty interior etc.)

Let μ = Lebesgue measure on $X = \mathbb{T}^2$.

In case 2 we have

$\mu(\Delta_{\alpha_0 \dots \alpha_{n-1}}) = 0.$

In case 1 we have:

the stable width of $\Delta_{\alpha_0 \dots \alpha_{n-1}}$ is $\sim \lambda^{-n}$ (each time the stable width is multiplied by λ^{-1})

the unstable width is ~ 1 (one of the widths of A, B)

So $\mu(\Delta_{\alpha_0 \dots \alpha_{n-1}}) \sim \lambda^{-n}$.

This is compatible with the fact that $\xi^{(n)}_{\varphi^{-1}} = \{\Delta_{\alpha_0 \dots \alpha_{n-1}}\}$ is a partition and the # of paths of length n in the graph grows like λ^n.

Now we can compute the entropy: (can show ξ is a generator...)

$H_\mu(\xi^{(n)}_{\varphi^{-1}}) \sim n \log \lambda$, so

$h_\mu(\varphi) = h_\mu(\varphi^{-1}) = h_\mu(\varphi^{-1}, \xi) = \log \lambda.$