CHAPTER I BASIC DEFINITIONS AND CONSTRUCTIONS

I.1 MOTIVATION

1.1 LAW OF THE LARGE NUMBERS

$X_1, X_2, \dots : \Omega \to \mathbb{R}$ iid random variables ($\mathbb{E}[X_1] < \infty$), then

- $\lim_{n\to\infty} \mathbb{P}\left[\left| \frac{1}{n}\sum_{k=1}^{n} X_k - \mathbb{E}[X_1] \right| > \varepsilon \right] = 0 \quad \forall \varepsilon > 0$ (weak)
- $\mathbb{P}\left[\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = \mathbb{E}[X_1] \right] = 1$ (strong)

Is independence necessary? $\to$ No!

1.2 A SIMPLE WEATHER MODEL

X_i : i^{th} day's weather $\in \{R, S\}$ "rainy"/"sunny"

$\mathbb{P}[X_{i+1} = b \mid X_i = a] = P_{ab} \in \mathbb{R}^{2\times 2}$, $a, b \in \{R, S\}$

Let

$$P = \begin{array}{cc} \begin{bmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix} & \begin{array}{c} R \\ S \end{array} \\ \begin{array}{cc} R & \quad S \end{array} & \end{array}$$

$\{X_i\}_{i\geq 0}$ is a Markov Chain

$$\mathbb{P}_S[X_i = S] = [0 \;\; 1]\, P^i \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

"$X_0 = S$"

Eigendecomposition: $M = \begin{bmatrix} \overbrace{1/3 \quad 2/3}^{v^T} \\ \underbrace{-1/2 \quad 1/2}_{w^T} \end{bmatrix}$, $MPM^{-1} = \begin{bmatrix} \overset{\lambda_1}{1} & \\ & \underset{\lambda_2}{0.4} \end{bmatrix}$

v, w : left eigenvectors

1.3 LONG-TERM BEHAVIOR

$$\mathbb{E}\left[\frac{\#\text{sunny days in the first } m \text{ days}}{m}\right] = \quad \begin{matrix} =: W_m^S \text{ for sunny} \\ W_m^R \text{ for rainy} \end{matrix}$$

$$= \frac{1}{m}\sum_{i=0}^{m-1} \mathbb{P}_f[X_i = S] \qquad f \in \mathbb{R}^2 : \text{ initial distribution}$$

$$= \frac{1}{m}\sum_{i=0}^{m-1} f^T P^i \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$= f^T M^{-1} \frac{1}{m}\sum_{i=0}^{m-1} \begin{bmatrix} 1 & \\ & 0.4 \end{bmatrix}^i M \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\xrightarrow[m\to\infty]{} f^T \begin{bmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 2/3 = \nu_2 = \nu_S = \nu_{\text{"sunny"}}$$

Observations:

(a) $\mathbb{E}[W_m^{S/R}] \xrightarrow[m\to\infty]{} \nu_{S/R}$

regardless the initial condition

(b) It holds even ("strong law"):

$$W_m^{S/R} \xrightarrow[m\to\infty]{\text{a.s.}} \nu_{S/R}$$

(c) $f \in \mathbb{R}^2_{\geq 0}, \|f\|_1 = 1 : \quad f^T P^m \xrightarrow[m\to\infty]{} \nu^T$ ("limiting distribution")

Is there something similar for deterministic systems as well?

1.4 DYNAMICAL SYSTEMS

Let $(G,+)$ be an (additive) group or semigroup, X some set, and $(T^g)_{g\in G}$ a family of maps $T^g : X \to X$, such that

(i) $T^{g_1} \circ T^{g_2} = T^{g_1+g_2} \qquad \forall g_{1,2} \in G$

(ii) $T^0 = id$

Then $(X, T^\bullet)$ is called a dynamical system with state space X, and $T^\bullet$ is called the flow.

- $G = \mathbb{N}_0$ or $\mathbb{Z}$: discrete time system, here $T := T^1$
- $G = \mathbb{R}_{\geq 0}$ or $\mathbb{R}$: continuous time system

1.5 THE STUDY OF DYNAMICAL SYSTEMS

(a) Differentiable dynamics : T differentiable, X smooth manifold

Local stretching $(\|DT\|)$, Lyapunov exponents, ...

$\leadsto$ see eg. [Ma]

(b) Topological dynamics : T continuous, X topological space

E.g. "transitivity": U, V open sets; is there a $n \geq 0$ such that

$$T^{-n}(U) \cap V \neq \emptyset \; ?$$

$\leadsto$ see e.g. [BS]

(c) Ergodic theory : T measurable, X measure space

E.g. long-term behavior as above

1.6 PREVIEW

Poincaré's Recurrence Theorem (PRT)

For almost all $x \in X$ (measured by the Lebesgue measure), the system starting at x will return arbitrary close to x at arbitrary large times.

Question: What other dynamical information can be extracted from measure-theoretic properties of a system?

The Ergodic Hypothesis

For certain measures μ, many functions $f: X \to \mathbb{R}$, and many states $x \in X$:

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(T^s x)\,ds = \frac{1}{\mu(X)}\int_X f(x)\,d\mu(x) = \text{"}\mathbb{E}_\mu[f]\text{"}$$

"time-average = space-average"; cf. law of the large numbers

Note: the ergodic hypothesis is a quantitative version of the PRT: with $f = \chi_{B_\varepsilon(x)}$ (characteristic function of an ε-ball) we get the relative time the system spends in $B_\varepsilon(x)$.

I.2 THE SETUP OF ERGODIC THEORY

1.7 MEASURE PRESERVING TRANSFORMATIONS

Def A: A measure space is a triplet $(X, \mathcal{B}, \mu)$ where

(i) X is a set (also called "space")

(ii) $\mathcal{B}$ is a σ-algebra (collection of subsets of X containing $\emptyset$ and being closed under complements and countable unions)

$A \in \mathcal{B}$ is called measurable

(iii) $\mu: \mathcal{B} \to [0,\infty]$ is a measure, i.e. a σ-additive function on $\mathcal{B}$.

$\mu(X) = 1 \rightsquigarrow$ "probability space"

Def B: A measure preserving transformation (mpt) is a quadruple $(X, \mathcal{B}, \mu, T)$ where $(X, \mathcal{B}, \mu)$ is a measure space, and

(i) T is measurable, i.e. $E \in \mathcal{B} \Rightarrow T^{-1}E := T^{-1}(E) \in \mathcal{B}$

(ii) μ is invariant, i.e. $\mu(T^{-1}E) = \mu(E) \quad \forall E \in \mathcal{B}$

A probability preserving transformation (ppt) is a mpt on a prob. space.

Ex:

$X = S^1$ (circle with unit circumference), $Tx = x + \alpha \mod 1$, where $\alpha \in \mathbb{R}$, μ is the Lebesgue measure

Def C: Two mpt's $(X_i, \mathcal{B}_i, \mu_i, T_i)$ are called measure-theoretically isomorphic if there is a measurable map $\pi : X_1 \to X_2$:

(i) there are $X_i' \in \mathcal{B}_i$ s.t. $\mu_i(X_i \setminus X_i') = 0$ and $\pi : X_1' \to X_2'$ is invertible with measurable inverse ("essentially a bijection")

(ii) for every $E \in \mathcal{B}_2$, $\pi^{-1}(E) \in \mathcal{B}_1$ and $\mu_1(\pi^{-1}E) = \mu_2(E)$

(iii) $T_2 \circ \pi = \pi \circ T_1$ on X_1

$$\begin{array}{ccc} X_1 & \xrightarrow{T_1} & X_1 \\ \pi \downarrow & & \downarrow \pi \\ X_2 & \xrightarrow{T_2} & X_2 \end{array}$$

diagram commutes

1.8 POINCARÉ'S RECURRENCE THEOREM

Thm: Let $(X, \mathcal{B}, \mu, T)$ be a ppt. If $E \in \mathcal{B}$ with $\mu(E) > 0$ then for almost every $x \in E$ there is a sequence $n_k \to \infty$ s.t. $T^{n_k}(x) \in E$.

Proof ([BG, Thm 3.2.1], extended)

Let $N_0 \subseteq E$ the set of points that never return to E, i.e.

$$N_0 := \{ x \in E \mid T^k(x) \notin E, \ k \geq 1 \}.$$

(a) We show that $T^{-i}(N_0) \cap T^{-j}(N_0) = \phi \quad \forall \; i > j \geq 0$. If $x \in X$ s.t. there are $i > j$ with $T^i(x) \in N_0 \ni T^j(x)$, then

$$T^{i-j}(\underbrace{T^j(x)}_{=: y \in N_0}) = T^i(x) \in N_0 \implies y \in N_0 \text{ is recurrent} \Rightarrow y \notin N_0 \quad \lightning$$

(b) N_0 is measurable, since $E \in \mathcal{A}$, T measurable, and

$$N_0 = \bigcap_{k=1}^{\infty} T^{-k}(E^c)$$

From ppt assumption & (b) $\Rightarrow \mu(T^{-i}(N_0)) = \mu(N_0) \; \forall i \geq 0$

Thus, from (a) $\Rightarrow$

$$1 = \mu(X) \geq \mu\left(\bigcup_{k=0}^{\infty} T^{-k}(N_0)\right) \overset{(a)}{=} \sum_{k=0}^{\infty} \mu(T^{-k}(N_0))$$
$$= \sum_{k=0}^{\infty} \mu(N_0)$$

Hence, $\mu(N_0) = 0$, and thus almost every point in E returns at least once to E: let $E_1 := E \setminus N_0$. Let N_i be the set that never returns to E_i, $i \geq 1$, and $E_{i+1} := E_i \setminus N_i$. We show iteratively that $\mu(N_i) = 0 \; \forall i \geq 0$.

The recurrent set, we want, is

$$E_\infty := E \setminus \bigcup_{i=0}^{\infty} N_i,$$

clearly measurable, and

$$\mu(E) \geq \mu(E_\infty) = \mu(E) - \mu\left(\bigcup_{i=0}^{\infty} N_i\right) \geq \mu(E) - \underbrace{\sum_{i=0}^{\infty} \mu(N_i)}_{< 0}$$

$\blacksquare$

1.9 REMARKS TO THE PRT

(a) The finiteness of μ is essential, as the proof shows.
E.g. take $X = \mathbb{Z}$, $\mathcal{B} = 2^X$, $T(x) = x+1$, and

$$\mu(E) = |E|, \text{ the counting measure}$$

Then, μ is invariant, but recurrence is clearly violated.

(b) $X = [0,1]$, $T(x) = 10x \pmod 1$

T leaves the Lebesgue measure invariant

PRT $\Longrightarrow$ for almost all $X \ni x = 0.\varepsilon_1\varepsilon_2\varepsilon_3\ldots\varepsilon_n\ldots$
the sequence $\varepsilon_1\varepsilon_2\ldots\varepsilon_n$ appears infinitely many times in the decimal expansion.

1.10 INVARIANT SETS

Def: $E \in \mathcal{B}$ is called *invariant*, if $T^{-1}(E) = E$.

Note: If E is invariant:

(i) so is $E^c := X \setminus E$;

(ii) $T: E \to E$ and $T: E^c \to E^c$ can be considered separately, they do not interact.

1.11 ERGODICITY

Def: A mpt $(X, \mathcal{B}, \mu, T)$ is *ergodic*, if every invariant set E satisfies $\mu(E) = 0$ or $\mu(E^c) = 0$. We call μ an ergodic measure.

"The system is essentially indecomposable"

Thm: For a mpt the following are equivalent:

(a) μ is ergodic

(b) if $E \in \mathcal{B}$ and $\mu(T^{-1}E \,\Delta\, E) = 0$, then $\mu(E) = 0$ or $\mu(E^c) = 0$

(c) if $f : X \to \mathbb{R}$ is measurable and $f \circ T = f$ a.e., then $\exists\, c \in \mathbb{R}$ s.t. $f = c$ a.e.

Remarks:

- $A \,\Delta\, B := (A \setminus B) \cup (B \setminus A)$ is called the symmetric difference
- A set E with $\mu(T^{-1}E \,\Delta\, E) = 0$ is called essentially T-invariant
- Part (c) of the Thm can be read as "a.e. trajectory of T goes almost everywhere in X"

Proof:

"(a) $\Rightarrow$ (b)": Let $E \in \mathcal{B}$ with $\mu(E \,\Delta\, T^{-1}E) = 0$. We'll construct an invariant $E_0 \in \mathcal{B}$ "close to E". Let

$$E_0 = \{ x \in X \mid T^k(x) \in E \text{ for infinitely many } k \in \mathbb{N} \}$$

E_0 is measurable and invariant (exercise). Now, to closeness:

- $x \in E_0 \setminus E$, then $\exists k : T^k(x) \in E$, but $x \notin E$, i.e. $x \in T^{-k}E \setminus E$
- $x \in E \setminus E_0$, then $\exists k : T^k(x) \notin E$, but $x \in E$, i.e. $x \in E \setminus T^{-k}E$

$\Longrightarrow$ $E \,\Delta\, E_0 \subset \bigcup_{k \geq 1} E \,\Delta\, T^{-k}E$. By $\underbrace{\mu(A \Delta C) \leq \mu(A \Delta B) + \mu(B \Delta C)}_{\circledast \text{ (prove this!)}}$

$$\begin{aligned}
\mu(E \Delta E_0) &\leq \mu\Big(\bigcup_{k\geq 1} E \,\Delta\, T^{-k}E\Big) \\
&\leq \sum_{k \geq 1} \mu(E \,\Delta\, T^{-k}E) \\
&\overset{\circledast}{\leq} \sum_{k \geq 1} \sum_{\ell=1}^{k} \mu(T^{-\ell+1}E \,\Delta\, T^{-\ell}E) \\
&\overset{\text{mpt}}{=} \sum_{k \geq 1} \sum_{\ell=1}^{k} \underbrace{\mu(E \,\Delta\, T^{-1}E)}_{=0} = 0
\end{aligned}$$

If μ ergodic, then $\mu(E_0) = 0$ or $\mu(E_0^c) = 0$. Since $\mu(E \Delta E_0) = 0$, it follows $\mu(E) = 0$ or $\mu(E^c) = 0$.

"(b) ⇒ (c)": Let f be measurable s.t. $f = f \circ T$ a.e. For $t \in \mathbb{R}$, $E := \{f > t\}$ is measurable, and we have

$$T^{-1}E = \{x \in X \mid f(Tx) > t\} \overset{a.e.}{=} \{x \in X \mid f(x) > c\} = E,$$

i.e. $\mu(E \Delta T^{-1}E) = 0$. By assumption $\mu(\{f > t\}) = 0$ or $\mu(\{f \leq t\}) = 0$. Assume $f \neq$ const a.e. Then $\exists t_*$, s.t. $\mu(\{f > t^*\}) > 0$ and $\mu(\{f \leq t^*\}) > 0$, in contradiction with what we just obtained.

"(c) ⇒ (a)": Let $E \in \mathcal{B}$ be an invariant set. Then $f = \chi_E$ satisfies

$$f \circ T = \chi_E \circ T = \chi_{T^{-1}E} = \chi_E = f,$$

hence $f =$ const a.e. Thus $\mu(E) = 0$ or $\mu(E^c) = 0$, and ergodicity follows.

1.12 ERGODICITY: INVARIANT UNDER MEASURE-THEORETIC ISOMORPHISM

Prop: Assume that the mpts $(X_i, \mathcal{B}_i, \mu_i, T_i)$, $i = 1, 2$, are meas. th. isomorph. Then, if one is ergodic, so is the other.

Proof: Suppose the mpt #1 is ergodic.
Let $f: X_2 \to \mathbb{R}$ be measurable and s.t. $f \circ T_2 = f$ a.e.
Then $f \circ T_2 \circ \pi = f \circ \pi$ a.e.

$$f \circ T_2 \circ \pi = f \circ \pi \circ T_1,$$

so by Thm 1.11: $f \circ \pi =$ const a.e.
Since π is an essential bijection, $f =$ const a.e.
$\overset{\text{Thm 1.11}}{\Longrightarrow}$ mpt #2 is ergodic ∎

1.13 INDEPENDENCE AND MIXING

- $(X, \mathcal{B}, \mu)$ prob. space
- $E \in \mathcal{B}$: events
- $\mathbb{P}[x \in E] = \mu(E)$: probability law

$E, F \in \mathcal{B}$ are independent, if

$$\mu(E) = \underbrace{\mu(E|F)}_{\text{conditional prob.}} := \frac{\mu(E \cap F)}{\mu(F)} \iff \mu(E \cap F) = \mu(E)\mu(F)$$

<u>Def</u>: A ppt $(X, \mathcal{B}, \mu, T)$ is called (strongly) mixing if for every $E, F \in \mathcal{B}$ we have $\mu(E \cap T^{-k}F) \xrightarrow[k\to\infty]{} \mu(E)\mu(F)$

"$T^{-k}(F)$ is asymptotically indipent of E" or

"$T^k(x)$ looses dependence on x for large k"

<u>Exercise</u>: Mixing is an invariant under measure-theoretic isom.

<u>PropA</u>: Strong mixing implies ergodicity.

<u>Proof</u>: $E \in \mathcal{B}$ invariant: $E = T^{-k}(E)$ $\forall k \geq 0$

$$\rightsquigarrow \mu(E) = \mu(E \cap T^{-k}E) \xrightarrow[k\to\infty]{} \mu(E)^2 \Rightarrow \mu(E) \in \{0, \underset{\mu(X)}{\overset{}{1}}\}$$

$\blacksquare$

<u>PropB</u>: Strong mixing $\iff$ $\forall f, g \in L^2(X, \mu)$:

$$\int f \cdot g \circ T^k \, d\mu \xrightarrow[k\to\infty]{} \int f \, d\mu \cdot \int g \, d\mu$$

<u>Proof</u>: "$\Leftarrow$": Take $f = \chi_E$, $g = \chi_F$

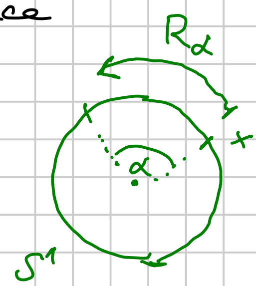
"$\Rightarrow$": [Sa, Prop 1.3]

$\blacksquare$

I.3 EXAMPLES

1.14 CIRCLE ROTATION

$\alpha \in \mathbb{R}$: $\quad R_\alpha : S^1 \to S^1,\quad x \mapsto x+\alpha \mod 1$ $\qquad S^1$: circle with unit circumference



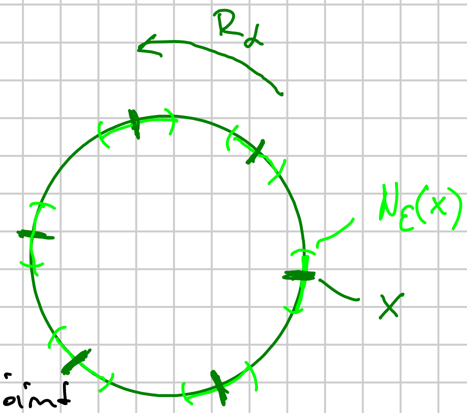
$\mathcal{B}$: Borel σ-alg, $\mu = m$: Lebesgue measure

<u>Prop</u>: (a) R_α is measure preserving
(b) R_α is ergodic iff $\alpha \notin \mathbb{Q}$
(c) R_α is never mixing

<u>Proof</u>: (a) $\rightsquigarrow$ Exercise

(b) • $\alpha = \frac{p}{q} \in \mathbb{Q}$ $\quad (p,q \in \mathbb{N})$

$\Rightarrow R_\alpha^q = id$



Pick $0<\varepsilon< 1/2q$

$\Rightarrow N_\varepsilon(x), N_\varepsilon(x+\frac{1}{q}), \dots, N_\varepsilon(x+\frac{q-1}{q})$ disjoint

"ε-neighborhood of x"

$\Rightarrow E = \bigcup_{i=1}^{q-1} N_\varepsilon(x+k\frac{1}{q})$ is invariant and $0<m(E)<1$

$\Rightarrow R_\alpha$ not ergodic

• $\alpha \notin \mathbb{Q}$. Let $f = \chi_E$, E invariant $\Rightarrow f \circ R_\alpha = f$

Fourier series: $f(x) = \sum_{n\in\mathbb{Z}} c_n e^{2\pi i n x}$ $\quad$ (in L^2)

$$\rightsquigarrow \sum_{n\in\mathbb{Z}} c_n e^{2\pi i n x} = f(x) = f(R_\alpha x) = \sum_{n\in\mathbb{Z}} c_n e^{2\pi i n(x+\alpha)}$$

$$\Leftrightarrow c_n = 0 \ \vee\ e^{2\pi i n \alpha} = 1 \quad \forall n \neq 0$$

$$\Leftrightarrow c_n = 0 \ \forall n\neq 0 \Rightarrow f = \text{const} \Rightarrow m(E) \in \{0,1\}$$

$$\Rightarrow R_\alpha \text{ ergodic}$$

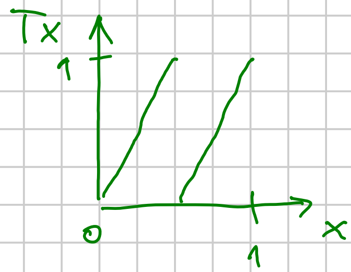
(c) Set $E=(0,\varepsilon)=F$, $0<\varepsilon<1$

Dirichlet thm: $\forall \alpha\in\mathbb{R}\ \exists m_k \xrightarrow[k\to\infty]{} \infty$ s.t. $m_k\alpha \bmod 1 \to 0$ $(k\to\infty)$

$\Rightarrow m(E\cap R_\alpha^{-m_k}F) \xrightarrow[k\to\infty]{} m(E)=\varepsilon \neq \varepsilon^2 = m(E)\,m(F)$

$\Rightarrow R_\alpha$ not mixing

1.15 ANGLE DOUBLING



$X=S^1$, $\mu=m$, $T(x)=2x \bmod 1$

<u>Prop</u>: T is measure preserving and strongly mixing

<u>Proof</u>: Exercise.

1.16 BERNOULLI SCHEMES

- S — alphabet (finite set)
- $X=S^{\mathbb{N}} \ni \{x_k\}=(x_0,x_1,x_2,\dots)$
- $d(\{x_k\},\{y_k\}) = 2^{-\min\{k \mid x_k\neq y_k\}}$ — metric

$\rightsquigarrow$ topology generated by cylinders

$$[a_0,a_1,\dots,a_{n-1}] = \{\{x_k\} \mid x_i=a_i \quad 0\le i\le n-1\}$$

Cylinders generate the Borel σ-alg w.r.t. metric d.

- $T: X\to X$, $(x_0,x_1,x_2,\dots)\longmapsto(x_1,x_2,\dots)$ — left shift
- $p=(p_a)_{a\in S}\in\mathbb{R}^{|S|}$ — probability vector (i.e. $p\geq 0$, $\sum_a p_a = 1$)

Def: The Bernoulli measure corresponding to p on $\mathcal{B}$ (Borel σ-alg. wrt d) is the unique measure μ, s.t.

$$\mu[a_0, \dots, a_{n-1}] = \prod_{i=0}^{n-1} p_{a_i}$$

Prop A: The $(\frac{1}{2}, \frac{1}{2})$-Bernoulli shift (i.e. $X = \{0,1\}$, $p = (\frac{1}{2}, \frac{1}{2})$, μ the corresp. Bernoulli measure, T the left shift) is measure theoretically isomorphic to the angle doubling map.

Proof: Let $\pi(x_0, x_1, \dots) = \sum_{i=0}^{\infty} 2^{-(i+1)} x_i$

$\pi: X' \to S^1$ is bijective, where

$$X' = \{\{x_k\} \in X \mid \nexists n \in \mathbb{N} \text{ s.t. } x_\ell = 1 \ \forall \ell \geq n\}$$

and $\mu(X') = \mu(X) = 1$ (exercise)

Clearly $\pi \circ T = \tilde{T} \circ \pi$ with $\tilde{T}: S^1 \to S^1, x \mapsto 2x \bmod 1$

Check that $m(\pi[a_0, \dots, a_{n-1}]) = 2^{-n} = \mu[a_0, \dots, a_{n-1}]$

$\overset{\text{Mon. class Thm.}}{\Longrightarrow}$ π preserves the measure $\overset{\text{Def. 1.7C}}{\Longrightarrow}$ claim

cylinders span $\mathcal{B}(X)$
+
dyadic intervals span $\mathcal{B}(S^1)$

$\blacksquare$

Prop B: Every Bernoulli scheme is mixing, hence ergodic.

Proof: The proof goes as in the case of the angle doubling map.
For the $(\frac{1}{2}, \frac{1}{2})$-Bernoulli shift we could use Prop 1.15 + Prop A (here) + the invariance of mixing under isomorphism to get the claim. $\blacksquare$

$\left(\frac{1}{2},\frac{1}{2}\right)$-Bernoulli shift — Coin toss

$\{x_k\} \in X$ sequence of events $Y_0, Y_1, \ldots$ $Y_i \in \{0, 1\}$ heads, tails

$\mu[a_0, \ldots, a_{n-1}] = 2^{-n}$ $\mathbb{P}[Y_0 = a_0, \ldots, Y_{n-1} = a_{n-1}] = 2^{-n}$

Invariance: $\mu = \mu \circ T^{-1}$

$\Rightarrow \mu[a_0] = \mu[*, a_0]$

$$\mathbb{P}[Y_0 = a_0] = \mathbb{P}[Y_0 = *, Y_1 = a_0] = \mathbb{P}[Y_1 = a_0]$$

independence of the Y_i

1.17 SUBSHIFT OF FINITE TYPE

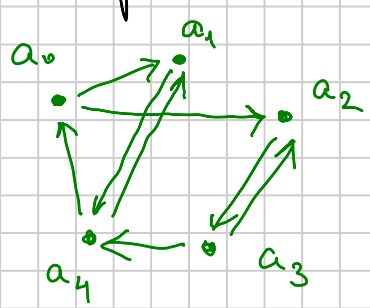
S — alphabet (finite), $A = \{t_{ij}\}_{i,j \in S} \in \{0,1\}^{|S| \times |S|}$

Def: The subshift of finite type (sft) with alphabet S and transition matrix A is:

- $\Sigma_A^+ = \left\{ \{x_k\} \in S^{\mathbb{N}} \;\middle|\; t_{x_i x_{i+1}} = 1 \;\; \forall i \geq 0 \right\}$, and
- metric $d(\{x_k\}, \{y_k\}) = 2^{-\min\{k \mid x_k \neq y_k\}}$, and
- the action $T(x_0, x_1, x_2, \ldots) = (x_1, x_2, \ldots)$, the left shift

The sft is a compact metric space, and $T: \Sigma_A^+ \to \Sigma_A^+$ is continuous

Idea:



$a_i \in S$, $i = 1, \ldots, |S|$

Edge from a_i to a_j iff $t_{a_i a_j} = 1$

$\Rightarrow \Sigma_A^+$: the set of all possible paths on this graph

Further terminology:

- Stochastic matrix: $P = (p_{ab})_{a,b \in S} \in [0,1]^{|S| \times |S|}$, $P\mathbb{1} = \mathbb{1} = (1, \ldots, 1)^T$

  compatible with A: $t_{ab} = 0 \Rightarrow p_{ab} = 0$

- Probability vector: $p = (p_a)_{a \in S} \in [0,1]^{|S|}$ s.t. $\sum p_a = 1$
- Stationary probability vector: p s.t. $p^T P = p^T$

1.18 Markov shift

Def: Given: prob. vector p, stoch. matrix P compatible with A.
The Markov measure on Σ_A^+ (with the Borel σ-alg $\mathcal{B}(\Sigma_A^+)$) is defined through

$$\mu[a_0, a_1, \ldots, a_{n-1}] = p_{a_0} \cdot p_{a_0 a_1} \cdot \ldots \cdot p_{a_{n-2} a_{n-1}}$$

(Remark: P stochastic $\Rightarrow \mu$ is well-defined Borel prob. measure)

Then, $(\Sigma_A^+, \mathcal{B}(\Sigma_A^+), \mu, T)$ is called the (p,P)-Markov shift.
$\uparrow$ left shift

Prop: (a) p stationary wrt $P \iff (p,P)$-Markov shift is a ppt

(b) Every stoch. matrix has a stat. prob. vector

Proof: (a) Sufficient to show for cylinders: $\mu[\underline{b}] = \mu(T^{-1}[\underline{b}]) = \mu[*, \underline{b}]$:

$$\sum_a p_a p_{ab_0} \cdot p_{b_0 b_1} \cdot \ldots \cdot p_{b_{n-2} b_{n-1}} = p_{b_0} p_{b_0 b_1} \cdot \ldots \cdot p_{b_{n-2} b_{n-1}}$$

no division by 0 due to compatibility

$$\iff \sum_a p_a p_{ab_0} = p_{b_0}$$

(b) Let $\Delta = \{ x \in \mathbb{R}^{|S|} \mid x \geq 0, \sum x_i = 1 \}$, and

$$R: \Delta \to \mathbb{R}^{|S|}, \quad x \longmapsto (x^T P)^T = P^T x$$

Check that $R(\Delta) \subseteq \Delta$ due to stochasticity of P
Brouwer's fixed point theorem (continuous mapping on a convex, compact set has a fixed point) $\Rightarrow R$ has a fixpt ∎

Ergodicity & mixing of Markov shifts $\rightsquigarrow$ later

I.4 BASIC CONSTRUCTIONS

1.19 PRODUCTS

$(X_i, \mathcal{B}_i, \mu_i)$, $i = 1,2$, measure spaces

$(X_1 \times X_2, \mathcal{B}_1 \otimes \mathcal{B}_2, \mu_1 \times \mu_2)$ — product (measure) space

- $\mathcal{B}_1 \otimes \mathcal{B}_2 = \sigma\left(\{B_1 \times B_2 \mid B_i \in \mathcal{B}_i\}\right)$
- $(\mu_1 \times \mu_2)(B_1 \times B_2) := \mu_1(B_1)\mu_2(B_2)$ (unique!)

Def: The product of two mpts $(X_i, \mathcal{B}_i, \mu_i, T_i)$, $i = 1,2$, is the mpt $(X_1 \times X_2, \mathcal{B}_1 \otimes \mathcal{B}_2, \mu_1 \times \mu_2, T_1 \times T_2)$ with

$$(T_1 \times T_2)(x_1, x_2) = (T_1 x_1, T_2 x_2)$$

Prop: The product of two ergodic mpts is not necessarily ergodic. The product of two mixing mpts is always mixing.

Proof: The product of two (identical) ergodic circle rotations:

$$\underset{\substack{\| \\ R_\alpha \times R_\alpha}}{T} : S^1 \times S^1 \to S^1 \times S^1, \quad (x_1, x_2) \longmapsto (x_1 + \alpha, x_2 + \alpha) \bmod 1$$

is not ergodic: $f(x,y) := x - y \bmod 1$ is a non-constant invariant function (cf. Thm. 1.11)

Mixing: see [Sa, Prop. 1.9]

1.20 SKEW-PRODUCTS

Example: $(\Sigma, \mathcal{B}, \mu, \sigma)$: $(\frac{1}{2}, \frac{1}{2})$-Bernoulli shift $\quad (\Sigma = \{0,1\}^{\mathbb{N}})$

$f : \Sigma \to \mathbb{Z}$, $f(\{x_k\}) = (-1)^{x_0}$

Let : $T : \Sigma \times \mathbb{Z} \to \Sigma \times \mathbb{Z}$, $T(\{x_k\}, \ell) = (\sigma\{x_k\}, \ell + f(\{x_k\}))$

T is mpt wrt $\mu \times m_{\mathbb{Z}}$ ($m_{\mathbb{Z}}$: counting measure), and

$$T^n(\{x_k\}, \ell) = \left(\sigma^n\{x_k\}, \ell + Y_0 + Y_1 + \ldots + Y_{n-1}\right), \quad Y_i := (-1)^{x_i}$$

"random walk on $\mathbb{Z}$ driven by the noise process $(X, \mathcal{B}, \mu, \sigma)$"

<u>Def</u>: Let $(\Omega, \mathcal{F}, \rho, \sigma)$ be a mpt, $(X, \mathcal{B}, \mu)$ a prob. space, and $\{T_\omega\}_{\omega \in \Omega}$ a family of mpts on $(X, \mathcal{B}, \mu)$, s.t. $(\omega, x) \mapsto T_\omega x$ is measurable with respect to $\rho \times \mu$.

We define the skew-product as $(\Omega \times X, \mathcal{F} \otimes \mathcal{B}, \rho \times \mu, \tau)$ with

$$\tau(\omega, x) = (\sigma\omega, T_\omega x)$$

τ preserves $\rho \times \mu$ ⤳ exercise (use Fubini's theorem)

$$\tau^n(\omega, x) = \left(\sigma^n\omega, T_{\sigma^{n-1}\omega} \circ T_{\sigma^{n-2}\omega} \circ \ldots \circ T_{\sigma\omega} \circ T_\omega x\right)$$

⤳ skew-products model forced / driven / time-dependent / random dyn. systems

The skew-product is sometimes called a random dynamical system

$(\Omega, \mathcal{F}, \rho, \sigma)$ is also called base or driving system

<u>Example</u>: T_ω: change in the amount of ice in a glacier

ω: season ($\sigma^n\omega$ models planetary dynamics ⤳ amount of incoming solar radiation)

1.21 FACTORS AND EXTENSIONS

<u>Def</u>: A mpt $(X, \mathcal{B}, \mu, T)$ is called a factor of a mpt $(Y, \mathcal{C}, \nu, S)$ if $\exists X' \subseteq X$, $\exists Y' \subseteq Y$, both of full measure, s.t.

- $T(X') \subseteq X'$, $S(Y') \subseteq Y'$
- there is a $\pi: Y' \to X'$ measurable and onto (surjective), with $\nu \circ \pi^{-1} = \mu$ and $\pi \circ S = T \circ \pi$ on Y'

π is called the factor map. $(Y, \mathcal{C}, \nu, S)$ is called an extension of $(X, \mathcal{B}, \mu, T)$.

Examples

1) Meas.-theoretically isomorphic systems are factors and extensions of each other
2) A skew-product $T: \Omega \times X \circlearrowleft$ is an extension of its base $\sigma: X \to X$, if μ is a probability measure.
3) Assume, we can "measure" the state $x \in X$ of some mpt only through the observables $f: X \to \mathbb{R}$ (measurable). The dynamical information we obtain is encoded in the sub-σ-alg
$$\mathcal{B} := \sigma\left(f \circ T^n \mid n \geq 0\right) = \sigma\left\{T^{-n} \circ f^{-1}(A) \mid A \in \mathcal{B}(\mathbb{R})\right\}$$
(E.g. $f = \chi_A$ for some $A \in \mathcal{C}$, and then we observe as trajectories sequences in $\{0,1\}^{\mathbb{N}}$)
It can happen that $(X, \mathcal{B}, \mu, T)$ is mixing, while $(X, \mathcal{C}, \mu, T)$ isn't.