LECTURE NOTES ON "Selected topics in Dynamical Systems"

Andreas Strömbergsson

Contents

1. Introduction & ergodicity

Lecture #1: Introduction & ergodicity

Prerequisities: Measure theory (key!)

(Basic differential geometry; "manifolds")

(Lie groups)

I will try to intro/review concepts when they enter, but depending on your background you may have to accept a few "black boxes".

Literature: ....

Examination: ....

Let $(X, \mathcal{B}, \mu)$ be a measure space, i.e.

$$\begin{cases} X & - \text{ a set} \\ \mathcal{B} & - \text{ a } \sigma\text{-algebra on } X \\ \mu : \mathcal{B} \to [0,\infty] & \text{ a measure on } \underbrace{(X,\mathcal{B})}_{} \end{cases}$$

a measurable space

Also: $P(X) := \{\mu : \mu \text{ a probability measure on } (X,\mathcal{B})\}$

Recall def: A map $T: X \to X$ is measurable ("m'ble") if $\forall E \in \mathcal{B}: T^{-1}E \in \mathcal{B}$.

Caveat: For $T: X \to Y$ with Y a topological space, T m'ble $\overset{\text{def}}{\Longleftrightarrow}$ $\forall\, U$ open $\subset Y$: $T^{-1}U \in \mathcal{B}$.

Def 1 $(X, \mathcal{B}, \mu, T)$ is a <u>measure preserving transformation</u> (<u>mpt</u>) if $T : X \to X$ is m'ble and $\underbrace{\mu \text{ is } T\text{-invariant}}$

$$\overset{\text{def}}{\Longleftrightarrow} T_* \mu = \mu \quad \Leftrightarrow \left[\forall E \in \mathcal{B} : \mu(T^{-1}E) = \mu(E)\right]$$

Also: ~~p~~pt ... $\mu(X) = 1$.

We are here jumping directly to the set-up of <u>ergodic theory</u>. Often one starts from just X (with some structure, e.g. a mfld) and $T : X \circlearrowleft$ (e.g. cont, or C^∞), – a <u>dynamical system</u>, and finding a T-inv measure μ is a useful <u>tool</u>! Cf, e.g., <u>Poincaré's Recurrence Theorem</u>! An important problem: Characterize <u>all</u> T-inv μ!

<u>More general setting</u>: Let G be a <u>topological semigroup</u> with its Borel σ-algebra.

Def 2: A <u>measure-preserving G-action</u> on $(X, \mathcal{B}, \mu)$ is a m'ble map $\varphi : G \times X \to X$ s.t.

(1) $\forall g \in G$: $\phi_g := \varphi(g, \cdot)$ is m'ble and $\varphi_{g*}(\mu) = \mu$.

(2) $\phi_{g_1} \circ \phi_{g_2} = \phi_{g_1 g_2}$ $\quad \forall g_1, g_2 \in G$

(3) if $e \in G$: $\phi_e = 1_X$.

identity element, i.e. G a <u>monoid</u>

? – this is a <u>tentative</u> definition...

Here $\underline{G = \langle \mathbb{Z}_{\geq 0}, + \rangle}$ $\Leftrightarrow$ Def 1

~~[illegible]~~ $\boxed{\phi_n := T^n, \quad n \geq 0}$

$\underline{G = \langle \mathbb{Z}, + \rangle}$ $\Leftrightarrow$ an <u>invertible</u> mpt,

i.e. $(X, \mathcal{B}, \mu, T)$ is an mpt and T <u>bijective</u>,

T^{-1} m'ble $(\Rightarrow T_*^{-1}(\mu) = \mu)$ $\boxed{\phi_n := T^n, \quad n \in \mathbb{Z}}$

$\underline{G = \langle \mathbb{R}, + \rangle}$ $\Leftrightarrow$ a <u>flow</u> on $(X, \mathcal{B}, \mu)$

Often require: $(X, \mathcal{B}, \mu)$ is (the completion of) a <u>standard Borel space</u>.

Explain... Note: Taking completion or not is more or less equivalent, in Def 1!

In this course, 3 key examples:

① **"Homogeneous dynamics"**

Simple case: $X =$ torus $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$

$T : X \circlearrowleft \qquad T(\underline{x}) = \underline{x} + \underline{a} \qquad (\underline{a} \in \mathbb{R}^n \text{ fixed})$

$\mu =$ Lebesgue

– then $(X, \mathcal{B}, \mu, T)$ is an invertible ppt.

Also: flow $\{\phi_t\}_{t \in \mathbb{R}}$ on X; $\phi_t(\underline{x}) = \underline{x} + t\underline{a}$

Orbits:

General case: $X = \Gamma \backslash G$ where G – Lie group, Γ – a discrete subgp

$:= \{\Gamma h : h \in G\}$

s.t. X admits a G-inv. prob measure μ.

G-action on X: $\phi_g(\Gamma h) = \Gamma h g \quad \forall g, h \in G$

(This is a right action; $\phi_{g_1 g_2} = \phi_{g_2} \circ \phi_{g_1}$!)

Let $\{u_t\}_{t \in \mathbb{R}}$ be any 1-parameter subgroup of G.

Then $\phi_t(\Gamma h) := \Gamma h u_t$ is a flow on (X, μ).

E.g.: Geodesic flow on T^1 of any hyperbolic surface of finite area!

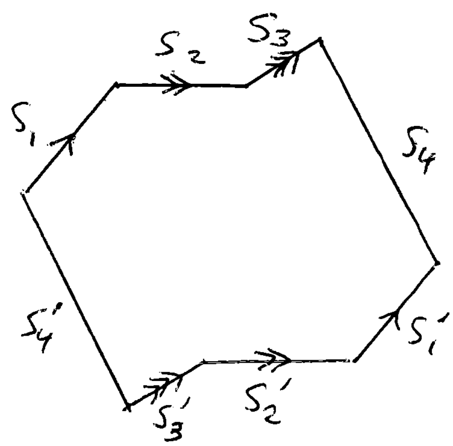
② Translation surfaces

<u>def</u> $\begin{cases} \text{a compact Riemann surface } M \text{ together with} \\ \text{a holomorphic 1-form } \alpha \ (\not\equiv 0) \text{ on } M. \end{cases}$

$=$ $\begin{cases} \text{a compact 2-dim mfld } M \text{ with a flat} \\ \text{Riemannian metric having conical singularities} \\ a_1, \dots, a_K \text{ with angles } 2\pi(m_i+1) \quad (i=1,\dots,K) \\ \qquad m_i \in \mathbb{Z}^+ \\ \text{together with a parallel unit vector field on} \\ M \setminus \{a_1,\dots,a_K\}. \end{cases}$

To see this, write $\alpha = dz$, or near a zero: $\alpha = z^m dz$.

<u>ex</u>:



$K=1,\ m_1=2$ ($\Rightarrow$ angle 6π)

$2-2g = \chi = -2$

i.e. $g=2$

$X :=$ the orbifold of <u>all</u> transl. surfaces with given $K, m_1, \dots, m_K$, and area 1.

Describe...

<u>$SL_2(\mathbb{R}) = \{T = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}) : \det T = 1\}$</u> <u>acts on X</u>!

$\exists$ "canonical" abs. cont. $SL_2(\mathbb{R})$-inv. measure μ on X.

Masur & Veech (80's) $\rightarrow$ μ <u>finite</u> (say $\mu(X)=1$)

$\left\{\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}\right\} \rightsquigarrow$ the "Teichmüller flow".

3 $N_s(\mathbb{R}^d) = \{A \subset \mathbb{R}^d : A \text{ locally finite}\}$

Topology? Identify $A \in N_s(\mathbb{R}^d) \longleftrightarrow$ measure $\sum_{\underline{x} \in A} \delta_{\underline{x}}$

Thus $N_s(\mathbb{R}^d) \subset \underbrace{N(\mathbb{R}^d)}_{\boxed{\text{counting Measures}}} \subset \underbrace{M(\mathbb{R}^d)}_{\boxed{\text{loc. finite Borel measures on } \mathbb{R}^d.}}$

Put the <u>vague topology</u> on $M(\mathbb{R}^d)$,

i.e. $\mu_n \to \mu$ iff $\underbrace{\mu_n(f)}_{\int_{\mathbb{R}^d} f \, d\mu_n} \to \mu(f)$, $\forall f \in C_c(\mathbb{R}^d)$.

but <u>not</u> loc. cpt!

$M(\mathbb{R}^d)$ <u>Polish</u>, $N(\mathbb{R}^d)$ closed ($N_s(\mathbb{R}^d)$ not closed).

$\mathbb{R}^d$ acts continuously on $N_s(\mathbb{R}^d)$ by translations:

$$\underline{x} + A := \{y + \underline{x} : y \in A\},$$

and $GL_d(\mathbb{R})$ acts continuously on $N_s(\mathbb{R}^d)$ by

$$TA := \{Ty : y \in A\}.$$

Now: An $\mathbb{R}^d$-inv Borel probability measure on $N_s(\mathbb{R}^d)$
$=$ a <u>stationary point process</u>!

<u>Ex</u>: A Poisson process in $\mathbb{R}^d$ with const. intensity!
This is both $\mathbb{R}^d$- and $SL_d(\mathbb{R})$-inv!
<u>OPEN PROBL</u>: Characterize all $[\mathbb{R}^d$- and $SL_d(\mathbb{R})]$-inv probability measures on $N_s(\mathbb{R}^d)$!
This type of problem can be seen as a vast generalization of certain key problems in homogeneous dynamics & Teichm dyn!

Some other examples

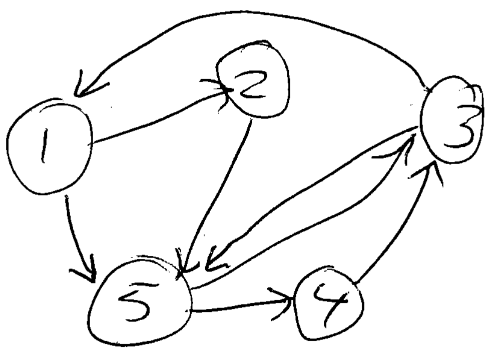
Bernoulli schemes

$X = S^{\mathbb{N}}$

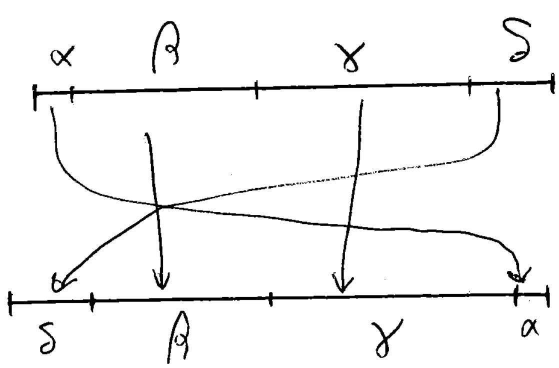
$T = \text{shift}$

more general

Markov chains



Interval Exchange Transformations (IETs)



We'll study IETs <u>a lot</u> in this course, as inroad to translation surfaces.

{Now back to general theory}

Ergodicity

$(X, \mathcal{B}, \mu, T)$ – an mpt.

We say $E \in \mathcal{B}$ is (T-)invariant if $T^{-1}(E) = E$.

Set $\underline{\mathrm{Inv}(T)} := \{E \in \mathcal{B} : T^{-1}(E) = E\}$

{a sub-σ-algebra of $\mathcal{B}$}

Def: $(X, \mathcal{B}, \mu, T)$ (or μ) is called ergodic if every $E \in \mathrm{Inv}(T)$ has $\mu(E) = 0$ or $\mu(X \setminus E) = 0$.

{Most often for us, $(X, \mathcal{B}, \mu, T)$ is a ppt here! The ergodic $\left[{m \atop p}\right]$pt's should be thought of as the "connected" or "irreducible" $\left[{m \atop p}\right]$pt's! Obviously, if μ is not ergodic, then we can split X into two disjoint mpt's!}

Note: If μ not ergodic, take $E \in \mathrm{Inv}(T)$ with $\mu(E) > 0$, $\mu(X \setminus E) > 0$; then

$$(E, \mathcal{B}_{|E}, \mu_{|E}, T_{|E}) \quad \text{and} \quad (X \setminus E, \mathcal{B}_{|X\setminus E}, \mu_{|X\setminus E}, T_{|X\setminus E})$$

are disjoint mpts. For ppt, get two disjoint ppt's; consider $\frac{1}{\mu(E)} \mu_{|E}$ and $\frac{1}{\mu(X \setminus E)} \cdot \mu_{|X\setminus E}$!

Def: T (or "$\langle X, \mathcal{B}, T\rangle$") is called uniquely ergodic if $\exists!\ \mu \in \mathcal{P}(X)$ s.t. μ is T-invariant. {Note: This μ must be ergodic!}

Prop 1: [μ ergodic]

$\overset{①}{\iff}$ $\Big[\forall E \in \mathcal{B}: \underbrace{E \text{ inv mod } 0}_{\text{i.e. } \mu(E \,\Delta\, T^{-1}E)=0} \Rightarrow \mu(E)=0 \text{ or } \mu(X \setminus E)=0\Big]$

$\overset{②}{\iff}$ $\Big[\forall f: X \to \mathbb{R} \text{ m'ble}: \underbrace{f \circ T = f \text{ a.e}}_{\text{i.e. } f(Tx)=f(x) \text{ for } \mu\text{-a.e. } x} \Rightarrow \underbrace{f = \text{const a.e.}}_{\text{i.e. } \exists c \in \mathbb{R}: f(x)=c,\ \mu\text{-a.e. } x}\Big]$

"proof": $\overset{①}{\Leftarrow}$ obvious

$\overset{②}{\Rightarrow}$ study sublevel sets of f.

$\overset{②}{\Leftarrow}$ (obvious) take $f = 1_E$

$\overset{①}{\Rightarrow}$ Consider $E_0 = \limsup_{k\to\infty} T^{-k}(E) = \bigcap_{n\geq 0} \bigcup_{k \geq n} T^{-k}(E)$

(or liminf)

Mixing

Def: A ppt (need μ a probability measure!) $(X, \mathcal{B}, \mu, T)$ is called ((strongly)) mixing if

$$\forall E, F \in \mathcal{B}: \quad \lim_{k\to\infty} \mu(E \cap T^{-k}F) = \mu(E)\mu(F)$$

mixing $\Rightarrow$ ergodicity !

mixing $\iff$ $\forall f, g \in L^2$: $\int_X f(x) g(T^k x)\, d\mu(x) \xrightarrow{k\to\infty} \int_X f\, d\mu \int_X g\, d\mu$

1.1. **Notes.** This lecture mainly covers some stuff from Sarig, [40, Ch. 1]. My Def. 1 is [40, Def. 1.2]. For Poincaré's Recurrence Theorem, see [40, Sec. 1.1].

Regarding my Def. 2, compare e.g. [25, p. 10, Def. 7] and [12, Ch. 8] (for the special case of G a group).

As I point out on p. 3 of the lecture, in the development of ergodic theory one often has to assume that $(X, \mathcal{B}, \mu)$ is (the completion of) a *standard Borel space.* See Sec. 5.2 in my notes to [40] for the relevant definitions and basic facts; note that I have chosen to use a somewhat different terminology that Sarig here. In particular I claimed in my lecture that if $(X, \mathcal{B})$ is a standard Borel space and μ is a probability measure on $(X, \mathcal{B})$ then it is "essentially equivalent" to specify an mpt T on $(X, \mathcal{B}, \mu)$ or on the completed space $(X, \mathcal{B}^{\mu}, \mu)$. See my notes regarding [40, Def. 1.2] for some remarks and lemmata making this precise.

In this connection, it may be worth emphasizing what is the definition of certain types of objects being *isomorphic:*

(1) Two measurable spaces $(X_1, \mathcal{B}_1)$ and $(X_2, \mathcal{B}_2)$ are said to be isomorphic if there exists a bijection $J : X_1 \to X_2$ such that both J and J^{-1} are measurable. (This is the natural definition; the conditions mean exactly that "J preserves all given structure".)

(2a) Two measure spaces $(X_1, \mathcal{B}_1, \mu_1)$ and $(X_2, \mathcal{B}_2, \mu_2)$ are said to be *strictly isomorphic* if there exists a measurable bijection $J : X_1 \to X_2$ such that also J^{-1} is measurable, and $J_*(\mu_1) = \mu_2$. (This is again the "natural definition". Note in particular that it follows that $J_*^{-1}(\mu_2) = \mu_1$.)

(2b) Two measure spaces $(X_1, \mathcal{B}_1, \mu_1)$ and $(X_2, \mathcal{B}_2, \mu_2)$ are said to be *almost isomorphic* (or *isomorphic mod 0*) if there exist $X_i' \in \mathcal{B}_i$ with $\mu_i(X_i \setminus X_i') = 0$ for $i = 1, 2$, such that $(X_1', \mathcal{B}_{1|X_1'}, \mu_{1|X_1'})$ and $(X_2', \mathcal{B}_{1|X_2'}, \mu_{1|X_2'})$ are strictly isomorphic.

(3) Two mpt's $(X_i, \mathcal{B}_i, \mu_i, T_i)$ (with $\mathcal{B}_i$ μ_i*-complete*[1]) are said to be isomorphic if there exist $X_i' \in \mathcal{B}_i$ for $i = 1, 2$ and a bijection $J : X_1' \to X_2'$, such that $\mu_i(X_i \setminus X_i') = 0$ and $T_i(X_i') \subset X_i'$ for $i = 1, 2$, J and J^{-1} are measurable, and $J_*(\mu_{1|X_1'}) = \mu_{2|X_2'}$ and $T_2 \circ J = J \circ T_{1|X_1'}$.

(Again see Sec. 5.2 in my notes to [40]; regarding (3) see also my notes regarding [40, Def. 1.3].)

Homogeneous dynamics: See the first two pages of [34] regarding the basic definitions; we will describe these in more detail in a later lecture. I mentioned the fact that the geodesic flow on the unit tangent bundle of any

[1] I think that if $(X_i, \mathcal{B}_i, \mu_i, T_i)$, $i = 1, 2$ are mpt's with $\mathcal{B}_i$ not necessarily μ_i-complete, then the most natural definition is to say that they are isomorphic *if* $(X_i, \mathcal{B}_i^{\mu_i}, \mu_i, T_i)$ *$(i = 1, 2)$ are isomorphic.*

hyperbolic surface (of finite area) can be obtained as a special case; for this cf. [34, Exc. 10–11 of Sec. 1.1] (I hope to say more about this later).

Also translation surfaces and $N_s(\mathbb{R}^d)$ (=the space of locally finite point sets in $\mathbb{R}^d$) we will define more carefully in later lectures.

The "Proposition 1" on p. 9 of my lecture is = Sarig's [40, Prop. 1.1 (p. 5)]. Also the (very brief) stuff which I mention on p. 9 about *mixing* is taken from Sarig's [40, Sec. 1.4].

2. Ergodic theorems: PET & consequences

Lecture #2: PET & consequences

Purpose: Formulate PET, understand what it says and some of its consequences, in particular equidistribution. Along the way: Discuss convergence in $P(X)$ (= weak conv = conv. in distr.)

From lecture #1:

I didn't finish the last page of that lecture. Good for renewing!

Prop 1: Let $(X, \mathcal{B}, \mu, T)$ be an mpt. Then

$[\mu \text{ ergodic}] \iff \ldots$

Def: Every $E \in \mathrm{Inv}(T)$ has $\mu(E)=0$ or $\mu(X \setminus E) = 0$

$E \in \mathcal{B}$ & $T^{-1}(E) = E$

...

Thm 2.2 in Sarig & p.34 remark 2

Thm 1 (PET, Birkhoff): Let $(X, \mathcal{B}, \mu, T)$ be a ppt and $f \in L^1 = L^1(X,\mu)$. Then

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

converges μ-a.e. and in L^1 to some $\bar{f} \in L^1$ which is T-invariant a.e. (i.e. $\bar{f} \circ T = \bar{f}$ μ-a.e.).

If μ is ergodic then $\bar{f} = \underbrace{\int_X f \, d\mu}_{\text{a constant!}}$ μ-a.e.

proof of the last statement: $\bar{f}$ T-inv a.e. and μ ergodic $\Rightarrow$ $\bar{f} = $ const. a.e.!

Which constant? Apply "$\int_X d\mu$" to

$$A_N^f \xrightarrow[N\to\infty]{L^1} \bar{f} \quad \Rightarrow \quad \int_X A_N^f \, d\mu \xrightarrow[N\to\infty]{} \text{const}$$

But $\int_X f\circ T^k \, d\mu = \int_X f \, d(T^k)_*(\mu) = \int_X f \, d\mu$ and thus

$\int_X A_N^f \, d\mu = \int_X f \, d\mu$, $\forall N$. $\therefore$ const $= \int_X f \, d\mu$! QED.

We'll prove the rest of Thm 1 in Lecture #3. Here we'll discuss & motivate Thm 1 = PET.

For $f = 1_E$ (some $E \in \mathcal{B}$), $A_N^f(x) = \dfrac{\#\{0 \le k < N : T^k(x) \in E\}}{N}$,

thus PET for μ ergodic $\Rightarrow$ "time spent in E is proportional to $\mu(E)$ as $N\to\infty$"

Cf. Boltzmann (pre 1900); the "Ergodic Hypothesis" in statistical physics; the time spent by a system in some region [of the phase space of microstates with some energy] is proportional to the volume of this region.

Notation, given $(X, \mathcal{B})$: a measurable space

$\underline{\underline{P(X)}} = \{\mu \ : \ \mu \text{ a prob. measure on } (X, \mathcal{B})\}$

Given $T : X \to X$ m'ble:

$\underline{\underline{P^T(X)}} = \{\mu \in P(X) \ : \ T_* \mu = \mu\}$

$\underline{\underline{\mathcal{E}^T(X)}} = \{\mu \in P^T(X) : \ \mu \text{ ergodic}\}$

Important & useful if these can be found explicitly! Ex: Ratner! Ex: Torus flow or torus map! See Probl. 12

Cor 1: (Given $T : X \to X$ m'ble.) If $\mu_1, \mu_2 \in \mathcal{E}^T(X)$ then either $\mu_1 = \mu_2$ or μ_1, μ_2 are (mutually) singular.

say def; see below...

proof: Assume $\mu_1 \neq \mu_2$; take $E \in \mathcal{B}$ with $\mu_1(E) \neq \mu_2(E)$.
Apply PET with $f = 1_E$

$\Rightarrow$ For μ_1-a.e. $x \in X$: $A_N^{1_E}(x) \xrightarrow{N \to \infty} \mu_1(E)$

For μ_2-a.e. $x \in X$: $A_N^{1_E}(x) \xrightarrow{N \to \infty} \mu_2(E)$.

Call the sets of these x: X_1 and X_2 resp.

Then $X_1 \cap X_2 = \emptyset$ (since $\mu_1(E) \neq \mu_2(E)$), $\mu_1(X_1) = 1$, $\mu_2(X_2) = 1$. (Also $X_1, X_2 \in \mathcal{B}$.) Done!

Next: Get $A_N^{1_E}(x) \to \mu(E)$ for "all" (reasonable) E, and fixed x? $\rightsquigarrow$ equidistribution (of the orbit of x)

important in dyn syst, number theory, probability theory, geometry, ...

DEF: Let X be a metric (or metrizable) space, $\mathcal{B}$ its Borel σ-algebra, and $\mu \in \mathcal{P}(X)$. A sequence $\{x_k\}_0^\infty \subset X$ is <u>equidistributed in X w.r.t. μ</u> if

$$\forall f \in C_b(X): \quad \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{\infty} f(x_k) = \int_X f \, d\mu$$

bounded cont. fcns on X

Special case: $X = \mathbb{T}^1$, $\mu = $ Leb; "equidistr. mod 1"

$$\Leftrightarrow \quad \forall E \in \mathcal{B}: \ \mu(\partial E) = 0 \Rightarrow \lim_{N\to\infty} \frac{\#\{0 \le k < N : x_k \in E\}}{N} = \mu(E)$$

$$\Leftrightarrow \quad \underbrace{\frac{1}{N} \sum_{k=0}^{N-1} \delta_{x_k}}_{\in \mathcal{P}(X)} \quad \text{converges weakly to } \mu \text{ as } N \to \infty.$$

These equivalences are part of the "Portmanteau Theorem" in probability theory, and def. of "weak convergence" $\leftrightarrow$ "convergence in distribution". I will explain more soon!

Ex: If $a_1, \dots, a_k \in \mathbb{R}$, not all in $\mathbb{Q}$, then $\{p(k) \bmod 1\}_{k=0}^\infty \subset \mathbb{T}^1$ is equidistr. w.r.t. Lebesgue (Weyl).

See Problem 15!

Ex: If $\underline{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$ with $1, a_1, a_2, \dots, a_n$ linearly independent over $\mathbb{Q}$, then for any $\underline{x} \in \mathbb{T}^n$, $\{\underline{x} + k\underline{a}\}_{k=0}^\infty$ equidistr. in $\mathbb{T}^n$ wrt Lebesgue.

orbit of $T: \underline{x} \to \underline{x} + \underline{a}$

Cor 2: If $(X, \mathcal{B}, \mu, T)$ is an <u>ergodic</u> ppt with X an <u>lcscH space</u> (and $\mathcal{B}$ its Borel σ-algebra)

ie locally compact, second countable, Hausdorff ($\Rightarrow$ Polish) cf. notes.

then $\exists\, X' \in \mathcal{B}$ s.t. $\mu(X')=1$ and $\forall x \in X'$: <u>$\{T^k(x)\}_{k=0}^{\infty}$ is equidistributed in X wrt μ.</u>

<u>proof outline:</u> Take $\{f_1, f_2, \dots\}$ dense in $C_c(X)$

The space of all $f \in C(X)$ with compact support.

Apply PET for each f_k

$\Rightarrow$ get set $X' \in \mathcal{B}$ such that

$$\forall x \in X': \quad \lim_{N\to\infty} A_N^{f_k}(x) = \int_X f_k \, d\mu, \quad (\forall k) \qquad \circledast$$

Approximation $\Longrightarrow$ $\circledast$ holds $\forall f \in C_c(X)$

$\Longrightarrow$ $\circledast$ holds for all $f \in C_b(X)$. $\square$

"μ is tight"

Weak convergence

DEF: The weakest topology for which $\mu \mapsto \mu(f)$ is continuous, $\forall f \in C_b(X)$.

DEF: For X a metric (or metrizable) space, we provide $P(X)$ with the weak topology; then $\mu_n \to \mu$ in $P(X)$ iff $\mu_n(f) \to \mu(f)$, $\forall f \in C_b(X)$.

μ_n converges weakly to μ

Standard notion in probability theory! Note from a functional analytic point of view, "weak-$*$ convergence" is a more appropriate name, at least for X LCH.

~~Note the above is not the definition of "weak topology"; see notes!~~

$\Leftrightarrow$ convergence in distribution

Basic facts

"Portmanteau Thm" $\mu_n \to \mu \iff \left[\mu_n(E) \to \mu(E) \text{ for every } E \in \mathcal{B} \text{ w/ } \mu(\partial E) = 0\right]$

$$\iff \left[\limsup_{n\to\infty} \mu_n(F) \le \mu(F) \quad \forall F \overset{\text{closed}}{\subset} X\right]$$

$$\iff \left[\liminf_{n\to\infty} \mu_n(U) \ge \mu(U) \quad \forall U \overset{\text{open}}{\subset} X\right]$$

X separable $\Rightarrow$ $P(X)$ metrizable, with metric = Prohorov distance:

$$d(\mu,\nu) = \inf\left\{\varepsilon > 0 : \mu(A) \le \nu(A^\varepsilon) + \varepsilon \ \&\ \nu(A) \le \mu(A^\varepsilon) + \varepsilon \quad \forall A \in \mathcal{B}\right\}$$

(the ε-neighborhood of A)

X Polish $\Rightarrow$ $P(X)$ Polish, X compact $\iff$ $P(X)$ compact

Clear since "$X \subset P(X)$" via $x \mapsto \delta_x$

For X lcscH

Aside: $M(X) := \left[\begin{array}{l}\text{The set of locally finite} \\ \text{Borel measures on } X\end{array}\right]$,

with the <u>vague</u> topology, i.e. $\mu_n \to \mu$ iff $\mu_n(f) \to \mu(f)$, $\forall f \in C_c(X)$.

We mentioned in lecture #1, for $X = \mathbb{R}^d$

$M(X)$ is a Polish space – see problem 14!

Note $\underline{P(X) \subset M(X)}$ and in fact the induced topology on $P(X)$ <u>is</u> the weak topology!

For X lcscH!! Cf Probl 6

But $\underline{P(X) \text{ is not closed in } M(X)}$, unless X compact!

Ex? Related to <u>tightness</u>...

Setting: X compact metric space, $T: X \to X$ continuous

"topological dynamics"

Then $T_*: P(X) \to P(X)$ is continuous.

Thm 2: Given $\mu \in P(X)$, set $\nu_N := \frac{1}{N}\sum_{k=0}^{N-1} T_*^k \mu$; then <u>any limit point of $\{\nu_N\}_1^\infty$ is T-invariant!</u>

"Proof": Use $T_*(\nu_N)(f) - \nu_N(f) = \frac{1}{N}\left(\sum_{k=1}^{N} \mu(f \circ T^k) - \sum_{k=0}^{N-1} \mu(f \circ T^k)\right)$

$$= \frac{1}{N}\mu(f \circ T^N - f) \to 0, \quad \forall f \in C(X) \ldots$$

Cor 1: $P^T(X) \neq \emptyset$ (Kryloff – Bogolioubo ff Thm)

Cor 2: <u>If</u> $P^T(X) = \{\mu\}$ {thus T <u>uniquely ergodic</u>!}

then <u>$\forall x \in X$: $\{T^k(x)\}_{k=0}^\infty$ is equidistributed in X wrt μ.</u>

{not just for <u>μ-a.e. x</u>!}

<u>proof</u>: Otherwise $\exists x \in X$ s.t. $\nu_N := \frac{1}{N}\sum_{k=0}^{N-1} \delta_{T^k(x)} \not\to \mu$,

so $\exists f \in C(X)$, $\varepsilon > 0$, $1 \le N_1 < N_2 < \ldots$ s.t.

$\circledast$ $|\nu_{N_j}(f) - \mu(f)| > \varepsilon$, $\forall j$. Take limit point $\nu \in P(X)$ of $\{\nu_{N_j}\}_{j=1}^\infty$. Thm 2 $\Rightarrow$ $\nu \in P^T(X)$

$\therefore \nu = \mu$, contradicting $\circledast$!

$\square$

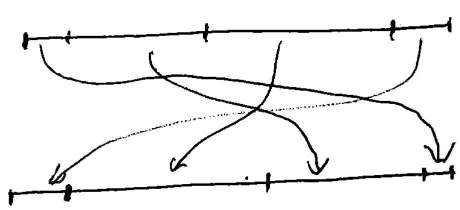
Note: The proof of Cor 2 is "easy" and does <u>not</u> use the PET!

Cor 3: If $P^T(X) = \{\mu\}$ and $\mathrm{supp}(\mu) = X$ then T is <u>minimal</u>, i.e. $\forall x \in X : \overline{\{T^k(x) : k \geq 0\}} = X$

$$\Leftrightarrow \nexists\ \emptyset \neq Y \underset{\neq}{\overset{\text{closed}}{\subseteq}} X : T(Y) \subset Y$$

A purely topological notion; "every orbit is dense" $\Leftrightarrow$ "(X, T) has no nontrivial closed subsystem"

<u>Q</u>: Does Cor 2 extend to an <u>IET</u>? (Then T is <u>not</u> continuous!) — See Probl. 10!



2.1. **Notes.** .

The first example on p. 4: This is a theorem by Weyl; cf. Problem 15 and (e.g.) [12, Thm. 1.4 (and Sec. 4.4)]. The second example on p. 4: See e.g. [12, Cor. 4.15]; note that this is also part of Problem 12.

In relation to Cor. 2 on p. 5, and for later use, we here discuss some classes of topological spaces: When formulating results for "a general topological space subject to some conditions", the classes of spaces which we will most often consider are (I think):
(1) Compact metrizable spaces.
(2) *lcscH* spaces, i.e. the topological spaces which are locally compact, second countable and Hausdorff.
(3) *Polish* spaces, i.e. the topological spaces which are separable and metrizable with a complete metric.

Let here us note that "(1) $\Rightarrow$ (2) $\Rightarrow$ (3)", i.e. any compact metric space is lcscH, and any lcscH space is Polish. [Details: The implication (1) $\Rightarrow$ (2) is quite basic; we need just point out that any compact metric space is totally bounded; hence separable; and for metric spaces separability and second countability are equivalent! The implication (2) $\Rightarrow$ (3) lies deeper; cf., e.g., [22, Thm. 5.3].]

Also in the proof of Cor. 2 (p. 5) we use the following fact: *If X is an lcscH space then $C_c(X)$ is separable.*[2] — This follows from the answer here (stackexchange), which applies since any lcscH space is easily seen to be σ-compact (viz., can be expressed as a countable union of compact sets). A key fact used there is that for any *compact* metric space K, $C(K)$ is separable; for this see e.g. [12, Lemma B.8].

Some more details for the end of the proof of Cor. 2 (p. 5): Here we are actually using two basic general facts about weak convergence: Let X be any metric space and let $\mu, \mu_1, \mu_2, \ldots \in P(X)$:
(1) If $S \subset C_b(X)$ and $\mu_n(f) \to \mu(f)$ for all $f \in S$, then $\mu_n(f) \to \mu(f)$ for all $f \in \overline{S}$ (the closure of S in $C_b(X)$).
(2) If X is lcscH, and $\mu_n(f) \to \mu(f)$ for all $f \in C_c(X)$, then $\mu_n \to \mu$ in $P(X)$ (weak convergence).

Proof of (1): Exercise! *Proof of (2):* This can be proved by using the fact that for any $\mu \in P(X)$ and any $\varepsilon > 0$ there is some compact set $K \subset X$ such that $\mu(K) > 1 - \varepsilon$; cf., e.g., [38, Thm. 2.18][3]; and also using the fact that for any compact subset $K \subset X$, there is some $h \in C_c(X)$ satisfying $0 \leq h \leq 1$ and $h_{|K} = 1$; cf., e.g., [38, Lemma 2.12]. Note that for any such

[2] $C_c(X)$ is a subspace of $C_b(X)$, the space of all bounded continuous functions on X. We always view $C_b(X)$ as a normed vector space with the norm $\|f\| := \sup_{x \in X} |f(x)|$ (i.e. the "supremum norm" or "L^∞ norm"). Of course this also makes $C_b(X)$ and its subspace $C_c(X)$ into metric spaces, with metric $d(f_1, f_2) = \|f_1 - f_2\|$.

[3] noticing that every open subset of an lcscH space is σ-compact.

h, and any $f \in C_b(X)$ the product hf is in $C_c(X)$! We leave it as Problem 6 to carry out the details of the argument.

[Application of these facts in the proof of Cor. 2: Fix $x \in X'$ and set $\mu_N := N^{-1} \sum_{k=0}^{N-1} \delta_{T^k(x)}$. Then $A_N^f(x) = \mu_N(f)$ for any $f \in C_b(X)$. Thus we know $\mu_N(f_k) \to \mu(f_k)$ for each $k \in \mathbb{Z}^+$; hence by (1) above we have $\mu_N(f) \to \mu(f)$ for all $f \in C_c(X)$, and in particular for all $f \in C_c(X)$. By (2) this implies $\mu_N \to \mu$ in $P(X)$, i.e. $\{T^k(x)\}_{k=0}^{\infty}$ is equidistributed in X wrt μ.]

Regarding the definition of the weak topology on $P(X)$ (for X a metric space), see, e.g., Billingsley, [5, Ch. 1] or Kallenberg, [20, Ch. 4]. In particular, for the "Portmanteau Theorem", see [5, Thm. 2.1], and for the facts I mentioned about the Prohorov distance, see [5, pp. 72–73]. Note that our space "$P(X)$" is called "$\mathcal{M}(X)$" in [12, Ch. 4], and "$\mathcal{M}$" in [34, p. 114]; however in this course I prefer to let $M(X)$ (for X a lcscH space) denote the set of all locally finite Borel measures on X; cf. p. 7 of the lecture. For basics about $M(X)$ we refer to Problem 14 and Kallenberg, [20, Thm. A2.3(i), (ii)].

Everything in our brief discussion on pp. 8–9 about the setting with X a compact metric space and $T : X \to X$ continuous can be found in [12, Ch. 4]. Indeed, our Thm. 2 is a special case of [12, Thm. 4.1]; our Cor. 1 is [12, Cor. 4.2]; our Cor. 2 is a special case of [12, Thm. 4.10].

3. Ergodic theorems: MET & PET – proofs

Lecture # 3 : MET and PET

Thm 1 (MET, von Neumann): Let $(X, \mathcal{B}, \mu, T)$ be a ppt. If $f \in L^2$ then $\boxed{A_N^f \xrightarrow[N\to\infty]{L^2} \bar{f}}$ ⊛

where $\bar{f} \in L^2$ is T-invariant. as L^2-element, i.e. a.e.!

If μ is ergodic then $\bar{f} = \int_X f \, d\mu$. constant!

Recall $A_N^f := \frac{1}{N} \sum_{k=0}^{N-1} f \circ T^k$

Note: $f \mapsto f \circ T$ is a unitary operator $L^2 \circlearrowleft$ (viz., $\langle f \circ T, g \circ T \rangle = \langle f, g \rangle$, $\forall f, g \in L^2$).

It is called "the Koopman operator."

Cf. Sarig's Ch. 3.

proof: Let

$$C = \{ g - g \circ T : g \in L^2 \}.$$

Space of coboundaries

For $f \in C$, ⊛ is easy, with $\bar{f} = 0$.

[Indeed, if $f = g - g \circ T$ then

$$A_N^f = N^{-1} \sum_{k=0}^{N-1} f \circ T^k = N^{-1}(g - g \circ T^N)$$

and $\| N^{-1}(g - g \circ T^N) \| \leq \frac{2}{N} \|g\| \to 0$ as $N \to \infty$]

Easy approximation $\Rightarrow$ ⊛ holds $\forall f \in \bar{C}$, with $\bar{f} = 0$.

Use $\|A_N^f\| \leq \|f\|$, $\forall f \in L^2$

However, $\boxed{L^2 = \overline{C} \oplus I}$

orthogonal direct sum

with $\underline{\underline{I = \{f \in L^2 : f \circ T = f\}}}$ in L^2, i.e. a.e.

[proof: $\underline{C \perp I}$ (for if $f \in I$, $g \in L^2$ then $\langle g - g\circ T, f\rangle = \langle g, f\rangle - \langle g\circ T, f\circ T\rangle = 0$).

$\therefore$ $\underline{\overline{C} \perp I}$, and remains to prove $\overline{C}^\perp = \underline{C^\perp \subset I}$.

Take $f \in C^\perp$. Then $\|f - f\circ T\|^2 = \langle f - f\circ T, f - f\circ T\rangle$

$= 2\|f\|^2 - 2\langle f, f\circ T\rangle = 2\langle f, f - f\circ T\rangle = 0$.

$f \perp C$!

$\therefore$ $f \in I$; done!]

Also, for $f \in I$, $\underline{\circledast \text{ obvious}}$, with $\overline{f} = f$!

$\underline{\text{Linearity}} \Rightarrow \circledast$ holds $\forall f \in L^2$, with

$\underline{\overline{f} = [\text{orthogonal projection of } f \text{ on } I]}$.

Last part of Thm 1: Proof just as for last part of PET (cf. lecture #2), i.e. note $\int \overline{f}\, d\mu = \lim \int A_N^f\, d\mu = \int f\, d\mu$, etc.

□

<u>Special case of Thm 1:</u> $\forall A, B \in \mathcal{B}$:

$$\frac{1}{N}\sum_{k=0}^{N-1} \underbrace{\mu(A \cap T^{-k}B)}_{=\langle 1_A, 1_B \circ T^k \rangle} = \langle 1_A, A_N^{1_B} \rangle \xrightarrow[N\to\infty]{} \underbrace{\langle 1_A, \overline{1_B} \rangle}_{\circledast}.$$

If μ ergodic then $\overline{1_B} = \mu(B)$ a.e., so that

$\circledast = \mu(A)\mu(B)$!

Thus μ ergodic $\Rightarrow$ "mixing on average" !
(Cf <u>mixing</u>, and <u>weak mixing</u>; stronger concepts!)

Next, recall Thm 1 from Lecture #2 = PET:

<u>Thm (PET)</u>: Let $(X, \mathcal{B}, \mu, T)$ be a ppt and $f \in L^1$. Then A_N^f converges μ-a.e. and in L^1 to some $\overline{f} \in L^1$ which is T-invariant a.e.
((If μ is ergodic then $\overline{f} = \int_X f\,d\mu$ a.e.))

Remark: $L^2 \subset L^1$, and for <u>$f \in L^2$</u> the L^1-conv in PET follows from ~~the~~ MET!
~~This can in fact be used to build a proof of PET.~~

In general, the L^1-conv is an "easy" consequence of the a.e.-conv! See Sarig's p. 34, Remark 2.

proof outline, assuming $f \in L^\infty$:

Then wlog assume $0 \leqslant f \leqslant 1$.

Set $\overline{A}(x) = \limsup_{n\to\infty} A_n(x) \in [0,1]$

$\underline{A}(x) = \liminf_{n\to\infty} A_n(x) \in [0,1]$

exist for all $x \in X$!

Easy: $0 \leqslant \underline{A} \leqslant \overline{A} \leqslant 1$, and $\overline{A}, \underline{A}$ are T-invariant.

Claim: $\boxed{\int_X \overline{A}\, d\mu \leqslant \int_X f\, d\mu}$ ⊛

[This implies PET, since $f \leftrightarrow 1-f \rightsquigarrow \int_X \underline{A}\, d\mu \geq \int f\, d\mu$,

thus $\int_X (\overline{A} - \underline{A})\, d\mu = 0$, i.e. $\overline{A} = \underline{A}$ a.e.,

i.e. $\lim_{n\to\infty} A_n(x)$ exists a.e.!

L^1-conv. then clear by Leb. bdd. conv.]

just "first idea"

proof of ⊛:

Write $S_N^f(x) := \sum_{k=0}^{N-1} f(T^k(x))$.

Fix $\varepsilon > 0$, take N large.

$$\int_X f\, d\mu = \frac{1}{N} \int_X S_N^f(x)\, d\mu(x)$$

since μ T-inv.

Given x, take smallest N_1 with

$$S^f_{N_1}(x) > N_1\left(\bar{A}(x)-\varepsilon\right),$$

next take smallest N_2 with

$$S^f_{N_2}\left(T^{N_1}(x)\right) > N_2\left(\bar{A}\left(T^{N_1}(x)\right)-\varepsilon\right) = N_2\left(\bar{A}(x)-\varepsilon\right),$$

next take smallest N_3 with

$$S^f_{N_3}\left(T^{N_1+N_2}(x)\right) > N_3\left(\bar{A}(x)-\varepsilon\right)$$

etc.

Get $$S^f_N(x) > \underbrace{(N_1+N_2+\dots+N_r)}_{<N \text{, but hopefully } \approx N\,!}\left(\bar{A}(x)-\varepsilon\right)$$

$$\Rightarrow \frac{1}{N}S^f_N(x) > \underbrace{\frac{N_1+\dots+N_r}{N}}_{<1 \text{ but hopefully } >0{,}999\,!}\left(\bar{A}(x)-\varepsilon\right)$$

$$\overset{?}{\Longrightarrow} \int_X f\,d\mu = \frac{1}{N}\int_X S^f_N\,d\mu > 0{,}999\int_X\left(\bar{A}(x)-\varepsilon\right)d\mu(x)$$

Technical, nice!!

□

Next: Understanding the limit function $\bar{f}$ in MET and PET! (Same if $f\in L^2$!)
Answer: $\bar{f}$ is the conditional expectation of f wrt the T-invariant sub-σ-algebra!

Thus we now turn to conditioning!

Conditioning

Let $(X, \mathcal{B}, \mu)$ - a probability space.

$\mathcal{F} \subset \mathcal{B}$ - a sub-σ-algebra.

$f \in L^1 = L^1(X, \mathcal{B}, \mu)$.

The conditional expectation of f given $\mathcal{F}$

sloppy: "μ"

<u>DEF</u>: $\mathbb{E}(f \mid \mathcal{F})$ is the unique element in $L^1(X, \mathcal{F}, \mu_{|\mathcal{F}})$ satisfying $\forall A \in \mathcal{F}$: $\underline{\int_A \mathbb{E}(f \mid \mathcal{F})\, d\mu = \int_A f\, d\mu}$.

Recall: $\mathbb{E}(f \mid \mathcal{F}) \in L^1(X, \mathcal{F}, \mu_{\mathcal{F}})$ means $\mathbb{E}(f \mid \mathcal{F})$ is equiv. class of <u>$\mathcal{F}$-m'ble</u> functions! (Hence Sarig's "I" in Def 2.1 is not needed.)

<u>proof of $\exists!$</u>

can avoid; write $f = f_+ - f_-$ etc

Consider the signed measure $\nu_f = f \cdot \mu_{|\mathcal{F}}$ on $(X, \mathcal{F})$. (<u>Def</u>: $\nu_f(A) = \int_A f\, d\mu$, $\forall A \in \mathcal{F}$)

Note $\nu_f \ll \mu_{|\mathcal{F}}$. Hence by <u>Radon-Nikodym</u>,

$\exists!\ g \in L^1(X, \mathcal{F}, \mu)$ s.t. $\nu_f(A) = \int_A g\, d\mu$, $\forall A \in \mathcal{F}$

a <u>density</u> of ν_f w.r.t. $\mu_{|\mathcal{F}}$

Take $\mathbb{E}(f \mid \mathcal{F}) := g$.

<u>Properties</u>: $f \mapsto \mathbb{E}(f \mid \mathcal{F})$ linear, with

$$\|\mathbb{E}(f \mid \mathcal{F})\|_{L^1} \leq \|f\|_{L^1}$$

$\{L^1(X,\mathcal{F},\mu) \subset L^1(X,\mathcal{B},\mu)\}$

$$\forall \varphi \in L^\infty(X,\mathcal{F}): \quad \mathbb{E}(\varphi f \mid \mathcal{F}) = \varphi \cdot \mathbb{E}(f \mid \mathcal{F})$$

special cases: $\mathbb{E}(\varphi \mid \mathcal{F}) = \varphi$ and $\mathbb{E}(c \mid \mathcal{F}) = c$

{a <u>constant</u>}

$$f_1 \leq f_2 \implies \mathbb{E}(f_1 \mid \mathcal{F}) \leq \mathbb{E}(f_2 \mid \mathcal{F}) \quad (\mu\text{-a.e.})$$

<u>If $\mathcal{F}_2 \subset \mathcal{F}_1$</u>: $\mathbb{E}\big(\mathbb{E}(f \mid \mathcal{F}_1) \mid \mathcal{F}_2\big) = \mathbb{E}(f \mid \mathcal{F}_2)$.

<u>Also DEF</u>: For $A \in \mathcal{B}$: $\underline{\mu(A \mid \mathcal{F})} := \mathbb{E}(1_A \mid \mathcal{F})$

{"<u>$(\mu$-$)$probability of A given $\mathcal{F}$</u>"
More usual to see: "Prob$(A \mid \mathcal{F})$"...}

<u>Discussion & motivation</u> {Consider the following special case!}

Assume $X = \bigsqcup_{j=1}^{\infty} B_j$ with $B_j \in \mathcal{B}$

{explain! <u>partition</u>}

Let $\underline{\mathcal{F} = \sigma(\{B_1, B_2, \ldots\})}$ ← {This is just the set of <u>all</u> <u>unions</u> of B_j's!}

Then for j with $\mu(B_j) > 0$: $\underline{\mu(A \mid \mathcal{F})(x) = \dfrac{\mu(A \cap B_j)}{\mu(B_j)}}$, $\forall x \in B_j$

{See probl 16!}

Hence we recovered the "undergrad def" (Bayes) of conditional probability. But if one wishes to condition on an event of probability 0 ($\mu(B_j)=0$) then this doesn't work. The trick is to instead view conditional probability (or expectation) as a <u>function</u> on X; this leads to a natural answer <u>a.e.</u>

For more discussion, cf, e.g. Billingsley!

<u>One more example</u>: Let μ be an abs. cont. prob. measure on $(\mathbb{R}^2, \mathcal{B}_2)$. $\mathcal{B}_n :=$ Borel σ-algebra of $\mathbb{R}^n$

Thus $\mu = \delta \cdot m$ ($m =$ Lebesgue on $\mathbb{R}^2$) for some $\delta \in L^1(\mathbb{R}^2, m)$, $\delta \geq 0$.

Let $\mathcal{F} = \{B \times \mathbb{R} : B \in \mathcal{B}_1\}$ $\Leftarrow$ The "experiment $\mathcal{F}$" $\Leftrightarrow$ "determine the x_1-coordinate"!

Then for any $A \in \mathcal{B}_1$:

$$\mu(\mathbb{R} \times A \mid \mathcal{F})(x,y) = \frac{\int_A \delta(x,t)\,dt}{\int_{\mathbb{R}} \delta(x,t)\,dt} \quad \text{for } \mu\text{-a.e. } (x,y).$$

indep of y!

This is the classical formula for the <u>conditional probability of $y \in A$ "given x"</u>!

See problem 17

Finally we now have:

Given ppt $(X, \mathcal{B}, \mu, T)$ and $f \in L^1$

Thm 2: In PET, $\bar{f} = \mathbb{E}(f \mid Inv(T))$ μ-a.e.

in fact both sides are def μ-a.e.

Recall here $Inv(T) = \{E \in \mathcal{B} : E = T^{-1}(E)\}$

Note $h : X \to \mathbb{R}$ is $Inv(T)$-m'ble $\Leftrightarrow$ h is T-inv.

proof of Thm 2: Recall the precise def:

$$\bar{f}(x) = \begin{cases} \lim_{N\to\infty} A_N^f(x) & \text{if exists} \\ 0 & \text{otherwise} \end{cases}$$

$\bar{f}$ is T-inv, i.e. $Inv(T)$-m'ble.

Hence it only remains to check that

$$\int_B \bar{f}\, d\mu = \int_B f\, d\mu, \quad \forall B \in Inv(T).$$

We noted this in Lecture #2 for B=X. The proof for general $B \in Inv(T)$ is "the same"!

But $\int_B \bar{f}\, d\mu = \int_B \left(\lim_{N\to\infty} A_N^f\right) d\mu \underset{L^1\text{-conv!}}{=} \lim \int_B A_N^f\, d\mu \underset{\mu \text{ and } B \text{ are } T\text{-inv!}}{=} \lim \int_B f\, d\mu$

$= \int_B f\, d\mu.$ Done! $\square$

3.1. **Notes.** .

This lecture corresponds to Sarig, [40, Sec. 2.1–2.3.1]. See also my notes to [40]; in particular (in the "details" section) I elaborate on several of the details in Sarig's proofs of the MET and the PET (for $f \in L^\infty$).

Regarding the remaining step of the proof of the PET, i.e. treating $f \in L^1$ and not only $f \in L^\infty$, see Sarig's [40, Sec. 2.4] where a more general result is proved. Alternatively, this is obtained as a special case of the Subadditive Ergodic Theorem, [40, Thm. 2.7], as we will discuss in Lecture # 5. (The proof of [40, Thm. 2.7] uses the PET, but *only for functions in* L^∞*.*)

On p. 8 in the lecture I refer to "Billingsley" for a more thorough discussion about conditioning; the precise reference which I have in mind is Billingsley, [4, Sec. 33–34].

4. Conditional probabilities; ergodic decomposition

Lecture #4: Conditional probabilities; ergodic decomposition

Def: For $(X, \mathcal{B}, \mu)$ a measure space, we set

$$\underline{\mathcal{L}^1_\mu} = \mathcal{L}^1(X, \mathcal{B}, \mu) := \left\{ f : X \to \mathbb{R} \text{ m'ble} : \int_X |f| \, d\mu < \infty \right\}.$$

Thus $\underline{L^1_\mu} = \underline{L^1(X, \mathcal{B}, \mu)} =$ " $\mathcal{L}^1(X, \mathcal{B}, \mu)$ mod μ-a.e. $=$ "

Thm 1: Let $(X, \mathcal{B})$ be a standard Borel space and $\mu \in P(X)$. Let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{B}$.

Then $\exists \, \underline{\{\mu_x\}_{x \in X}}$ (or $\{\mu_x^{\mathcal{F}}\}$) $\subset P(X)$ s.t. for every $f \in \mathcal{L}^1(X, \mathcal{B}, \mu)$, $\underline{x \mapsto \int_X f \, d\mu_x \text{ is a version of } \mathbb{E}(f|\mathcal{F})}.$

viz., $x \mapsto \int_X f \, d\mu_x$ is in $\mathcal{L}^1(X, \mathcal{F}, \mu)$ and
$\forall A \in \mathcal{F}: \int_A f \, d\mu = \int_A \left(\int_X f \, d\mu_x \right) d\mu(x).$

In particular $\boxed{\text{"} \mu = \int_X \mu_x \, d\mu(x) \text{"}}$ i.e. $\int_X f \, d\mu = \int_X \left(\int_X f \, d\mu_x \right) d\mu(x)$,
$\forall f \in \mathcal{L}^1(X, \mathcal{B}, \mu)$

These μ_x are uniquely determined μ-a.e., i.e. if $\{\mu'_x\}_{x \in X}$ is also "ok" then $\mu'_x = \mu_x$ for μ-a.e. x.

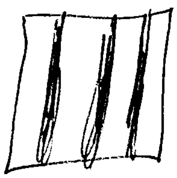
Example

$X = [0,1]^2$

$\mathcal{B}$ = Borel

μ = Lebesgue

$\mathcal{F} = \{B \times [0,1] \;:\; B \text{ Borel} \subset [0,1]\}$



Then <u>can</u> take:

Depends only on x

$\mu_{(x,y)}$ = Leb. on $\{x\} \times [0,1]$

i.e. $\mu_{(x,y)}(E) = m(\{t \;:\; (x,t) \in E\}), \quad \forall E \in \mathcal{B}.$

1-dim Leb. on $[0,1]$

<u>NOT</u> useful/reasonable to extend to $\mathcal{B}^{\mu}$!

Discussion: Recall, for $E \in \mathcal{B}$:

$$\mu(E \mid \mathcal{F}) := \underbrace{\mathbb{E}(1_E \mid \mathcal{F})}_{[0,1]\text{-valued!}} \in L^1(X, \mathcal{F}, \mu)$$

For $E_1, E_2, \ldots \in \mathcal{B}$ pairwise disjoint:

$$\circledast \quad \mu\Big(\bigcup_{j=1}^{\infty} E_j \mid \mathcal{F}\Big) = \sum_{j=1}^{\infty} \mu(E_j \mid \mathcal{F}) \quad \text{in } L^1(X, \mathcal{F}, \mu)$$

{proof: t.h. "ok" by mon. conv. thm.}

Also $\mu(\emptyset \mid \mathcal{F}) = 0$ a.e., $\mu(X \mid \mathcal{F}) = 1$ a.e.

$\therefore$ $\mu(\cdot \mid \mathcal{F})$ is an "$L^1(X, \mathcal{F}, \mu)$-valued measure" on $(X, \mathcal{B})$!

But the problem is that $\circledast$ holds only for μ-a.a. x, and the set of exceptions may depend on E. We need to select one version of $\mu(E \mid \mathcal{F})$, for each $E \subset \mathcal{B}$, so that for a.e. x, $\circledast$ holds for all $E \in \mathcal{B}$!
Natural: Start from some nice countable family of E's.
Technically nicer: Work with functions...

proof sketch: (following Sarig)

May assume X is a compact metric space.
Then $C(X)$ is separable; take dense subset $\{f_0, f_1, f_2, \ldots\}$ with $f_0 \equiv 1$. Let $\mathcal{A}_{\mathbb{Q}} = [\mathbb{Q}\text{-algebra generated by } \{f_j\}]$.
Then $\mathcal{A}_{\mathbb{Q}} \subset C(X)$, $\mathcal{A}_{\mathbb{Q}}$ is countable and dense in $C(X)$.
For each $f \in \mathcal{A}_{\mathbb{Q}}$, fix a version $\overline{\mathbb{E}}(f \mid \mathcal{F})$ of $\mathbb{E}(f \mid \mathcal{F})$.

Let $\underline{\underline{X_0}}$ be the set of all $x \in X$ s.t.

0. $\underline{\overline{\mathbb{E}}(1 \mid \mathcal{F})(x) = 1}$.
1. $\forall\, \alpha, \beta \in \mathbb{Q},\ g_1, g_2 \in \mathcal{A}_{\mathbb{Q}}$: $\underline{\overline{\mathbb{E}}(\alpha g_1 + \beta g_2 \mid \mathcal{F})(x) = \alpha \overline{\mathbb{E}}(g_1 \mid \mathcal{F})(x) + \beta \overline{\mathbb{E}}(g_2 \mid \mathcal{F})(x)}$
2. $\forall g \in \mathcal{A}_{\mathbb{Q}}$: $\underline{\min g \leq \overline{\mathbb{E}}(g \mid \mathcal{F})(x) \leq \max g}$.

Then $X_0 \in \mathcal{F}$ and $\mu(X_0) = 1$.

For each $x \in X_0$, the map $g \longmapsto \overline{\mathbb{E}}(g \mid \mathcal{F})(x)$ is a $\mathbb{Q}$-linear functional on $\mathcal{A}_{\mathbb{Q}}$ of norm ≤ 1.

Now $\underline{\exists!\ \varphi_x \in C(X)^*;\ \|\varphi_x\| \leq 1}$ s.t.
$\underline{\varphi_x(g) = \overline{\mathbb{E}}(g \mid \mathcal{F})(x),\ \forall g \in \mathcal{A}_{\mathbb{Q}}}$.

Then $\forall g \in C(X)$: $\underline{g \geq 0 \Rightarrow \varphi_x(g) \geq 0}$.

{Riesz repr. Thm}

Hence: $\exists!\ \underline{\mu_x \in M(X)}$ s.t. $\underline{\varphi_x(g) = \mu_x(g) = \int_X g \, d\mu_x}$, $\forall g \in C(X)$

In fact $\underline{\underline{\mu_x \in P(X)}}$, since $\mu_x(X) = \int_X d\mu_x = \overline{\mathbb{E}}(1 \mid \mathcal{F})(x) = 1$.

For $x \in X \setminus X_0$, set $\mu_x := \mu$, say!

Step 1: $\forall f \in C(X)$: $x \mapsto \int_X f \, d\mu_x$ is a version of $\mathbb{E}(f|\mathcal{F})$.

(Viz., $x \to \int_X f \, d\mu_x$ is $\mathcal{F}$-m'ble, and

$\forall A \in \mathcal{F}$: $\int_A \left(\int_X f \, d\mu_x \right) d\mu(x) = \int_A f \, d\mu$.)

proof: True by constr. for $f \in \mathcal{A}_{\mathbb{Q}}$; next use the fact that $\mathcal{A}_{\mathbb{Q}}$ is dense in $C(X)$.

Step 2: $\forall E \in \mathcal{B}$: $x \mapsto \int_X 1_E \, d\mu_x$ is a version of $\mathbb{E}(1_E|\mathcal{F})$.

(Viz. $x \mapsto \int_X 1_E \, d\mu_x$ is $\mathcal{F}$-m'ble and
$\forall A \in \mathcal{F}$: $\mu(A \cap E) = \int_A \mu_x(E) \, d\mu$.) $(*)$

proof: Let $\mathcal{M} = \{ E \in \mathcal{B} : E \text{ satisfies } (*) \}$.

"Note" $\mathcal{M} \supset \mathcal{A} := \{ U \subset X : 1_U \text{ is a pointwise limit of some (bounded) sequence in } C(X) \}$

$\mathcal{A}$ is an algebra, $\mathcal{M}$ is a "monotone class";
so Monotone Class Theorem $\Rightarrow \mathcal{M} \supset \sigma(\mathcal{A}) \supset \mathcal{B}$!

Step 3: $\forall f \in L^1(X, \mathcal{B}, \mu)$: $x \mapsto \int_X f \, d\mu_x$ is a version of $\mathbb{E}(f \mid \mathcal{F})$.

proof: Split $f = f^+ + f^-$ $\Rightarrow$ may assume $f \geq 0$

{Careful: For which x is $\int_X f \, d\mu_x$ undefined?}

Write f as pointwise limit of simple functions $0 \leq f_1 \leq f_2 \leq \cdots$ The claim holds for each f_n, by Step 2 (+ linearity). Now use Monotone Convergence Theorem, etc $\rightsquigarrow$ Done!

Step 4: Uniqueness.

proof: Let $\mathcal{U}$ be a countable base for the topology of X. We may assume $X \in \mathcal{U}$ and that $\mathcal{U}$ is closed under finite intersections.

$$\forall U \in \mathcal{U}: \qquad \underbrace{x \to \int_X 1_U \, d\mu_x}_{\mu_x(U)} \text{ and } x \mapsto \int_X 1_U \, d\mu'_x$$

are versions of $\mathbb{E}(1_U \mid \mathcal{F})$; hence equal μ-a.e.

$\therefore \exists X_0 \in \mathcal{F}$ s.t. $\mu(X_0) = 1$ and $\forall x \in X_0$:

$$\underbrace{[\forall U \in \mathcal{U}: \mu_x(U) = \mu'_x(U)]}.$$

$\Rightarrow \mu_x = \mu'_x$, by Monotone Class argument!

Addendum to Thm 1

Assume $\mathcal{F}$ is countably generated, i.e.

$\mathcal{F} = \sigma(\underbrace{\{A_1, A_2, \dots\}}_{\text{a countable set.}})$ for some $A_1, A_2, \dots \in \mathcal{F}$.

Then for any $x \in X$, set

$$\underbrace{[x]_{\mathcal{F}}}_{\text{the atom of } x} := \underbrace{\bigcap_{\substack{A \in \mathcal{F} \\ (x \in A)}} A}_{\text{is in } \mathcal{F}\text{? Yes!}} = \bigcap_{j=1}^{\infty} \begin{Bmatrix} A_j & \text{if } x \in A_j \\ X \setminus A_j & \text{if } x \notin A_j \end{Bmatrix}$$

Note: The atoms partition X. Also, if $E \in \mathcal{F}$ and $x \in E$, then $[x]_{\mathcal{F}} \subset E$!

Thm 2: In the above situation, $\exists X_0 \in \mathcal{F}$ s.t. $\mu(X_0) = 1$ and

(1) $\mu_x([x]_{\mathcal{F}}) = 1, \quad \forall x \in X_0$

(2) $\forall x_1, x_2 \underbrace{\in X}_{\text{All!}}: \ [x_1]_{\mathcal{F}} = [x_2]_{\mathcal{F}} \Rightarrow \mu_{x_1} = \mu_{x_2}$.

proof: Set $X_0 = \{x \in X : \mu_x(A_j) = 1_{A_j}(x), \ \forall j \geq 1\}$.

Then $X_0 \in \mathcal{F}$ and $\mu(X_0) = 1$ since for each j we have $\mu_x(A_j) = \mathbb{E}(1_{A_j} | \mathcal{F})(x) = 1_{A_j}(x)$ for μ-a.e. x.

Take $x \in X_0$. Then $\forall j \geq 1$:

$$x \in A_j \Rightarrow \mu_x(A_j) = 1$$

$$x \notin A_j \Rightarrow \mu_x(X \setminus A_j) = 1 - \mu_x(A_j) = 1 - 0 = 1$$

Hence $\mu_x([x]_{\mathcal{F}}) = \mu_x\left(\bigcap_{j=1}^{\infty} \begin{cases} A_j & \text{if } x \in A_j \\ X \setminus A_j & \text{if } x \notin A_j \end{cases}\right) = \underline{\underline{1}}$.

Next take any $x_1, x_2 \in X$ with $[x_1]_{\mathcal{F}} = [x_2]_{\mathcal{F}}$.

For each $f \in C(X)$, $x \mapsto \int_X f \, d\mu_x$ is

X compact!

$\mathcal{F}$-m'ble; hence the set

$$A := \left\{ x \in X : \int_X f \, d\mu_x = \int_X f \, d\mu_{x_1} \right\}$$

is in $\mathcal{F}$. Also $x_1 \in A$; hence $[x_1]_{\mathcal{F}} \subset A$,

and so $x_2 \in A$, i.e. $\underline{\int_X f \, d\mu_{x_2} = \int_X f \, d\mu_{x_1}}$.

□

Thm 3 (ergodic decomposition): Let $(X, \mathcal{B}, \mu, T)$ be a ppt where $(X, \mathcal{B})$ is a standard Borel space, and let $\{\mu_x\}_{x \in X}$ be the conditional probabilities w.r.t. $\mathcal{F} := \mathrm{Inv}(T)$. Then for μ-a.e. $x \in X$, μ_x is T-invariant and ergodic.

Recall also: "$\mu = \int_X \mu_x \, d\mu(x)$"!

proof: For every $f \in L^1$, set

$$X_f := \{x \in X : \mu_x(f) = \lim_{n\to\infty} A_n^f(x)\}.$$

Then $X_f \in \mathrm{Inv}(T) = \mathcal{F}$ and $\mu(X_f) = 1$.

[proof: $x \mapsto \mu_x(f)$ is a version of $\mathbb{E}(f \mid \mathrm{Inv}(T))$, hence $\mathrm{Inv}(T)$-m'ble and (by PET) for μ-a.e. x:
$\mu_x(f) = \mathbb{E}(f \mid \mathrm{Inv}(T))(x) = \lim_{n\to\infty} A_n^f(x)$.]

Fix $\{f_1, f_2, \dots\}$ dense in $C(X)$.

Set $$X' := \bigcap_{j=1}^{\infty} \left(X_{f_j} \cap X_{f_j \circ T}\right).$$

Now for all $j \geq 1$, $x \in X'$: $T_*(\mu_x)(f_j) = \int_X f_j \circ T \, d\mu_x =$

$$= \lim_{n\to\infty} A_n^{f_j \circ T}(x) = \lim_{n\to\infty} A_n^{f_j}(T(x)) = \lim_{n\to\infty} A_n^{f_j}(x) = \mu_x(f_j)$$

Since $\{f_1, f_2, \dots\}$ is dense in $C(X)$, it follows that (for $x \in X'$): $\underline{T_*(\mu_x)(f) = \mu_x(f), \quad \forall f \in C(X)}$.

Hence $T_*(\mu_x) = \mu_x$, i.e. $\underline{\mu_x \text{ is } T\text{-invariant}, \forall x \in X'}$.

$\underline{\mu_x \text{ ergodic ?}}$

This is possible by basic fact about standard Borel spaces.

Take $\underline{\text{countably generated}}$ sub-σ-algebra

$\underline{\mathcal{E} \subset \mathcal{F}} = \mathrm{Inv}(T)$ s.t. $\underline{\mathcal{E} \underset{\mu}{=} \mathcal{F}}$ $\left(\text{viz. } \forall B \in \mathcal{F} : \exists B' \in \mathcal{E} \text{ s.t. } \mu(B \Delta B') = 0\right)$.

Now we can take $N \in \mathcal{B}$ with $\mu(N) = 0$ s.t

$$\forall x \in X \setminus N : \begin{cases} \mu_x^{\mathcal{E}} = \mu_x^{\mathcal{F}} & \leftarrow \text{easy from } \mathcal{E} \underset{\mu}{=} \mathcal{F} \\ \mu_x^{\mathcal{E}} \text{ is } T\text{-invariant} & \leftarrow \text{cf. part 1 of this Thm 3} \\ \mu_x^{\mathcal{E}}(f_j) = \lim_{n\to\infty} A_n^{f_j}(x), \ \forall j \geq 1 & \leftarrow \\ \mu_x^{\mathcal{E}}([x]_{\mathcal{E}}) = 1 & \leftarrow \text{ok by Thm 2} \end{cases}$$

Note $0 = \mu(N) = \int_X \mu_x(N)\, d\mu(x) \Rightarrow \mu_x(N) = 0$, μ-a.e. x.

Set $\underline{N_1 = N \cup \{x \in X : \mu_x(N) > 0\}}$. Then $\underline{\mu(N_1) = 0}$.

fixed through the rest of the proof!

Take any $\underline{\underline{x \in X \setminus N_1}}$, then any $\underline{y \in [x]_{\mathcal{E}} \setminus N}$.

Then $\forall j$: $\underline{\underline{\lim_{n\to\infty} A_n^{f_j}(y)}} \underset{\substack{\uparrow \\ y \notin N}}{=} \mu_y^{\mathcal{E}}(f_j) \underset{\substack{\uparrow \\ \mu_y^{\mathcal{E}} = \mu_x^{\mathcal{E}} \text{ by Thm. 2}}}{=} \underline{\underline{\mu_x^{\mathcal{E}}(f_j)}}$.

Also $x \notin N_1 \Rightarrow \mu_x^\varepsilon(N) = 0$ and $\underline{\mu_x^\varepsilon([x]_\varepsilon \setminus N) = 1}$.

$\therefore$ $\underline{\left[\lim_{n\to\infty} A_n^{f_j}(y) = \mu_x^\varepsilon(f_j),\ \forall j \geq 1\right]}$ for $\underline{\underline{\mu_x^\varepsilon\text{-a.e. } y \in X}}$.

Hence by Lebesgue bounded convergence Theorem:

$$\lim_{n\to\infty} \Big\| \underset{\uparrow}{A_n^{f_j}} - \underbrace{\mu_x^\varepsilon(f_j)} \Big\|_{L^2_{\mu_x^\varepsilon}} = 0, \quad \forall j \geq 1.$$

{function on X} (constant)

Now $\{f_1, f_2, \ldots\}$ is dense in $C(X)$, which is dense in $L^2_{\mu_x^\varepsilon}$.

$\therefore$ $\underline{\lim_{n\to\infty} \|A_n^f - \mu_x^\varepsilon(f)\|_{L^2_{\mu_x^\varepsilon}} = 0, \quad \forall f \in L^2_{\mu_x^\varepsilon}}$

Detail: We use the fact that $f \mapsto A_n^f - \mu_x^\varepsilon(f)$ (operator $L^2_{\mu_x^\varepsilon} \to L^2_{\mu_x^\varepsilon}$) has norm ≤ 2. For this, we use the fact that μ_x^ε is T-invariant!

Apply this for an arbitrary $\underline{T\text{-invariant}}$ $f \in L^2_{\mu_x^\varepsilon}$!

Then $A_n^f = f$, $(\forall n)$ {trivially; we used this in the proof of MET.}

and so $\lim_{n\to\infty} \|f - \mu_x^\varepsilon(f)\|_{L^2_{\mu_x^\varepsilon}} = 0$,

i.e. $\underline{f = \mu_x^\varepsilon(f) \quad \mu_x^\varepsilon\text{-a.e.}}$ Hence: Every T-invariant $f \in L^2_{\mu_x^\varepsilon}$ is $\underline{\mu_x^\varepsilon\text{-a.e. constant}}$!

$\therefore$ $\underline{\mu_x^\varepsilon \text{ is ergodic}}$!

{Prop #1:1}

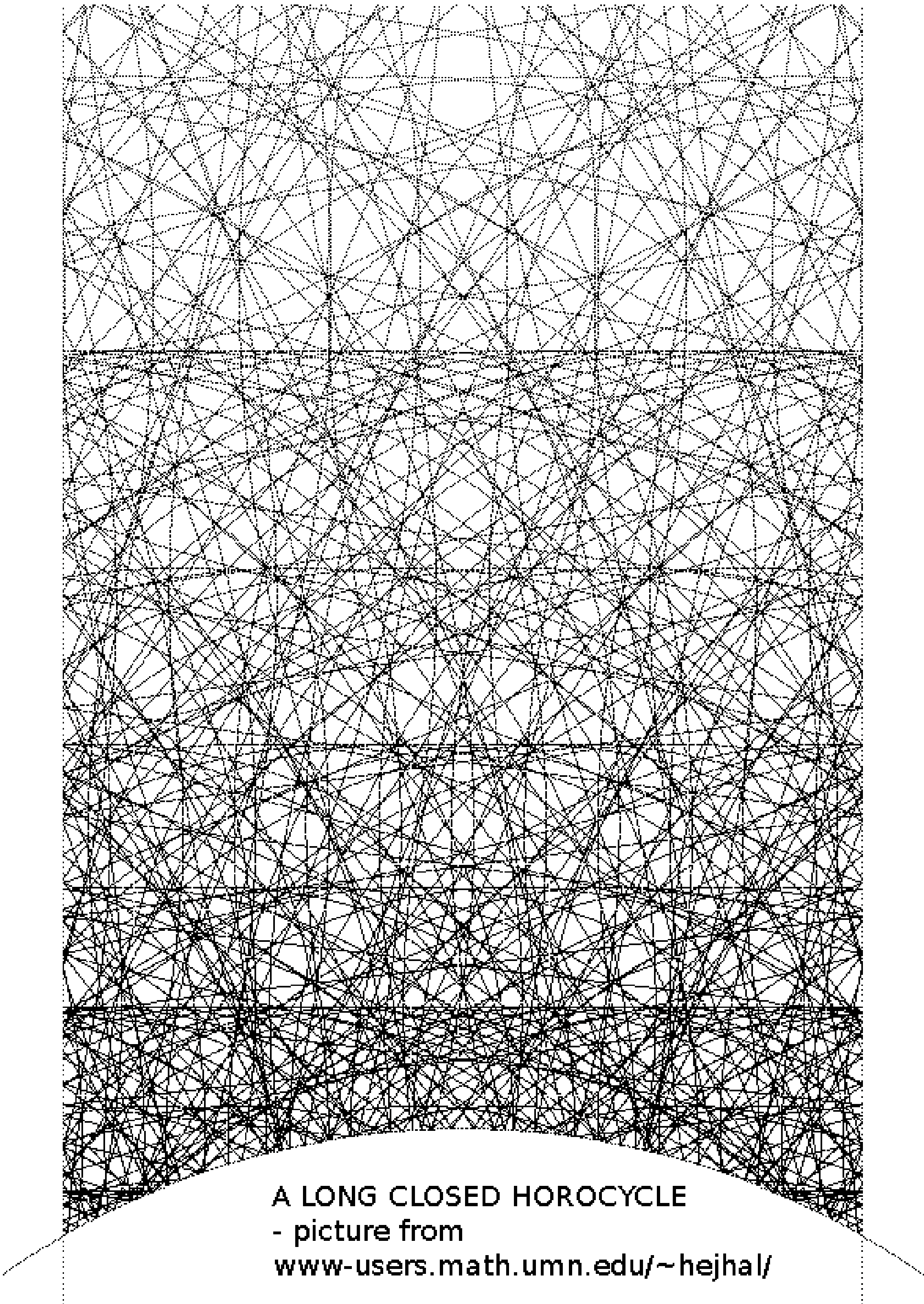
$\square\square$

4.1. **Notes.** .

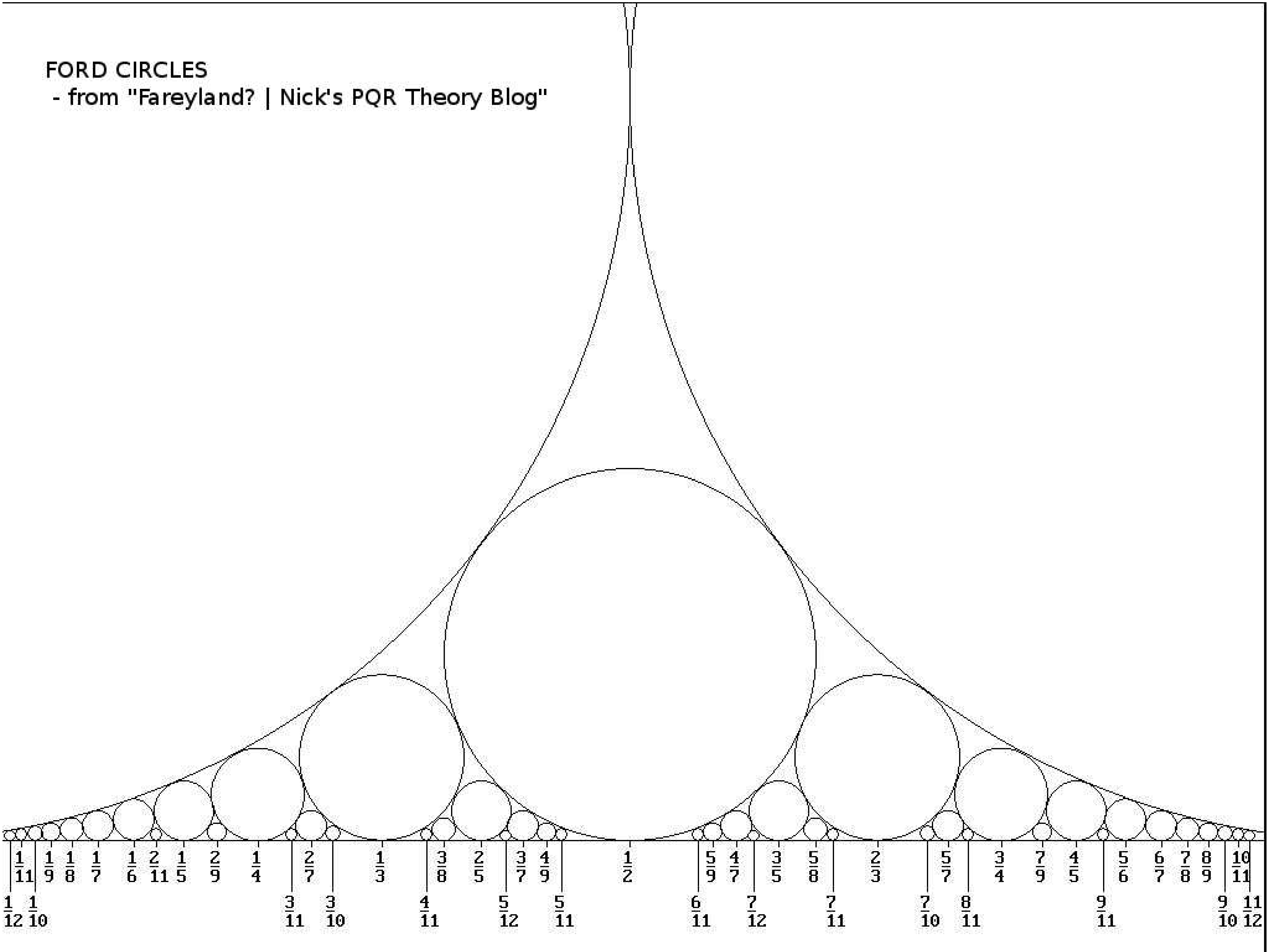
This lecture corresponds to Sarig, [40, Sec. 2.3.2–3]. See my notes to [40] for many more details on the proofs.

Regarding Theorem 2 ("addendum to Theorem 1") in my lecture; cf. Einsiedler and Ward, [12, Thm. 5.14(2)].

Regarding the proof of Theorem 3 (ergodic decomposition, [40, Thm. 2.5]), I was not able to follow Sarig's proof of the ergodicity of μ_x for a.e. $x \in X$. Instead I give a similar proof as in Einsiedler and Ward, [12, Thm. 6.2]. Again see my notes to [40] for more details.



A LONG CLOSED HOROCYCLE
- picture from
www-users.math.umn.edu/~hejhal/


FORD CIRCLES
- from "Fareyland? | Nick's PQR Theory Blog"
1/12 1/11 1/10 1/9 1/8 1/7 1/6 2/11 1/5 2/9 1/4 3/11 2/7 3/10 1/3 4/11 3/8 2/5 5/12 3/7 4/9 5/11 1/2 6/11 5/9 4/7 7/12 3/5 5/8 7/11 2/3 7/10 5/7 8/11 3/4 7/9 4/5 9/11 5/6 6/7 7/8 8/9 9/10 10/11 11/12

Lecture #5: "Introduction to homogeneous dynamics"

Let G be a locally compact group which is second countable.

{Viz. a topological group which is Hausdorff & locally compact.}

{For simplicity! Thus G is lcscH!}

Theorem 1: There is a left-invariant ⊛ Borel measure μ on G which is finite on all compact sets. This μ is unique up to multiplication by scalar (in $\mathbb{R}^+$); it is called (left) Haar measure.

⊛ viz: $\mu(gE) = \mu(E)$, $\forall g \in G$, $E \overset{\text{Borel}}{\subset} G$.

$$\Leftrightarrow \int_G f(h)\,d\mu(h) = \int_G f(gh)\,d\mu(h), \quad \forall g \in G,\ f \in C_c(G).$$

{Proof: See notes! For Lie group: Easy; a left-inv volume form. Of course there's an analogous right Haar measure.}

Let Γ be a discrete subgroup of G.
Set $X = \Gamma\backslash G = \{\Gamma g : g \in G\}$

Let $\pi: G \to X$ be the projection, $\pi(g) = \Gamma g$.

{Note: X inherits any "local" structure from G; thus X is always a lcscH, and G acts on X by homeos. If G Lie gp: X is a C^∞ mfld, G acts by diffeomorphisms.}

{Now μ induces a Borel measure μ_X on X:}

Def: Let $F \subset G$ be a (Borel) fundamental domain for $\Gamma\backslash G$, i.e. $\#(\Gamma g \cap F) = 1$, $\forall g \in G$ $\Leftrightarrow G = \bigsqcup_{\gamma \in \Gamma} \gamma F$.

{exists?? Probl 21}

Then for any Borel set $E \subset X$; $\mu_X(E) := \mu(\pi^{-1}(E) \cap F)$

{Note: μ_X is independent of the choice of F!}

Def: Γ is called a <u>lattice</u> (in G) if $\mu_X(X) < \infty$.

Theorem 2: If Γ is a lattice then μ is also right invariant, and thus <u>μ_X is G-invariant.</u>

{i.e. G <u>unimodular</u>}   (viz: $\mu_X(Eg) = \mu_X(E)$, $\forall E \overset{\text{Borel}}{\subset} X$, $g \in G$)

Conversely, if there is a finite G-invariant Borel measure ν on X, then Γ is a lattice and <u>$\nu = c \cdot \mu_X$</u> for some $c > 0$.

{Recall from #1: Thus get ppt's and flows on (X, μ_X) from 1-para subgroups of G.}

We skip the proof, except proof that μ is right inv.
Fix $g \in G$. Then Fg is <u>also</u> a f.d. for $\Gamma \backslash G$; hence <u>$\mu(Fg) = \mu_X(X) = \mu(F)$</u>. But $E \mapsto \mu(Eg)$ is a left Haar measure on G, hence $\exists\, c > 0$ s.t. <u>$\mu(Eg) = c\,\mu(E)$</u>, $\forall E$ Borel $\subset G$. Thus $c = 1$, etc!

Examples: $G = \langle \mathbb{R}^d, + \rangle$, $\Gamma = \mathbb{Z}^d$ ((or other lattice!))

$$X = \Gamma \backslash G = \mathbb{R}^d / \mathbb{Z}^d = \mathbb{T}^d, \quad \mu_X = \text{"Lebesgue"}$$

$$F = [0,1)^d \quad \text{(say)}$$

$G = SL_d(\mathbb{R})$, $\Gamma = SL_d(\mathbb{Z})$ see p. 3 $\longrightarrow$

$G = SL_d(\mathbb{Q}_p)$

$G = SL_d(\mathbb{R} \times \mathbb{Q}_p)$, $\Gamma = SL_d(\mathbb{Z}[\tfrac{1}{p}])$

$G = SL_d(\mathbb{A})$, $\Gamma = SL_d(\mathbb{Q})$

Ex: Let $G = SL_d(\mathbb{R})$, $\Gamma = SL_d(\mathbb{Z})$, a lattice!

{due to Siegel.}

Normalize μ s.t. $\mu_X(X) = 1$.

Now $X = \Gamma \backslash G \cong \{\Lambda : \Lambda \text{ a lattice in } \mathbb{R}^d, \ vol(\mathbb{R}^d/\Lambda) = 1\}$

by $\Gamma g \leftrightarrow \mathbb{Z}^d g$

{so the rows of g form a lattice basis.}

Some interesting flows on (X, μ_X):

Fix $d_1, d_2 \geq 1$ s.t. $d_1 + d_2 = 1$. Set

$$a_t = \operatorname{diag}\left[e^{-t/d_1}, \ldots, e^{-t/d_1}, e^{t/d_2}, \ldots, e^{t/d_2}\right],$$

$$\gamma_t(\Gamma g) = \Gamma g a_t$$

Then $\{\gamma_t\}$ – a "diagonal flow".

{Cf Samuel's lecture; he proved orbits of γ_t corresp. to Diophantine properties of matrices! $\{\gamma_t\}$ is a "highly chaotic" flow!}

Also fix $y \in \mathbb{R}^{d-1} \setminus \{\underline{0}\}$, $y = (y_1, \ldots, y_{d-1})$, Set

$$u_t = \begin{pmatrix} 1 & ty_1 & ty_2 & \cdots & ty_{d-1} \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

{Acts by shearing on the lattice!}

$\varphi_t(\Gamma g) = \Gamma g u_t$. Then $\{\varphi_t\}$ a {an example of} unipotent flow.

Theorem 3 (Ratner's Measure Classification):
Let G be a Lie group and Γ a lattice in G, and let $\varphi_t(\Gamma g) = \Gamma g u_t$ be a unipotent flow on $X = \Gamma \backslash G$. Then every φ_t-invariant ergodic $\nu \in P(X)$ is homogeneous, i.e. there exist $x \in X$ and a closed connected subgroup $S \subset G$ s.t. $\{u_t\} \subset S$, $xS = \overline{\varphi_{\mathbb{R}}(x)}$, $\nu(xS) = 1$ and ν is S-invariant.

$\varphi_{\mathbb{R}}(x) := \{\varphi_t(x) : t \in \mathbb{R}\}$

In particular: xS closed.

- It follows that ν is the unique S-invariant probability measure on xS; $\nu = \mu_{x,S}$.

- Also: $\varphi_{\mathbb{R}}(x)$ is equidistributed in xS – see DM, Thm 1.3.4.

More explicitly: Say $x = \Gamma g$, set $\boxed{\tilde{S} = gSg^{-1},\ \tilde{\Gamma} = \Gamma \cap \tilde{S},\ \tilde{X} = \tilde{\Gamma} \backslash \tilde{S}}$

Note $xS = (\Gamma g) S = \left[\begin{array}{l}\text{image of } (\Gamma e)\tilde{S} \text{ under the} \\ \text{diffeomorphism } x_1 \mapsto x_1 g^{-1},\ X \xrightarrow{\sim} X\end{array}\right]$

Define $J : \tilde{X} \to X$, $J(\tilde{\Gamma}\tilde{s}) = \Gamma \tilde{s}$
– this is a C^∞ immersion.

Check: well-def, injective, C^∞ immersion, – "clear"!

Now $J(\tilde{X}) = (\Gamma e)\tilde{S}$; hence $(\Gamma e)\tilde{S} \cong \tilde{X}$

isomorphism of measurable spaces.

Thm. 2 $\Rightarrow$ $\tilde{\Gamma}$ is a lattice in $\tilde{S}$ and "$\tilde{\nu} = c \cdot \mu_{\tilde{X}}$".
By some more work it follows that J is proper, and so $(\Gamma e)\tilde{S}$ and xS are closed regular submanifolds of X.

Ex: $G = SL_2(\mathbb{R})$, Γ any lattice with $-\begin{pmatrix}1&0\\0&1\end{pmatrix}\in\Gamma$.

Then $\boxed{\Gamma\backslash G = T^1(\Gamma\backslash\mathcal{H}) \ \cancel{\text{hyp. surface}}\ \underbrace{\Gamma\backslash\mathcal{H}}_{\text{hyp surface of finite area, possibly with singular cone points}}}$ ⊛

Explanation: $\mathcal{H} = \{z = x+iy;\ x,y\in\mathbb{R},\ y>0\}$ with (or (x,y))

Riemannian metric $(ds)^2 = \dfrac{(dx)^2+(dy)^2}{y^2}$ – the Poincaré upper half plane model of the hyperbolic plane. Its group of orientation preserving isometries is

$G' := SL_2(\mathbb{R})/\{\pm\begin{pmatrix}1&0\\0&1\end{pmatrix}\}$, via $\begin{pmatrix}a&b\\c&d\end{pmatrix}(z) := \dfrac{az+b}{cz+d}$.

(G' acts simply transitively on $T^1\mathcal{H}$; hence we get an identification of G' and $T^1\mathcal{H}$, namely:)

$\rightsquigarrow$ (non-canonical (depends on choice of "origin" in $T^1\mathcal{H}$)) identification $G' \longleftrightarrow T^1\mathcal{H}$.

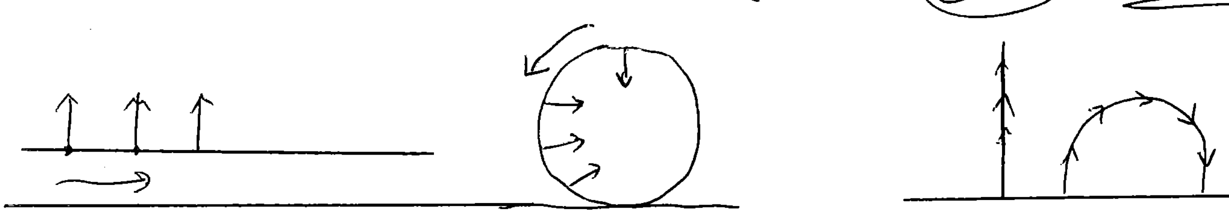
(and: left G'-multiplication $\longleftrightarrow$ action on $T^1\mathcal{H}$!)

$\Rightarrow$ also ⊛!

Flows: $\gamma_t(\Gamma g) = \Gamma g a_t = \Gamma g \begin{pmatrix}e^{t/2}&0\\0&e^{-t/2}\end{pmatrix}$ – geodesic flow

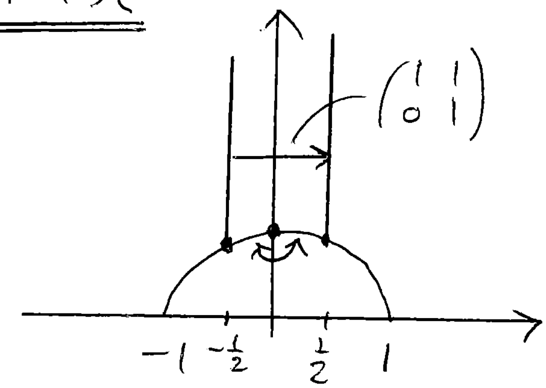
$\varphi_t(\Gamma g) = \Gamma g u_t = \Gamma g\begin{pmatrix}1&t\\0&1\end{pmatrix}$ – horocycle flow

~~Horocycl~~ horocycle flow $\{\varphi_t\}$ $\longleftrightarrow$ geodesic flow $\{\gamma_t\}$

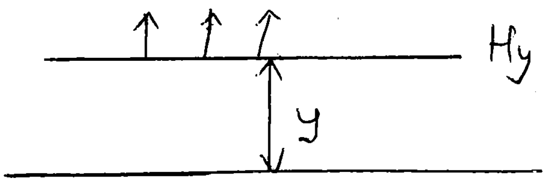


Ex $\Gamma = SL_2(\mathbb{Z})$. Then

$\Gamma \backslash \mathcal{H}$



Closed horocycles in $X = \Gamma \backslash G$



Let $\underline{\underline{H_y = \varphi_{\mathbb{R}}\begin{pmatrix}\sqrt{y} & 0\\ 0 & 1/\sqrt{y}\end{pmatrix}}}$ (any $y>0$)

Note $\varphi_{1/y}\begin{pmatrix}\sqrt{y} & 0\\ 0 & 1/\sqrt{y}\end{pmatrix} = \begin{pmatrix}\sqrt{y} & 1/\sqrt{y}\\ 0 & 1/\sqrt{y}\end{pmatrix} = \begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}\sqrt{y} & 0\\ 0 & 1/\sqrt{y}\end{pmatrix} \overset{\text{on } \Gamma\backslash G}{=} \begin{pmatrix}\sqrt{y} & 0\\ 0 & 1/\sqrt{y}\end{pmatrix}$,

i.e. $\underline{H_y \text{ has length } \tfrac{1}{y}}$.

Let $\underline{h_y = \text{uniform Lebesgue along } H_y}$,

i.e. $\int_X f\,dh_y = \int_0^1 f\left(\begin{pmatrix}1&x\\0&1\end{pmatrix}\begin{pmatrix}\sqrt{y} & 0\\ 0 & 1/\sqrt{y}\end{pmatrix}\right) dx \qquad \forall f \in C_c(X)$.

$\boxed{\text{Ratner (Thm 3)} \overset{\circledast}{\Rightarrow} \text{If } \nu \in P(X) \text{ is ergodic and } \varphi_t\text{-invariant, then } \nu = \mu_X \text{ or } \nu = h_y \text{ for some } y>0.}$

{For general lattice $\Gamma < SL_2(\mathbb{R})$: 1-para of h_y's around each cusp.}

{First proved by Dani, 1978!}

proof of $\circledast$ (outline): $\{u_t\} = \begin{pmatrix}1&*\\0&1\end{pmatrix} \subset S \Rightarrow$

$\underbrace{S = \begin{pmatrix}1&*\\0&1\end{pmatrix}}_{\Rightarrow \nu = h_y \text{ for some } y>0 \text{ (after some work)}}$ or $\underbrace{S = \begin{pmatrix}*&*\\0&*\end{pmatrix}}_{\text{not unimodular!}}$ or $\underbrace{S = G}_{\Rightarrow \nu = \mu_X}$

$\square$

Remarks: Also, $\forall x \in X$, if $\varphi_{\mathbb{R}}(x)$ non-closed, then it is equidistributed in X w.r.t. μ_X.

{Cf. DWM Thm 1.3.4. First proved by Dani-Smillie.}

If X compact, then $\{\varphi_t\}$ is uniquely ergodic.

{First proved by Furstenberg, 1973}

{We next give an application of the above classification (and of ergodic decomposition).}

{They proved much more precise result on the rate of convergence, connecting with spectral theory ~~&~~ and Eisenstein series.}

Theorem 4 (Selberg, Sarnak, Zagier):

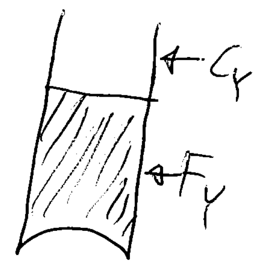
$h_y \xrightarrow[(y\to 0)]{} \mu_X$ in $P(X)$.

proof: Assume not; then $\exists\, y_1 > y_2 > \cdots \to 0$ and $f \in C_b(X)$ and $\varepsilon > 0$ s.t. $|h_{y_j}(f) - \mu_X(f)| > \varepsilon$, $\forall j$. View (by Riesz representation theorem) $\{h_{y_j}\}_{j=1}^{\infty}$ as a sequence in $C_c(X)^*$; note $\|h_{y_j}\| = h_{y_j}(X) = 1$, $\forall j$, and by Alaoglu's Theorem, the unit ball in $C_c(X)^*$ is weak-$*$ compact. {Also ~~[illegible]~~ this unit ball is metrizable, for the weak$*$ topology} Hence wlog {after passing to a subsequence}

We assume $h_{y_j} \xrightarrow[\text{(weak-*)}]{}$ some $\nu \in C_c(X)^*$, $\|\nu\| \leq 1$.

{ Is $\nu \in P(X)$? Need to prove <u>tightness</u>! }

Define subsets $F_Y, C_Y \subset X$ (Y large)

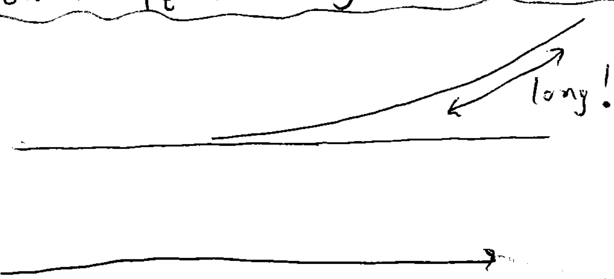


<u>Claim</u>: For Y large, $\limsup_{j \to \infty} h_{y_j}(C_Y) \leq \frac{18}{Y}$

$\therefore$ $\boxed{\nu \in P(X), \text{ and } h_{y_j} \to \nu \text{ in } P(X)}$ {i.e. <u>weak conv</u>!}

<u>proof of claim</u>:

{ We choose to give a proof connecting with number theory and Ford circles. One can also give a more "dynamical" proof, using the fact that if x is <u>far</u> out in a cusp, then $\varphi_t(x)$ stays in that cusp region for a <u>long time</u>.



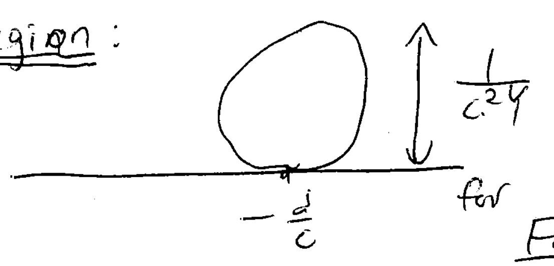
}

$\overset{z = x+iy}{\pi(z)} \in C_Y \Rightarrow \exists\, \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$ s.t. $\operatorname{Im} \frac{az+b}{cz+d} > Y$

$\Rightarrow \underbrace{\frac{y}{|cz+d|^2} > Y}_{\Updownarrow} \iff (cx+d)^2 + (cy)^2 < \frac{y}{Y}$

(keep $y < Y$; then $c \neq 0$; may choose <u>$c > 0$</u>!)

<u>region</u>:



for $Y=1$ these are <u>Ford circles</u>

$\Rightarrow |x + \frac{d}{c}| < \frac{\sqrt{y/Y}}{c}$ and $c < \frac{1}{\sqrt{Yy}}$.

Note $h_{y_j}(C_Y) = \mathrm{Leb}\left(\left\{x \in \mathbb{R}/\mathbb{Z} : \underbrace{\pi\left(\begin{pmatrix}1 & x\\ 0 & 1\end{pmatrix}\begin{pmatrix}\sqrt{y_j} & 0\\ 0 & 1/\sqrt{y_j}\end{pmatrix}\right)}_{\text{"}\pi(x+iy_j)\text{"}} \in C_Y\right\}\right)$.

Now <u>count</u> how large part of $\mathbb{R}/\mathbb{Z}$ can give $\pi(x+iy_j) \in C_Y$.

$\underline{c, d \in \mathbb{Z}}$, $\underline{0 < c < \frac{1}{\sqrt{Yy}}}$, $\underline{\gcd(c,d) = 1}$, <u>"$d$ mod c"</u>.

Note: Any two distinct $\langle c, d\rangle$, $\langle c', d'\rangle$ have

$$\left|\frac{d}{c} - \frac{d'}{c'}\right| = \frac{|c'd - cd'|}{cc'} \geq \frac{1}{cc'}.$$

This corresponds to "Ford circles being disjoint"! This geometrical viewpoint generalizes to <u>arbitrary</u> <u>lattice</u> $\Gamma < SL_2(\mathbb{R})$

For $n \geq 0$, <u>contribution from $\{\langle c, d\rangle : 2^n \leq c < 2^{n+1}\}$</u>:

Any two such $\frac{d}{c}, \frac{d'}{c'}$ are separated by $\geq \frac{1}{cc'} > 2^{-2-2n}$; hence $\underline{\#\{\langle c,d\rangle : 2^n < c < 2^{n+1}\} < 2^{2+2n}}$. Also each such $\frac{d}{c}$ contributes an interval $\subset \mathbb{R}/\mathbb{Z}$ of <u>length</u> $\frac{2\sqrt{y/Y}}{c} \leq \underline{2^{1-n}\sqrt{\frac{y}{Y}}}$.

$\therefore$ Length contribution $< 2^{2+2n} \cdot 2^{1-n}\sqrt{\frac{y}{Y}} = \underline{8 \cdot 2^n \sqrt{\frac{y}{Y}}}$

Note also $2^n \leq c < \frac{1}{\sqrt{yY}}$.

$$\therefore \underline{\underline{h_{y_j}(C_Y)}} < \sum_{\substack{n \geq 0 \\ (2^n \leq 1/\sqrt{yY})}} 8 \cdot 2^n \sqrt{\frac{y}{Y}} \leq 16 \cdot \frac{1}{\sqrt{yY}}\sqrt{\frac{y}{Y}} = \underline{\underline{\frac{16}{Y}}},$$

QED (<u>Claim</u> proved!)

Next: ν is φ_t-invariant!

— This is an easy consequence of the fact that H_{y_j} is a $\{\varphi_t\}$-orbit of length $y_j^{-1} \longrightarrow \infty$.

Easier: $h_{y_j} \to \nu$ and each h_{y_j} is φ_t-invariant! However the other argument is needed in Probl 24.

Hence by ergodic decomposition,

Thm 3, Lecture #4, but for flows.

"$\nu = \int_X \nu_x \, d\nu(x)$", where $\{\nu_x\}_{x\in X} \subset P(X)$ are the conditional probabilities for the $\{\varphi_t\}$-invariant sub-σ-algebra, and ν_x is φ_t-invariant and ergodic for ν-a.e. $x \in X$.

Modify ν_x for $x \in$ null set so this holds for all $x \in X$!

Hence by Ratner (Thm 3): $\nu_x = \mu_X$ or h_y (some $y>0$).

Thus: "$\nu = c\mu_X + \int_{\mathbb{R}^+} h_y \, d\eta(y)$" for some $c \geq 0$ and Borel $\eta \in M(\mathbb{R}^+)$ with $c + \eta(\mathbb{R}^+) = 1$.

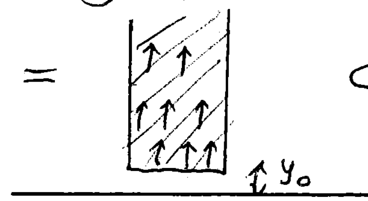
We want to prove $c=1, \eta=0$! This implies $\nu = \mu_X$, i.e. $h_{y_j} \to \mu_X$ in $P(X)$, contradicting our assumption from start (p. 7), thus completing the proof of Thm 4.

Assume $\eta(\mathbb{R}^+) > 0$. Then $\exists y_0 > 0$ s.t.

$\eta_0 := \eta([y_0, \infty)) > 0$, i.e. $\nu(S_{y_0}) = \eta_0 > 0$

see p. 11

Thus as $j \to \infty$, h_{y_j} has an (η_0-)portion concentrating more and more closely to the 2-dim "singular" surface $S_{y_0} =$



$\subset X$

We show this is impossible using a "trick":

Fix $Y > 1$ so large that $\underline{\frac{16}{Y} < \eta_0/2}$

Take $T > 0$ so that $\underline{y_0 e^T = Y+1}.$

Then $\underbrace{\gamma_T}_{\text{Geodesic flow}}(S_{y_0}) = S_{Y+1}$, so that $\underline{[\gamma_{T*}(\nu)](S_{Y+1}) = \eta_0}$

and $\underbrace{\gamma_{T*}(h_{y_j})}_{= h_{e^T y_j}} \longrightarrow \gamma_{T*}(\nu).$

Hence for all large j:

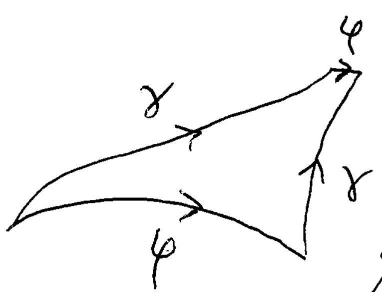
$$\underline{h_{e^T y_j}(C_Y) \geq \frac{\eta_0}{2} > \frac{16}{Y}},$$

contradicting our "Claim" (p. 8)

Done! $\square\ \square\ \square_{(\text{Thm } 4)}$

Related fact:

$\underline{\gamma_T \circ \varphi_s = \varphi_{e^{-T}s} \circ \gamma_T}$



– important!

The above proof is easily generalized to show that any <u>subsegment</u> of H_y of (Euclidean) length $\geq y^{\frac{1}{2}-\varepsilon}$ also goes equidistributed as $y \to 0$. See Problem 24.

5.1. **Notes.**

pp. 1–2: As stated in the lecture, in this course we will generally not work with other groups than *Lie groups,* and in fact seldom other Lie groups than $G = \mathrm{SL}_d(\mathbb{R})$ or $G = \mathrm{ASL}_d(\mathbb{R})$. However it is convenient to be a bit familiar with the more general framework of an arbitrary *locally compact group* G. Theorems 1 and 2 (appropriately formulated) hold in this general framework. A common simplifying assumption is to require G to also be σ-compact; cf., e.g., [30, (0.36)] and [14, Ch. 2.3]. In our lecture we make the even stronger assumption that G is *second countable.* This makes life simple in certain ways. First of all, note that G is now an *lcscH* space. Also, by Struble, [43], there exists a metric d on G which realizes the topology of G, and which is *left invariant,* and which also has the property that all the open balls $B_r(g) := \{h \in G \,:\, d(g,h) < r\}$ ($g \in G$, $r > 0$) have compact closure.

To illustrate, let us prove some useful basic facts in this setting, making use of a fixed metric d as above. Let Γ be a discrete subgroup of G.

Fact 1: $d_\Gamma := \inf\{d(\gamma_1, \gamma_2) \,:\, \gamma_1 \neq \gamma_2 \in \Gamma\} > 0$.
(Proof: Since d is left invariant, $d_\Gamma := \inf\{d(\gamma, e) \,:\, \gamma \in \Gamma \setminus \{e\}\}$. Since Γ is discrete there is an open set $U \subset G$ with $U \cap \Gamma = \{e\}$. Since U is open and $e \in U$, there is some $r > 0$ such that $B_r(e) \subset U$, and it follows that $d_\Gamma \geq r$.)

Fact 2: For any compact set $K \subset G$, the intersection $\Gamma \cap K$ is finite.
(Proof: Otherwise there exist distinct $\gamma_1, \gamma_2, \ldots \in \Gamma \cap K$, and since K is compact we can find a convergent subsequence, say $\{\gamma_{n_j}\}_{j \geq 1}$ where $1 \leq n_1 < n_2 < \cdots$. Then $d(\gamma_{n_j}, \gamma_{n_i}) \to 0$ as $j, i \to \infty$, contradicting $d_\Gamma > 0$.)

Fact 3: Γ is countable.
(Proof: We have $X = \cup_{n=1}^{\infty} \overline{B_n(e)}$ and each closed ball $\overline{B_n(e)}$ is compact; hence the statement follows by using Fact 2.)

Let us also note that since X is a lcscH space, every open subset of X is σ-compact (easy to see using [22, Thm. 5.3(i)$\Rightarrow$(v)]), and hence by [38, Thm. 2.18], if λ is a Borel measure on X satisfying $\lambda(K) < \infty$ for every compact set K, then λ is *regular*, and thus λ is a Radon measure in the sense used in [14, p. vii]. We have used this to make our formulation of Theorem 1 a bit simpler. For a proof of Theorem 1 (in the setting of general locally compact groups), cf., e.g., [14, Thm. 2.10].

We point out that the study of invariant measures on $X = \Gamma \backslash G$ can be carried out in the more general setting of G any locally compact group, and Γ any *closed* subgroup of G, and one does not need to introduce a fundamental domain for $\Gamma \backslash G$ in this development. Cf., e.g., [14, 2.6] and [36, Ch. 1].

For proofs of the claims surrounding the definition of μ_X in the lecture, see Problems 21 and 22.

For completeness, we give here a proof of the last part of Theorem 2 in the lecture (but this proof requires some understanding of [36, Ch. 1]): Assume that there is a finite G-invariant Borel measure ν on X. Then Γ is a "lattice" in the sense of [36, Def. 1.8], and by [36, Remark 1.9], G is unimodular, i.e. the Haar measure μ on G is both left and right invariant. Using this fact, as in the lecture it follows that μ_X is a G-invariant Borel measure on X (possibly with $\mu_X(X) = \infty$). However by [36, Lemma 1.4] (applied with $H = \Gamma$ and $\chi \equiv 1$, and switching sides left $\leftrightarrow$ right) a G-invariant Borel measure on X is *unique* up to a scalar multiple; hence $\nu = c \cdot \mu_X$ for some $c > 0$; and now we also see that $\mu_X(X) < \infty$ since $\nu(X) < \infty$, and so Γ is a lattice (in the sense defined in our lecture).

---

p. 4: Ratner proved her measure classification theorem in [37] (1991); we follow [34, Cor. 1.3.7] rather closely in our statement; cf. also [34, Thm. 1.3.4] for the claim that $\varphi_{\mathbb{R}}(x)$ is equidistributed in xS.

In the discussion making the conclusion more explicit, after having proved that $\widetilde{\Gamma}$ is a lattice in $\widetilde{S}$ we claim that "by some more work" this implies that J is *proper;* for details cf. [36, Thm. 1.13]. For the fact that this implies that $(\Gamma e)\widetilde{S}$ (and thus also xS) is a closed regular submanifold of X, cf., e.g., [7, p. 81, Exc. 1].

Ratner's Theorem plays a crucial role in the proofs of quite a large number of startling results in several different areas of mathematics. See [34, Sec. 1.4] for a discussion of a few of these.

---

p. 5: For more details regarding the identification of $\Gamma\backslash G$ with $T^1(\Gamma\backslash\mathcal{H})$, facts about the geodesic and horocycle flows, etc., see Problem 8 (= [34, pp. 8–9, Exc. 10–11]); and also [29].

---

p. 6: The classification of ergodic φ_t-invariant measures for $G = \mathrm{SL}_2(\mathbb{R})$ (and more generally for G semisimple and *horospherical* flows) was obtained by Dani (1978) [9]; in the special case of X *compact* this had been done by Furstenberg (1973) [16]; cf. also Veech, [44].

---

p. 7: The references to Dani and Dani-Smillie: [10] and [11].

References for Theorem 4: Selberg (unpublished work), Zagier [53], Sarnak [41].

Regarding weak-*-compactness and metrizability of the unit ball in $C_c(X)^*$, cf., e.g., Folland [15, Thm. 5.18 and p. 171 (Exc. 50)]. We discussed the fact that $C_c(X)$ is separable (for X any lcscH space) in our notes to Lecture #2; cf. Sec. 2.1.

(One may note that the subset of *positive* functionals in $C_c(X)^*$ embeds as a subset of the space $M(X)$ of locally finite Borel measures on X, which we introduced in Lecture #2 (p. 7), and the vague topology on $M(X)$ induces the weak-* topology on this subset.)

Let us remark that instead of Alaoglu's Theorem, we could have referred to *Prohorov's* Theorem: Indeed, from the beginning of our proof of Theorem 4 we have sequence $\{h_{y_j}\}_{j\geq 1}$ in $P(X)$, and our "Claim" on p. 8 shows that this sequence is *tight* (viz., for any $\varepsilon > 0$ there is a compact set $K \subset X$ such that $\liminf_{j\to\infty} h_{y_j}(K) > 1-\varepsilon$). Hence by Prohorov's theorem (cf., e.g., [20, Thm. 16.3]), there is a subsequence of $\{h_{y_j}\}_{j\geq 1}$, say $\{h_{y_{j_n}}\}_{n\geq 1}$ where $1 \leq j_1 < j_2 < \cdots$, which converges to some $\nu \in P(X)$ (weak convergence) as $n \to \infty$!

---

p. 10: Here we apply ergodic decomposition for the *flow* $\{\varphi_t\}$; the proof should be an easy modification of the proof of Theorem 3 in Lecture #4 (one first proves the pointwise ergodic theorem for flows; cf. Problem 23). For a precise statement and proof, cf., e.g., [12, Thm. 8.20]; however note that the proof for our special case (namely $G = \langle \mathbb{R}, + \rangle$) should be easier since we do have a pointwise ergodic theorem in this case.

Details on going from "$\nu = \int_X \nu_x \, d\nu(x)$" to "$\nu = c\mu_X + \int_{\mathbb{R}^+} h_y \, d\eta(y)$": As stated in the lecture, we first modify the ν_x's on a null set – e.g. by setting $\nu_x := \mu_X$ for any "bad" x – so that ν_x is $\{\varphi_t\}$-invariant and ergodic for *all* $x \in X$. As noted in the lecture, for each $x \in X$ we now have $\nu_x = \mu_X$ or $\nu_x = h_y$ for some $y > 0$. In other words, if we set $X_1 := \{x \in X \,:\, \nu_x = \mu_X\}$ and $X_2 := X \setminus X_1$ then there is a function $\tau : X_2 \to \mathbb{R}_{>0}$ such that $\nu_x = h_{\tau(x)}$ for all $x \in X_2$. Let us prove that $X_1, X_2 \in \mathcal{B}$ (the Borel σ-algebra of X) and that τ is Borel measurable. For any Borel subset $B \subset \mathbb{R}_{>0}$ we set $H_B := \cup_{y\in B} H_y \subset X$; this is a Borel subset of X. Note that $\mu_X(H_{\mathbb{R}_{>0}}) = 0$ but $h_y(H_{\mathbb{R}_{>0}}) = 1$ for all $y > 0$. Hence $X_1 = \{x \in X \,:\, \nu_x(H_{\mathbb{R}_{>0}}) = 0\}$. Now recall that the ν_x's are conditional probabilities for the appropriate invariant sub-σ-algebra $\mathcal{F} \subset \mathcal{B}$; hence the function $x \mapsto \int_X f \, d\mu_x$ is $\mathcal{B}$-measurable[4] for every $f \in \mathcal{L}^1(X, \mathcal{B}, \mu)$. Applying this with $f = 1_{H_{\mathbb{R}_{>0}}}$ it follows that $X_1 \in \mathcal{B}$; hence also $X_2 \in \mathcal{B}$. Furthermore for any Borel set $B \subset \mathbb{R}_{>0}$ we have $\mu_X(H_B) = 0$ and $h_y(H_B) = 1$ for all $y \in B$ while $h_y(H_B) = 0$ for all $y \in \mathbb{R}_{>0} \setminus B$; hence $\tau^{-1}(B) = \{x \in X \,:\, \nu_x(H_B) = 1\}$, and again using the fact that the ν_x's are conditional probabilities it follows that $\tau^{-1}(B) \in \mathcal{B}$. Hence $\tau : X_2 \to \mathbb{R}_{>0}$ is indeed Borel measurable.

---

[4] even $\mathcal{F}$-measurable.

Now the relation "$\nu = \int_X \nu_x \, d\nu(x)$" means that for every $f \in \mathcal{L}^1(X, \mathcal{B}, \nu)$ we have

$$\begin{aligned} \nu(f) &= \int_X \nu_x(f) \, d\nu(x) \\ &= \int_{X_1} \mu_X(f) \, d\nu(x) + \int_{X_2} h_{\tau(x)}(f) \, d\nu(x) \\ &= \nu(X_1) \cdot \mu_X(f) + \int_{\mathbb{R}_{>0}} h_y(f) \, d\tau_*(\nu)(x) \end{aligned}$$

This proves the desired relation "$\nu = c\mu_X + \int_{\mathbb{R}+} h_y \, d\eta(y)$", with $c := \nu(X_1)$ and $\eta = \tau_*(\nu)$. $\square$

6. The Subadditive Ergodic Theorem

Lecture #6: The Subadditive Ergodic Theorem

Theorem 1 The Subadditive Ergodic Theorem, due to John Kingman.

Let $(X, \mathcal{B}, \mu, T)$ be a ppt, and let $g^{(n)} : X \to \mathbb{R}$ $(n \in \mathbb{Z}^+)$ be $\mathcal{B}$-m'ble functions satisfying

$$\underline{g^{(n+m)} \leq g^{(n)} + g^{(m)} \circ T^n} \qquad \forall\, n, m \in \mathbb{Z}^+,$$

viz., $\{g^{(n)}\}_n$ is a subadditive cocycle

and

$$\underline{\int_X \max(0, g^{(1)})\, d\mu < \infty.}$$

Then $g(x) := \lim_{n\to\infty} \frac{g^{(n)}(x)}{n}$ exists in $[-\infty, \infty)$ for μ-a.e. $x \in X$, and g is (μ-a.e.) T-invariant.

If furthermore μ is ergodic then $g = \underbrace{\inf_{n\geq 1} \frac{1}{n} \int g^{(n)} d\mu}_{\text{constant!}}$ μ-a.e.

Special case: $g^{(n)} = n \cdot A_n^f = \sum_{k=0}^{n-1} f \circ T^k$ for some $f \in L^1_\mu$.

Then $\underline{g^{(n+m)} = g^{(n)} + g^{(m)} \circ T^n}$, equality!, $\forall\, n, m \geq 1$; additive cocycle!

Hence $\underline{\text{Thm } 1 \Rightarrow PET}$!

Indeed recall that once we know that $g(x)$ exists a.e. it follows "easily" that $g \in L^1$ and $A_n^f \xrightarrow[L^1]{} g$!

Cor. 1: On growth-rate of product of random matrices: Fürstenberg–Kesten, actually proved before Thm 1.

Let $(X, \mathcal{B}, \mu, T)$ be a ppt and let $A : X \to GL_d(\mathbb{R})$ be a m'ble function satisfying $\int_X \log^+ \|A\| \, d\mu < \infty$.

Operator norm, $A(x) : \mathbb{R}^d \to \mathbb{R}^d$, or other norm with $\|AB\| \leq \|A\| \cdot \|B\|$.

Set $A_n(x) = A(T^{n-1}x) \cdots A(x)$. Then the following limit exists μ-a.e. and is T-invariant: $\lim_{n\to\infty} \frac{1}{n} \log \|A_n(x)\|$

proof: Apply Thm 1 to $g_n(x) = \log \|A_n(x)\|$. $\square$

proof of Thm 1

Wlog assume $g^{(n)} \leq 0, \ \forall n$.

(Otherwise consider $\underbrace{h^{(n)}}_{\substack{\leq 0 \\ \&\text{subadditive}}} = g^{(n)} - \underbrace{\left(g^{(1)+} + g^{(1)+}\circ T + \dots + g^{(1)+}\circ T^{n-1}\right)}_{\text{PET applies!}}$)

Set $\underline{G(x) := \liminf_{n\to\infty} \frac{g^{(n)}(x)}{n}} \in [-\infty, 0]$

Claim: $G\circ T = G$ μ-a.e. $\circledast$

proof: $\frac{g^{(n+1)}(x)}{n} \leq \frac{g^{(n)}(Tx)}{n} + \frac{g^{(1)}(x)}{n}$ & apply liminf

$\Rightarrow \ G(x) \leq G(Tx), \ \forall x \in X.$

If $\circledast$ fails then $\exists \varepsilon > 0$ st. $E = \{x : G(Tx) > G(x) + \varepsilon\}$ has $\mu(E) > 0$. Then by Poincaré recurrence (Discuss...), for μ-a.e. $x \in E$:

$\exists\ 1 \leq n_1 < n_2 < \dots$ s.t. $T^{n_j}(x) \in E$ $(\forall j)$, and so

$G(x) \leq G(Tx) \leq G(T^2x) \leq \dots$

∞-many of these steps are $> \varepsilon$-increase!

$\Rightarrow \exists n$ s.t. $G(T^n x) > 0$, contradiction!

Replace X by $X_0 := \bigcap_{n=0}^{\infty} T^{-n}\{x : G(Tx) = G(x)\}$;

note $T(X_0) \subset X_0$, $\mu(X_0) = 1$, and $G\circ T = G$ on X_0!

Hence from now on we may assume $\underline{G\circ T = G, \ \forall x \in X}$!

Fix $M \in \mathbb{Z}^+$ {large}; set $\underline{G_M(x) := \max(G(x), -M)}$.

(Thus $-M \leq G_M \leq 0$.) Note $\underline{G_M \circ T = G_M}$.

Claim: $\boxed{\limsup_{n \to \infty} \frac{g^{(n)}(x)}{n} \leq G_M(x) \quad \text{for } \mu\text{-a.e. } x}$

This suffices, for letting $M \to \infty$ it implies $\limsup_{n \to \infty} \frac{g^{(n)}(x)}{n} \leq G(x) = \liminf_{n\to\infty} \frac{g^{(n)}(x)}{n}$ a.e., hence **equal** a.e., i.e. $\underline{\lim_{n\to\infty} \frac{g^{(n)}(x)}{n} \text{ exists a.e}}$, q.e.d.!

Fix $N \in \mathbb{Z}^+$, $\varepsilon > 0$.

Let $\underline{\tau(x) := \min \left\{ \ell \geq 1 : \frac{g^{(\ell)}(x)}{\ell} \leq G_M(x) + \varepsilon \right\}}$.

Note $\tau(x)$ exists in $\mathbb{Z}^+$, for all $x \in X$!

Given $n > N$ and $x \in X$:

Find $1 \leq k_1 < k_1 + \ell_1 \leq k_2 < k_2 + \ell_2 \leq \dots \leq k_b < k_b + \ell_b \leq n$

so that $\boxed{\ell_i = \tau(T^{k_i} x) \leq N}$ $(\forall i \in \{1, \dots, b\})$

Then
$$\begin{aligned}
\underline{g^{(n)}(x)} &\leq g^{(n-1)}(x) + g^{(1)}(T^{n-1}x) \\
&\leq g^{(n-2)}(x) + g^{(1)}(T^{n-2}x) + g^{(1)}(T^{n-1}x) \\
&\leq \dots \leq g^{(k_b+\ell_b)}(x) + \sum_{j=k_b+\ell_b}^{n-1} g^{(1)}(T^j x) \\
&\leq g^{(k_b)}(x) + g^{(\ell_b)}(T^{k_b}x) + \sum_{j=k_b+\ell_b}^{n-1} g^{(1)}(T^j x)
\end{aligned}$$

$$\leqslant g^{(k_b)}(x) + l_b\,(\underbrace{G_M(T^{k_b}x)}_{=G_M(x)} + \varepsilon) + \sum_{j=k_b+l_b}^{n-1} g^{(1)}(T^j x)$$

... repeat! ...

$$\leqslant \Big(\sum_{i=1}^{b} l_i\Big)\big(G_M(x)+\varepsilon\big) + \sum_{j\in\{0,1,\dots,n-1\}\setminus\bigcup_{i=1}^{b}[k_i,k_i+l_i)} g^{(1)}(T^j x).$$

def: $B := \sum l_i$

$$\Rightarrow \frac{g^{(n)}(x)}{n} \leqslant \frac{B}{n}\big(G_M(x)+\varepsilon\big) + 0$$

By choosing the k_i, l_i "greedily", ~~we~~ one can ensure that

$$B \geqslant n - N - \sum_{j=1}^{n-N} I\big(\tau(T^j x) > N\big)$$

overshooting!

$$\Rightarrow \frac{B}{n} \geqslant 1 - \frac{N}{n} - \frac{1}{n}\sum_{j=0}^{n-1} I\big(\tau(T^j x) > N\big)$$

By PET!

$$\xrightarrow[(n\to\infty)]{} 1 - 0 - \mu\big(\tau > N \mid \mathrm{Inv}(T)\big)(x)$$

for μ-a.e. x

$$\xrightarrow[(N\to\infty)]{} 1 - 0 - 0$$

See Sarig's footnote, p. 51; easy from def. of conditional probability, and Lebesgue Bounded Conv.

Trivial if $G_M(x)+\varepsilon > 0$; otherwise use above $\frac{B}{n} \geqslant \dots$!

$\therefore \limsup_{n\to\infty} \frac{g^{(n)}(x)}{n} \leqslant G_M(x) + \varepsilon$ for μ-a.e. x. $\Rightarrow$ GET CLAIM! □□

Finally, for ~~μ ergodic~~ : $\quad g = \inf_{n\geq 1} \frac{1}{n} \int_X g^{(n)} d\mu \quad$ a.e. ?

We know g exists a.e., and is T-inv., hence μ ergodic $\Rightarrow$ $g = c$, a constant, μ-a.e.

Proof of $\quad c \leq \inf_n \frac{1}{n} \int_X g^{(n)} d\mu$:

Fix $n \geq 1$. Now $g^{(n)}, g^{(2n)}, g^{(3n)}, \ldots$ is subadditive cocycle, hence $\quad g^{(kn)} \leq \sum_{\ell=0}^{k-1} g^{(n)} \circ T^{\ell n} \qquad (\forall k \geq 1,\ x \in X)$

For any $j \geq 0$, substituting $x \leftarrow T^j x$ in the above gives: $\quad g^{(kn)}(T^j x) \leq \sum_{\ell=0}^{k-1} g^{(n)}(T^{j+\ell n}(x)). \qquad (\forall x \in X)$

~~Letting $k \to \infty$ gives~~ Dividing by kn and letting $k \to \infty$ gives:

$$c \leq \liminf_{k\to\infty} \frac{1}{kn} \sum_{\ell=0}^{k-1} g^{(n)} \circ T^{j+\ell n}, \qquad \mu\text{-a.e.}$$

Adding over $j = 0, 1, \ldots, n-1$ gives:

$$nc \leq \underbrace{\liminf_{k\to\infty} \frac{1}{kn} \sum_{\ell=0}^{kn-1} g^{(n)} \circ T^{\ell}}, \qquad \mu\text{-a.e.}$$

This limit exists, by "PET, extended to $\int g^{(n)+} < \infty$", and the limit equals $\int g^{(n)}$ a.e.!

$\therefore \quad c \leq \frac{1}{n} \int_X g^{(n)} d\mu, \qquad \forall n \geq 1$, as desired!

Proof of $c \geq \inf_n \frac{1}{n} \int_X g^{(n)} d\mu$:

As earlier, wlog assume $g^{(n)} \leq 0, \quad \forall n.$

For $N \in \mathbb{Z}^+$, set $g_N^{(n)}(x) := \max(g^{(n)}(x), -nN)$.

Then $(g_N^{(n)})_n$ is subadditive

Since (1) "$n \mapsto -nN$" is subadditive, and (2) "max of any two subadditive cocycles is again subadditive"!

a constant!

Now $\lim_{n\to\infty} \frac{g_N^{(n)}(x)}{n} = c_N$ for μ-a.e. x.

By first part of Thm. 1, applied to $(g_N^{(n)})$ + ergodicity.

Also $-N \leq \frac{g_N^{(n)}(x)}{n} \leq 0$ $(\forall x, n)$; hence

Lebesgue Bounded Convergence Thm

$$c_N = \lim_{n\to\infty} \int_X \frac{g_N^{(n)}(x)}{n} d\mu(x) \geq \inf_n \int_X \frac{g_N^{(n)}}{n} d\mu \geq \inf_n \frac{1}{n} \int_X g^{(n)} d\mu.$$

Also $c_N \to c$ as $N \to \infty$.

proof: If $c = -\infty$ then for μ-a.e. x: $\frac{g^{(n)}(x)}{n} \to -\infty$

$\Rightarrow \forall N : \exists M : \forall n \geq M : g_N^{(n)}(x) = -Nn$

$\Rightarrow \forall N : c_N = -N \Rightarrow$ OK!

If $c > -\infty$ then take $N > |c| + 1$; then for μ-a.e. x we get $g_N^{(n)}(x) = g^{(n)}(x)$ for all large n, and so $c_N = c$. OK!

Done!

$\square\square\square$

6.1. **Notes.** .

p. 1, Theorem 1: This is a combination of Sarig, [40, Thm. 2.7 and Prop. 2.3], and we have replaced Sarig's assumption "$g^{(n)} \in L^1 \ \forall n$" by the weaker assumption $g^{(1)+} \in L^1$ (where $g^{(1)+}(x) := \max(0, g^{(1)}(x))$). The fact that Sarig's proof extends to this more general case is discussed in detail in my notes to [40].

p. 1, the remark just below Theorem 1: See (my notes to) Sarig, [40, p. 34, Remark 2] regarding the fact that once we know that the limit $g(x) := \lim_{n\to\infty} A_n^f(x)$ *exists* for μ-a.e. x, it is fairly easy to show that $g \in L^1$ and that the convergence $A_n^f \to g$ also holds in the L^1 norm.

7. The Multiplicative Ergodic Theorem I

Lecture 7: The Multiplicative Ergodic Theorem

Let $(X, \mathcal{B}, \mu, T)$ be a ppt and let $A: X \to GL_d(\mathbb{R})$ be m'ble. For $n \in \mathbb{Z}^+$ define $A_n: X \to GL_d(\mathbb{R})$ by

$$A_n(x) := A(T^{n-1}x) \cdot A(T^{n-2}x) \cdots A(x).$$

Note $A_{n+m}(x) = A_n(T^m x) A_m(x)$, $\forall n, m \in \mathbb{Z}^+$

the cocycle identity

{Note: μ irrelevant for this def!}

{Thus $A_n(x)$ is a product of random matrices, exactly as considered by Fürstenberg-Kesten; cf lecture #6, Cor 1. Thms 1, 2 below give info about "asymptotic direction". More perspective/viewpoints: We'll discuss below!}

{It is also natural to set $A_0(x) \equiv I$; then the cocycle identity holds $\forall n, m \geq 0$. In fact if T is invertible then there is a natural def. of $A_n(x)$, $\forall n \in \mathbb{Z}$ s.t. the cocycle identity holds $\forall n, m \in \mathbb{Z}$!}

{Understanding/motivation for the def. of $A_n(x)$:}

The linear cocycle defined by A over T:

$$\tilde{T}: X \times \mathbb{R}^d \to X \times \mathbb{R}^d; \quad (x, \underline{v}) \mapsto (Tx, A(x)\underline{v})$$

{cf. Sarig §1.6.1; "skew-products"}

Then $\tilde{T}^n(x, \underline{v}) = (T^n x, A_n(x)\underline{v})$, and the cocycle identity corresponds to $\tilde{T}^{n+m} = \tilde{T}^n \circ \tilde{T}^m$.

Def: A flag in $\mathbb{R}^d$ is a family of vector subspaces $0 \subsetneq V^1 \subsetneq V^2 \subsetneq \dots \subsetneq V^k$ of $\mathbb{R}^d$.
It is called complete if $k = d$ (thus $\dim V^j = j,\ \forall j$).

Theorem 1 (Oseledets' Multiplicative Ergodic Theorem):

Assume $\log \|A^{\pm 1}\| \in L^1$. Actually $\log^+ \|A\| \in L^1$ suffices, if we allow $\lambda_1 = -\infty$

see ⊛ below

Then there is $X' \in \mathcal{B}$, $\mu(X') = 1$ with $T(X') \subset X'$, such that for every $x \in X'$ there are $s = s(x) \in \mathbb{Z}^+$,
$\chi_1(x) < \dots < \chi_s(x)$ (in $\mathbb{R}$) and a flag
$0 \subsetneq V_x^1 \subsetneq \dots \subsetneq V_x^s = \mathbb{R}^d$ such that

Lyapunov exponents

$\forall x \in X'$, $i \in \{1, \dots, s(x)\}$:

(a) $s(T(x)) = s(x)$ and $\chi_i(T(x)) = \chi_i(x)$
and $A(x) \cdot V_x^i = V_{T(x)}^i$

(b) the maps $x \mapsto s(x)$, $x \mapsto \chi_i(x)$ and $x \mapsto V_x^i$ are m'ble.

measurability of a map into the Grassmannian $Gr(d)$; cf. notes below!

(c) $\lim_{n\to\infty} \frac{1}{n} \log \|A_n(x)\underline{v}\| = \chi_i(x)$, $\forall\, \underline{v} \in V_x^i \setminus V_x^{i-1}$

with $V_x^0 := \{0\}$

⊛ Any norm on $\mathbb{R}^d$; the limit is independent of the choice of norm.

Theorem 2 (the invertible case):

Assume also that T is invertible.
Then we can take X' in Thm. 1 s.t. for every $x \in X'$ there is a decomposition

$$\mathbb{R}^d = H_x^1 \oplus \dots \oplus H_x^{s(x)} \quad \text{such that } \forall x \in X',\ i \in \{1, \dots, s(x)\}$$

"Oseledets subspaces"

a) $A(x) H_x^i = H_{T(x)}^i$ and $V_x^i = \bigoplus_{j=1}^{i} H_x^j$

b) The map $x \mapsto H_x^i$ is m'ble.

c) $\displaystyle\lim_{n \to \pm\infty} \frac{1}{n} \log \|A_n(x) \underline{v}\| = \chi_i(x), \qquad \forall \underline{v} \in H_x^i \setminus \{0\}$

d) $\displaystyle\lim_{n \to \pm\infty} \frac{1}{n} \log \sin \angle\Big(\bigoplus_{i \in I} H^i_{T^n(x)},\ \bigoplus_{j \in J} H^j_{T^n(x)} \Big) = 0$

for any two $\emptyset \neq I, J \subset \{1, \dots, s(x)\}$, $I \cap J = \emptyset$.

In (c), the angle between any two subspaces $V, W \subset \mathbb{R}^d$ is

$$\angle(V, W) := \min\{\angle(\underline{v}, \underline{w}) : \underline{v} \in V \setminus \{\underline{0}\},\ \underline{w} \in W \setminus \{\underline{0}\}\} \in [0, \tfrac{\pi}{2}],$$

where (of course) $\angle(\underline{v}, \underline{w}) = \arccos\left(\frac{\underline{v} \cdot \underline{w}}{\|\underline{v}\| \cdot \|\underline{w}\|}\right) \in [0, \pi]$.

~~Thm 1~~ (c) $\iff \quad \frac{1}{n} \log \|A_n(x)\underline{v}\| = \chi_i(x) + o(1)$

$$\iff \quad \|A_n(x)\underline{v}\| = \exp\big((\chi_i(x) + o(1))\,n\big)$$

as $n \to \infty$

<u>Thm 2 (c) $\Longrightarrow$ Thm 1(c) !</u>

Indeed, take $\underline{v} \in V_x^i \setminus V_x^{i-1}$,

i.e. $\underline{v} = \sum_{j=1}^{i} \underline{w}_j \;,\quad \underline{w}_j \in H_x^j ,$

$\underline{w}_i \neq \underline{0}$.

$$\underbrace{\|A_n(x)\,\underline{v}\|}\; \approx \; ? \qquad \text{as } n \to \cancel{0}\,\infty$$

$$= \Big\| \sum_{j=1}^{i} A_n(x)\,\underline{w}_j \Big\|$$

$$\gtreqless \|A_n(x)\,\underline{w}_i\| \pm \sum_{j=1}^{i-1} \|A_n(x)\,\underline{w}_j\|$$

$$\gtreqless \exp\big((\chi_i(x)+o(1))\,n\big) \pm \sum_{j=1}^{i-1} \exp\big((\chi_j(x)+o(1))\,n\big)$$

smaller than $\chi_i(x)$!

Hence
$\|A_n(x)\underline{v}\| = \exp\big((\chi_i(x)+o(1))\,n\big)$!

Example

("one-point ppt")

Assume $X = \{x\}$ and $A(x) = A \in GL_d(\mathbb{R})$ in Thms 1,2.
Then the Lyapunov exponents $\chi_1(x) < \dots < \chi_s(x)$ are the logarithms of the absolute values of the eigenvalues of A, and the Oseledets spaces $H_x^1, \dots, H_x^s$ are the corresponding sums of generalized eigenspaces.

— See problem 27

Example: The derivative cocycle

(An important, general example!)

Let M — a C^∞ manifold,
$f : M \to M$ a C^∞ immersion (e.g. a diffeomorphism)

Then $df : T(M) \to T(M)$ is also C^∞.

$$\begin{tikzcd} T(M) \arrow[r, "df"] \arrow[d, "\pi"'] & T(M) \arrow[d, "\pi"] \\ M \arrow[r, "f"] & M \end{tikzcd}$$

(the standard projection)

Note: df is a linear isomorphism on every fiber: $df|_{T_x M} : T_x M \xrightarrow{\sim} T_{f(x)} M$ $(\forall x \in M)$

Note: $(df)^n = d(f^n)$!

Thms. 1,2 extend to this setting, and ~~even~~ more generally to a "linear cocycle" over f on any (fin. dim.) vector bundle over M. — Need to fix a norm on each fiber! Precise assumptions?

If $\underline{T(M) \text{ trivializes}}$, i.e. $\exists$ diffeomorphism φ:

$$\begin{tikzcd} M \times \mathbb{R}^d \arrow[r, "\varphi", "\sim"'] \arrow[dr, "pr_1"'] & T(M) \arrow[d, "\pi"] \\ & M \end{tikzcd}$$

s.t. $\varphi(x, \cdot)$ is a $\underline{\text{linear isomorphism}}$ $(\forall x \in M)$, and the norm on $T_x(M)$ comes from a $\underline{\text{fixed}}$ norm on $\mathbb{R}^d$ via $\varphi(x, \cdot)$. This can be significantly loosened up!

– then we can immediately reduce back to the situation of Thms 1, 2, namely by considering the linear cocycle

$$\tilde{f} : M \times \mathbb{R}^d \xrightarrow[\sim]{\varphi} T(M) \xrightarrow{df} T(M) \xrightarrow[\sim]{\varphi^{-1}} M \times \mathbb{R}^d$$

Of course: $\tilde{f} =$ "df in explicit coordinates".

$\underline{\underline{\text{Special case}}}$: $\underline{\underline{M = \Gamma \backslash G}}$, G a Lie group, Γ a lattice $< G$. Let $\underline{\mathfrak{g} = \mathrm{Lie}(G)}$. the Lie algebra of $\mathfrak{g}$

Fix an $\mathbb{R}$-basis $X_1, \ldots, X_d \in \mathfrak{g}$.

$\leadsto$ Left-invariant vector fields $X_1, \ldots, X_d$ on G,
s.t. $X_1(g), \ldots, X_d(g)$ is a basis of $T_g G$, $\forall g \in G$

$\leadsto$ Vector fields on $M = \Gamma \backslash G$ with the same property.

Viz: $\underline{M \text{ is parallelizable}}$.

$\leadsto$ Trivialization of $T(M)$;

$$\varphi : M \times \mathbb{R}^d \xrightarrow{\sim} T(M)$$
$$\big(\Gamma g, (v_1, \dots, v_d)\big) \longmapsto \sum_{j=1}^{d} v_j \cdot X_j(\Gamma g)$$

This construction clearly applies to <u>any</u> manifold with a parallelization.

Also fix a <u>Riemannian metric on $M = \Gamma \backslash G$</u> by

$$\Big\| \sum_{j=1}^{d} v_j \cdot X_j(p) \Big\| := \sqrt{\sum_{j=1}^{d} v_j^2}\,, \quad \forall\, p \in M,\ (v_1, \dots, v_d) \in \mathbb{R}^d.$$

This is an instance of [an inner product on $\mathfrak{g}$]
$\longrightarrow$ [left inv Riemannian metric on G]
$\longrightarrow$ [Riemannian metric on $M = \Gamma \backslash G$]

Now let $f : M \to M$ be the diffeomorphism given by
<u>$f(\Gamma g) = \Gamma g g_1$</u>, for a fixed $g_1 \in G$.

"$\mathbb{R}^d = \mathfrak{g}$" via $X_1, \dots, X_d$

Then under the trivialization $\varphi : M \times \mathbb{R}^d \xrightarrow{\sim} T(M)$,

$$\boxed{df = \big[\text{the } \underline{\text{constant}} \text{ linear map } Ad(g_1^{-1})\big]!} \quad \circledast$$

Hence: The Lyapunov exponents of f are <u>constant</u>, and determined by the eigenvalues of $Ad(g_1^{-1})$, just as in the "one-point ppt" example!

$\underline{\text{Proof of } \circledast}$: At any $p = \Gamma g \in M$, $X_j(p) \in T_p(M)$ is the derivative of the curve $h \mapsto \Gamma g \exp(hX_j)$ at $h=0$; hence $df(X_j(p)) \in T_p(M)$ is the derivative at $h=0$ of

$$h \mapsto f(\Gamma g \exp(hX_j)) = \Gamma g \exp(hX_j) g_1 =$$
$$= \Gamma g g_1\, g_1^{-1} \exp(hX_j) g_1$$
$$= f(p) \exp\big(h \cdot Ad(g_1^{-1})(X_j)\big).$$

"Done!"

$\underline{\underline{\text{Special case (from Lecture \# 5)}}}$:

$G = SL_2(\mathbb{R})$,

$$\gamma_t(\Gamma g) = \Gamma g a_t = \Gamma g \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$$ geodesic flow

$$\varphi_t(\Gamma g) = \Gamma g u_t = \Gamma g \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$ horocycle flow

$$\overline{\varphi_t}(\Gamma g) = \Gamma g \overline{u}_t = \Gamma g \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$$ "opposite horocycle flow"

$$\mathfrak{g} = sl_2(\mathbb{R}) = \{X \in M_2(\mathbb{R}) : Tr\, X = 0\}$$

$$X_1 = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

The generators of $\{\gamma_t\}$, $\{\varphi_t\}$, $\{\overline{\varphi_t}\}$, resp.!

Then $\underline{Ad(\gamma_t^{-1})(X_1) = X_1}$, $\underline{Ad(\gamma_t^{-1})(X_2) = e^{-t}X_2}$, $\underline{Ad(\gamma_t^{-1})(X_3) = e^{t}X_3}$

Thus: $$d\gamma_t\big(X_1(p)\big) = X_1(\gamma_t(p))$$
$$d\gamma_t\big(X_2(p)\big) = e^{-t}\cdot X_2(\gamma_t(p)) \qquad \forall p \in M, \; t \in \mathbb{R}.$$
$$d\gamma_t\big(X_3(p)\big) = e^{t}\cdot X_3(\gamma_t(p))$$

Hence <u>the Lyapunov exponents of $\gamma_t : \Gamma\backslash G \circlearrowleft$</u> are <u>$-t,\ 0,\ t$</u> (at every $x \in M = \Gamma\backslash G$)!

This also shows that the direct sum decomposition

$$T_p(M) = \mathbb{R}\cdot X_1(p) \oplus \underbrace{\mathbb{R}\cdot X_2(p)}_{E_p^s} \oplus \underbrace{\mathbb{R}\cdot X_3(p)}_{E_p^u}$$

The stable & unstable subspaces

of each fiber of $T(M)$ is preserved by $\{\gamma_t\}$ and vectors in E_p^s (E_p^u) are <u>contracted</u> (expanded) at an exponential rate w.r.t. t (as $t \to \infty$), while $X_1(p)$ is the flow direction.
Thus $\{\gamma_t\}$ is a <u>hyperbolic flow</u>.

(Note also: $X_1(p) = \frac{d}{dt}\gamma_t(p)_{|t=0}$ — flow direction.
$X_2(p) = \frac{d}{ds}\varphi_s(p)_{|s=0}$,
and $d\gamma_t(X_2(p)) = e^{-t}\cdot X_2(\gamma_t(p))$ is closely related to the relation $\gamma_t \circ \varphi_s = \varphi_{e^{-t}s} \circ \gamma_t$ cf. lecture #5, p.11.)

Similarly, a diffeomorphism $f: M \longrightarrow M$ is said to be <u>Anosov</u> if $\forall x \in M$ there is a decomposition $T_p M = E_x^s \oplus E_x^u$ which is preserved by df, and such that $(df)^n$ contracts vectors in E_x^s at an exponential rate $(n \to \infty)$ and $(df)^{-n}$ contracts vectors in E_x^u at an exponential rate $(n \to \infty)$.

<u>Example</u>: $f: \mathbb{T}^n \circlearrowleft$ coming from a hyperbolic linear map $A \in GL_n(\mathbb{R})$, $\det A = \pm 1$.

Cf. Problem 30.

Another example for Theorem 1:

The <u>Rauzy-Veech-Zorich cocycle</u>;

cf. Lecture 13!

7.1. **Notes.** .

p. 2 (Theorem 1(b)): Here when saying that $x \mapsto V_x^i$ is measurable, we need to have a σ-algebra on the *Grassmannian* $\mathrm{Gr}(d)$, the set of all linear subspaces of $\mathbb{R}^d$. In fact $\mathrm{Gr}(d)$ equals the disjoint union $\sqcup_{l=0}^{d} \mathrm{Gr}(d,l)$ where $\mathrm{Gr}(d,l)$ is the set of all l-dimensional linear subspaces of $\mathbb{R}^d$, and as we will now describe, each $\mathrm{Gr}(d,l)$ has the structure of a *connected C^∞-manifold.* The σ-algebra in question is simply the corresponding Borel σ-algebra.

The quickest way of giving $\mathrm{Gr}(d,l)$ a structure of a manifold is to express it as a *homogeneous space.* Thus let $G = \mathrm{GL}_d(\mathbb{R})$, and note that G acts on the set $\mathrm{Gr}(d,l)$ through $V \mapsto gV = \{g\boldsymbol{v} \,:\, \boldsymbol{v} \in V\}$ (any $V \in \mathrm{Gr}(d,l)$, $g \in G$), and this action is transitive. Hence if we fix any $V_0 \in \mathrm{Gr}(d,l)$ and let H be the corresponding stabilizer,

$$H := \{h \in G \,:\, hV_0 = V_0\},$$

then we get an identification (at the level of *sets*)

$$\text{“}G/H = \mathrm{Gr}(d,l)\text{”},$$

through

$$gH \leftrightarrow gV_0 \qquad (\text{any } g \in G).$$

Note that H is a closed subgroup of G; hence G/H has a natural structure as a C^∞-manifold, of dimension $\dim G - \dim H$; cf., e.g., [19, Thm. 4.2]. (Any quotient G/H where G is a Lie group and H is a closed subgroup is called a *homogeneous space,* although in this course we almost exclusively consider the case when $H = \Gamma$, a discrete subgroup of G.) Alternatively one may take $G = O(d)$ in the above discussion. Cf. Problem 31.

p. 2: Regarding Theorem 1(c), the fact that the limit $\lim_{n\to\infty} \frac{1}{n} \log \|A_n(x)v\|$ is independent of the choice of the norm $\|\cdot\|$ on $\mathbb{R}^d$: This is immediate from the fact that any two norms on $\mathbb{R}^d$ are *equivalent;* cf. e.g. [26, Thm. 2.4-5]. (More explicit statement: For any two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on $\mathbb{R}^d$, there exist constants $0 < c_1 \leq c_2$ such that $c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_1$ for all $x \in \mathbb{R}^d$.)

p. 4: Regarding the claim that Theorems 1,2 extend to the more general setting of a linear cocycle on an arbitrary vector bundle over the manifold M, cf., e.g., Viana [50, Thms. 2.1, 2.2]. In that text, Viana is considering a finite-dimensional vector bundle $\pi : \mathcal{E} \to M$ *over an arbitrary probability space M,* and assumes that $\mathcal{E}$ is endowed with a “Riemannian norm”. I am not completely sure what the precise definitions of those things are. In Problem 28, I ask you to find a way to clarify this.

p. 5: For the claim that the assumption that the norm on $T_x(M)$ comes from a fixed norm on $\mathbb{R}^d$ can be significantly loosened up: Again cf. Viana [50, p. 16, around (36)].

p. 8: Definition of a *hyperbolic flow:* Cf., e.g., [21, Def. 17.4.1].

p. 9: Definition of an *Anosov diffeomorphism:* Cf., e.g., [21, Def. 6.4.2].

8. The Multiplicative Ergodic Theorem II

Lecture 8: The Multiplicative Ergodic Theorem (proofs)

Review of spectral theorem for symmetric matrices

Let $C \in M_d(\mathbb{R})$ be symmetric, i.e. $C^t = C$.

Then $Sp(C) \subset \mathbb{R}$ (finite), and

$$\mathbb{R}^d = \underbrace{\bigoplus_{\lambda \in Sp(C)} E_\lambda}_{\text{orthogonal direct sum}} \quad \text{with} \quad E_\lambda = E_\lambda^{(C)} = \{\underline{v} \in \mathbb{R}^d : C\underline{v} = \lambda \underline{v}\}$$

Thus: $$C(\underline{v}) = \sum_{\lambda \in Sp(C)} \lambda \cdot \underbrace{(\underline{v}|E_\lambda)}_{\text{The orthogonal projection of } \underline{v} \text{ on } E_\lambda}, \quad \forall \underline{v} \in \mathbb{R}^d.$$

"Functional calculus": For any $f : Sp(C) \to \mathbb{R}$,
$f(C) \in M_d(\mathbb{R})$ is defined by $f(C)(\underline{v}) = f(\lambda)\underline{v}$, $\forall \underline{v} \in E_\lambda$.

In particular, if C is positive definite,

$\left(\overset{\text{def}}{\Longleftrightarrow} \langle C\underline{v}, \underline{v}\rangle > 0, \ \forall \underline{v} \in \mathbb{R}^d \iff Sp(C) \subset \mathbb{R}_{>0}\right)$

then $f(C)$ is defined for any $f : \mathbb{R}_{>0} \to \mathbb{R}$;
in particular C^α is defined (& positive definite) $\forall \alpha \in \mathbb{R}$! E.g.: $\sqrt{C}$, $C^{1/2n}$!

We now turn to discussing (some points in) the proof of Oseledets' Theorem.

Let $(X, \mathcal{B}, \mu, T)$ be a ppt and let $A: X \to GL_d(\mathbb{R})$ be m'ble with $\log \|A^{\pm 1}\| \in L^1$. Set

$$\underline{A_n(x) = A(T^{n-1}x)A(T^{n-2}x)\cdots A(x)} \qquad (n \in \mathbb{Z}^+)$$

Recall $\underline{A_{n+m}(x) \equiv A_n(T^m x) A_m(x)}$.

Theorem: Oseledets' Mult. Erg. Thm = Thm 1 of Lecture #7, here stated sloppily.

There is $X' \in \mathcal{B}$ with $\mu(X') = 1$, $T(X') \subset X'$ and for every $x \in X'$ there are $s = s(x) \in \mathbb{Z}^+$, $\chi_1(x) < \cdots < \chi_s(x)$ and a flag $0 \subsetneq V_x^1 \subsetneq \cdots \subsetneq V_x^s = \mathbb{R}^d$ such that

a), b) s, χ_i, V^i are T-inv & m'ble;

c) $\forall x \in X'$, $\underline{v} \in V_x^i \setminus V_x^{i-1}$: $\lim\limits_{n\to\infty} \frac{1}{n} \log \|A_n(x)\underline{v}\| = \chi_i(x)$.

$V_x^0 := 0$

proof: WLOG, $\|\cdot\|$ is the Euclidean norm. Then:

$$\|A_n(x)\underline{v}\|^2 = \langle A_n(x)\underline{v}, A_n(x)\underline{v}\rangle = \langle \underbrace{A_n(x)^t A_n(x)}_{\text{pos. def!}}\underline{v}, \underline{v}\rangle.$$

Set $B_n(x) = \sqrt{A_n(x)^t A_n(x)}$.

We'll see $\lim\limits_{n\to\infty} B_n(x)^{1/n}$ exists, for μ-a.e. x!

Let $0 < t_n^1(x) \leqslant \cdots \leqslant t_n^d(x)$ be the eigenval's of $P_n(x)$.

We'll now prove $\lim_{n\to\infty} t_n^j(x)^{1/n}$ exists a.e. using a clever trick by Raghunathan.

$$\forall i \in \{1,\dots,d\}: \quad \prod_{j=d-i+1}^{d} t_n^j(x) = \|\underbrace{A_n(x)^{\wedge i}}\|$$

The i:th exterior product of $A_n(x)$, ~~[illegible]~~ i.e. the linear map on $\Omega^i(\mathbb{R}^d)$ induced by $A_n(x)$, where $\Omega^i(\mathbb{R}^d)$ is the space of alternating i-forms on $\mathbb{R}^d$, provided with natural "Euclidean" norm.

Also $(A_n(x)^{\wedge i})_n$ satisfies cocycle identity since $(A_n(x))_n$ does (and "$\wedge i$ respects multiplication"); hence

$$g_i^{(n)}(x) := \log \|A_n(x)^{\wedge i}\| = \log \prod_{j=d-i+1}^{d} t_n^j(x)$$

is a subadditive cocycle ! Also $g_i^{(n)} \in L^1_\mu$ (from $\log \|A^{\pm}\| \in L^1 \dots$); hence the Subadditive Ergodic Theorem applies, (and in fact get limit $\neq -\infty$):

$$\lim_{n\to\infty} \frac{g_i^{(n)}(x)}{n} \text{ exists in } \mathbb{R}, \text{ for } \mu\text{-a.e. } x\,!$$

$\therefore$ (1) $\boxed{t_j(x) := \lim_{n\to\infty} t_n^j(x)^{1/n} \in \mathbb{R}_{>0} \text{ exists}}$ $(\forall j)$ for μ-a.e. x!

{ Now we only need one small extra input from "dynamics" }

Note: $\lim_{n\to\infty} \frac{1}{n} \log \|A(T^n x)\| = 0 \qquad (\mu\text{-a.e. } x)$

{ Can prove using PET, but there's also a more direct proof. }

Hence $\|A_{n+1}(x)\,\underline{u}\| = \|A(T^n x) A_n(x)\,\underline{u}\| \leqslant \|A(T^n x)\| \cdot \|A_n(x)\underline{u}\|$

$$\textcircled{2} \qquad \leqslant e^{o(n)} \cdot \|A_n(x)\underline{u}\|.$$

Take $X' \in \mathcal{B}$, $\mu(X')=1$ s.t. ①, ② hold $\forall x \in X'$.

Replace X' by $\bigcap_{n\geqslant 0} T^{-n}(X') \Rightarrow$ <u>may assume $T(X') \subset X'$.</u>

{ Now almost all that remains can be done for any <u>fixed</u> $x \in X'$, and it boils down to the following <u>linear algebra result.</u> ("LA" = Linear Algebra) }

<u>LA</u>: Let $A_1, A_2, \ldots \in GL_d(\mathbb{R})$, $B_n = \sqrt{A_n^t A_n}$, and let $0 < t_n^1 \leqslant \ldots t_n^d$ be the eigenvalues of B_n. Assume

a) $\forall j: \quad t_j := \lim_{n\to\infty} (t_n^j)^{1/n}$ exists in $\mathbb{R}_{>0}$

b) $\forall \delta > 0 : \exists N \geqslant 1 : \forall n \geqslant N : \forall \underline{u} \in \mathbb{R}^d : \|A_{n+1}\underline{u}\| \leqslant e^{\delta n} \|A_n \underline{u}\|$.

Then $\Lambda := \lim_{n\to\infty} B_n^{1/n}$ exists, and there are $s \in \mathbb{Z}^+$, $\chi_1 < \ldots < \chi_s$ and a flag $0 \subsetneq V^1 \subsetneq \ldots \subsetneq V^s = \mathbb{R}^d$ s.t.

$$\forall \underline{v} \in V^r \setminus V^{r-1} : \lim_{n\to\infty} \frac{1}{n} \|A_n \underline{v}\| = \chi_r \qquad \circledast$$

{ In fact, we'll see that the V^i:s are built out of eigenspaces of Λ, and $\{\chi_1, \ldots, \chi_s\} = \{\log t_j : t = 1, \ldots, d\}$. }

proof of LA:

Let $\underline{s = \#\{t_j : 1 \le j \le d\}}$.

Take $\underline{\chi_1 < \dots < \chi_s}$ s.t. $\{t_j\} = \{e^{\chi_1}, \dots, e^{\chi_s}\}$.

Set $\underline{I_i := \{j : t_j = e^{\chi_i}\}}$.

— Then $I_1, \dots, I_s$ partition $\{1, \dots, d\}$ $\{$and "$I_1 < I_2 < \dots < I_s$".$\}$

Set $\underline{U_n^i := \sum_{j \in I_i} E_{t_n^j}^{(B_n)}}$

For n large: $\underline{\dim U_n^i = \# I_i}$ and $\underline{\mathbb{R}^d = \bigoplus_{i=1}^{s} U_n^i}$ (ON).

$\{$Namely so large that $t_n^j < t_n^{j'}$ whenever $j \in I_i$, $j' \in I_{i'}$, $i < i'$. We'll always assume that n is this large.$\}$

Set $\underline{V_n^r := \bigoplus_{i \le r} U_n^i}$

$\underline{\tilde{V}_n^r := \bigoplus_{i \ge r} U_n^i}$.

Key Lemma (KL): $\forall \delta > 0 : \exists N \ge 1 : \forall n' > n \ge N$, $1 \le r < r' \le s$:

$\forall \underline{u} \in V_n^r$: $\underline{\| \underline{u} \,|\, \tilde{V}_{n'}^{r'} \| \le \|\underline{u}\| \cdot \exp\left(-n(\chi_{r'} - \chi_r - \delta)\right)}$

Equivalently: $\underline{\cos \angle(V_n^r, \tilde{V}_{n'}^{r'})} =$

$= \sup\{\langle \underline{u}, \underline{u}' \rangle : \underline{u} \in V_n^r, \underline{u}' \in \tilde{V}_{n'}^{r'}, \|\underline{u}\| = \|\underline{u}'\| = 1\} \le \exp\left(-n(\chi_{r'} - \chi_r - \delta)\right)$

$KL \Rightarrow \underline{\underline{V_n^r \xrightarrow{n\to\infty} \text{some } V^r \subset \mathbb{R}^d}}$

This gives the flag $0 \subsetneq V^1 \subsetneq \dots \subsetneq V^s = \mathbb{R}^d$ claimed to exist in LA! Note $\dim V^r = \sum_{i\le r} \# I_i$

The ~~topo~~ convergence is in the topology of the Grassmannian, $Gr(d)$. Note that the convergence is "intuitively obvious", namely $KL \Rightarrow \angle(V_n^r, \tilde{V}_{n'}^{r+1}) \approx \frac{\pi}{2} \Rightarrow V_n^r, (V_{n'}^r)^\perp$ "nearly ON" $\Rightarrow V_n^r \approx V_{n'}^r$, i.e. the $(V_n^r)_n$ is a "Cauchy sequence".

Sarig gives a careful proof showing that a recursively defined ON-basis on V_n^r converges, vector by vector.

Hence: $U_n^r = V_n^r \ominus V_n^{r-1} \xrightarrow{n\to\infty} V^r \ominus V^{r-1} \underline{\underline{:= U^r}}$

Note $\mathbb{R}^d = \bigoplus_{r=1}^{s} U^r$, $\dim U^r = \# I_r$.

Hence; easily $(\forall \underline{v} \in \mathbb{R}^d)$:

$$B_n^{1/n}(\underline{v}) = \sum_{t\in Sp(B_n)} t^{1/n} \left(\underline{v} \,|\, E_t^{(B_n)}\right) \to \underline{\underline{\sum_{i=1}^{s} t_i (\underline{v} \,|\, U^i) \quad := \Lambda(\underline{v})}}$$

Thus: We have proved that $\underline{\Lambda = \lim_{n\to\infty} B_n^{1/n}}$ exists, and we have an explicit formula for Λ (in terms of t_i, U^i)

Also, $\underline{\underline{KL^\infty}}$: $\forall \delta > 0 : \exists N \ge 1 : \forall n \ge N,\ 1 \le r < r' \le s$:
$\underline{\forall \underline{u} \in V_n^r : \ \|\underline{u} \,|\, \tilde{V}^{r'}\| \le \|\underline{u}\| \cdot \exp(-n(\chi_{r'} - \chi_r - \delta))}$

"Key Lemma in the limit" ~~[illegible]~~. Proof: ~~[illegible]~~ simply let $n' \to \infty$ in KL!

Now take $\underline{v} \in V^r \setminus V^{r-1}$; we wish to prove ⊛ in LA. Clearly it suffices to do this for $\underline{\underline{v} \in U^r,\ \underline{v} \neq \underline{0}}$. {By same argument as #7, p. 3.}

Write $\underline{\underline{v} = (\underline{v} \mid V_n^{r-1}) + (\underline{v} \mid U_n^r) + \sum_{r'=r+1}^{s} (\underline{v} \mid U_n^{r'})}$

Now

$$\|A_n(\underline{v} \mid V_n^{r-1})\| \leq e^{(\chi_{r-1} + o(1))n} \cdot \underbrace{\|\underline{v} \mid V_n^{r-1}\|}_{\leq \|\underline{v}\|}$$

"as before"

$$\|A_n(\underline{v} \mid U_n^r)\| \lesseqgtr e^{(\chi_r \pm o(1))n} \cdot \underbrace{\|\underline{v} \mid U_n^r\|}_{\to \|\underline{v}\| \text{ as } n \to \infty}$$

"as before"

$\underline{\text{For } r' > r:}$

$$\|A_n(\underline{v} \mid U_n^{r'})\| \leq e^{(\chi_{r'} + o(1))n} \cdot \|\underline{v} \mid U_n^{r'}\|$$

{By KL^∞}

$$\leq e^{(\chi_{r'} + o(1))n} \cdot e^{-(\chi_{r'} - \chi_r - o(1))n} \cdot \|\underline{v}\|$$

$$\leq e^{(\chi_r + o(1))n} \cdot \|\underline{v}\|.$$

Hence $\|A_n \underline{v}\| \lesseqgtr e^{(\chi_r + o(1))n} \cdot \|\underline{v}\|$ as $n \to \infty$.

{For $\geq$, use also the fact that the vectors in our decomposition are $\underline{\text{orthogonal}}$.}

$\therefore$ ⊛ in LA holds, and so $\underline{\text{all LA is proved}}$ (modulo KL). □

Proof of KL (= Key Lemma, p. 5), start:

Assume $n' = n+1$. Then for $\underline{u} \in V_n^r$,

$$\|A_{n+1}\underline{u}\| = \|A_{n+1}((\underline{u}|V_{n+1}^{r'-1}) + (\underline{u}|\tilde{V}_{n+1}^{r'}))\| \geq e^{(\chi_{r'}-o(1))(n+1)} \|\underline{u}|\tilde{V}_{n+1}^{r'}\|$$

{use orthogonality, "as before"}

$$\geq e^{(\chi_{r'}-o(1))n} \cdot \|\underline{u}|\tilde{V}_{n+1}^{r'}\|$$

but also

$$\|A_{n+1}\underline{u}\| \leq e^{o(n)} \|A_n\underline{u}\| \leq e^{o(n)} e^{(\chi_r+o(1))n} \|\underline{u}\| \leq e^{(\chi_r+o(1))n} \|\underline{u}\|$$

$$\therefore \|\underline{u}|\tilde{V}_{n+1}^{r'}\| \leq e^{(\chi_r-\chi_{r'}+o(1))n} \cdot \|\underline{u}\|$$

$\therefore$ Done, for $n' = n+1$!

The same proof easily extends to $n' = n+k$ for k bounded as $n \to \infty$; however to get uniformity over all $n' > n$ one needs to keep careful track of the exponential decay etc; in the end do double induction over k (in $n' = n+k$) and t (in $r' = r+t$) with precise explicit bound!

To conclude the proof of Oseledets' Thm:

Measurability (of s, χ_i, V^i): "standard" — be technically complicated!

Invariance (of s, χ_i, V^i):

For any $x \in X'$, $\underline{v} \in \mathbb{R}^d \setminus \{\underline{0}\}$,

$$\|A_n(Tx)\,\underline{v}\| = \|A_{n+1}(x)\,A(x)^{-1}\,\underline{v}\|;$$

{By cocycle identity; $A_{n+1}(x) = A_n(Tx)\,A(x)$}

hence $\lim_{n\to\infty} \frac{1}{n} \log \|A_n(Tx)\,\underline{v}\| = \chi_i(x)$ ~~...~~

iff $A(x)^{-1}\underline{v} \in V_x^i \setminus V_x^{i-1}$;

hence, $V^i_{T(x)} = A(x)\,V_x^i$ $(i = 1, \dots, s)$, and $\chi_i(Tx) = \chi_i(x)$!

$s(Tx) = s(x)$,

Finally, we comment briefly on

Oseledets' Thm in the invertible case: Thm 2 of #7, here very sloppily stated.

If also T invertible, then $\exists$ decomposition

$$\underline{\mathbb{R}^d = \bigoplus_{i=1}^{s(x)} H_x^i} \quad \text{s.t.} \quad \underline{V_x^j = \bigoplus_{i \leq j} H_x^i} \quad \text{and} \quad \underline{H_{T(x)}^i = A(x) H_x^i}$$

Remark: Certainly $H_x^i \neq U_x^i$ in general; indeed the H_x^i are typically <u>not orthogonal</u> (cf. Thm 2(d)!)

<u>Key principle used to get started in proof of Thm 2:</u>

For T invertible, PET applies also to T^{-1}, hence for $f \in L^1_\mu$, both

$$\underline{\tilde{f}(x) := \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} f(T^{-k}(x))}$$

or $\sum_1^N$!

and

$$\underline{\bar{f}(x) := \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} f(T^{k}(x))}$$

exist μ-a.e.! Now [ergodic decomposition] & $[\mathrm{Inv}(T) = \mathrm{Inv}(T^{-1})] \Rightarrow \boxed{\tilde{f} = \bar{f} \quad \mu\text{-a.e.}}$

This extends to the <u>Subadditive</u> Ergodic Theorem (see Sarig's Remark, p. 47)!

8.1. **Notes.** .

In this lecture we mainly follow Sarig, [40, Sec. 2.6.2] (cf. also my notes to Sarig's notes). See also Viana, [51, Ch. 4], especially regarding measurability issues.

p. 3, the identity $\prod_{j=d-i+1}^{d} t_n^j(x) = \|A_n(x)^{\wedge i}\|$: Note that Sarig discusses this in detail, starting from the basic definitions, in his [40, Sec. 2.6.1] (see also my notes to Sarig's notes).

p. 6: The intuitive argument given here for the existence of the limit space V^r can be made rigorous; cf. Problems 35 and 36. (But I should stress that my solutions to those problems use the same type of arguments as in Sarig, [40, p. 55]; hence this does not really give a simplification of Sarig's proof; but perhaps a more conceptual perspective.)

p. 10, some more details regarding the PET in the invertible case: Note that the ("original") PET applied to T^{-1} says that

$$\widetilde{f}(x) := \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} f(T^{-k}(x))$$

exists μ-a.e., and is (μ-a.e.) T^{-1}-invariant. Using $\sum_{k=1}^{N} f(T^{-k}(x)) = -f(x) + \sum_{k=0}^{N} f(T^{-k}(x))$ it follows that also

$$\widetilde{f}(x) := \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} f(T^{-k}(x)) \qquad \text{for } \mu\text{-a.e. } x.$$

Arguing now as in Sarig, [40, p. 47, Remark] (using ergodic decomposition and the fact that the two σ-algebras $\mathfrak{Inv}(T)$ and $\mathfrak{Inv}(T^{-1})$ are *the same;* cf. also my notes regarding some details in Sarig's proof), it follows that $\boxed{\widetilde{f}(x) = \overline{f}(x) \text{ for } \mu\text{-a.e. } x.}$ Indeed, this is in fact a *special case* of [40, p. 47, Remark]; since if we set $g^{(n)} = \sum_{k=0}^{n-1} f \circ T^k$ (this is a subadditive – and even additive – cocycle) then $g^{(n)} \circ T^{-n} = \sum_{k=1}^{n} f \circ T^{-k}$.

9. Entropy I

Lecture 9: Measure Theoretic Entropy

Let $(X, \mathcal{B}, \mu)$ be a probability space.

In the following, any <u>partition</u> $\alpha = \{A_1, A_2, \dots\}$ of X will be assumed to be <u>finite or countable</u>, and <u>measurable</u> (i.e. $A_j \in \mathcal{B}, \forall j$).

<u>Def:</u> a) The <u>Information content</u> of a set $A \in \mathcal{B}$ is $\underline{I_\mu(A) := -\log_2 \mu(A)} \in [0, \infty]$

{From now on we'll write "log" for "$\log_2$"}

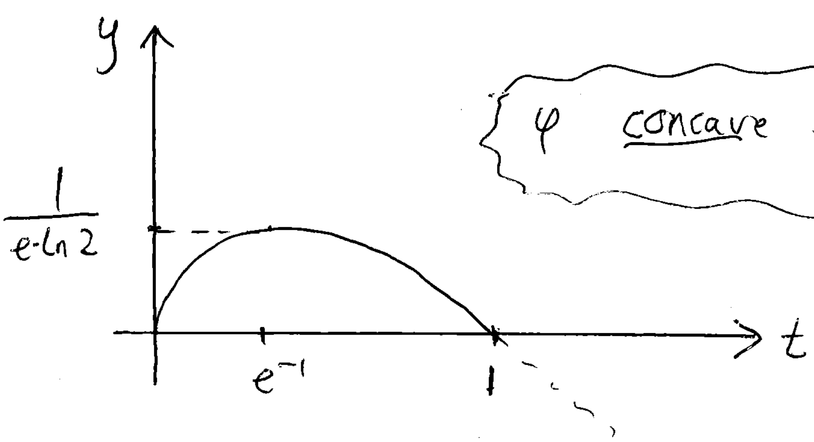
b) The <u>Information Function</u> of a partition α of X is $\underline{I_\mu(\alpha) : X \to [0,\infty]}$; $\underline{I_\mu(\alpha)(x) = \sum_{A\in\alpha} I_\mu(A) \cdot 1_A(x)}$

c) The <u>Entropy</u> of α is

$$\underline{H_\mu(\alpha) := \int_X I_\mu(\alpha)\, d\mu = \sum_{A \in \alpha} \underbrace{\mu(A)(-\log \mu(A))}_{\text{convention: } 0\cdot\log 0 := 0\ !}} \in [0,\infty]$$

We'll often write $\underline{\varphi(t) := -t \log t}$;

{φ <u>concave</u> $\Rightarrow$ $\underline{0 \le H_\mu(\alpha) \le \log(\#\alpha)}$!}



Motivation: $I_\mu(A)$ tells how much information the event "$x \in A$" contains, measured in "bits". It is natural to require $I_\mu(A)$ to be a continuous function of $\mu(A)$, and to require $I_\mu(A \cap B) = I_\mu(A) + I_\mu(B)$ for any two independent events A, B (i.e. $\mu(A \cap B) = \mu(A)\mu(B)$). This makes I_μ uniquely determined up to scaling (and unit "bits" $\Rightarrow I_\mu(A) = 1$ if $\mu(A) = \frac{1}{2}$).
Also the Entropy of α is the expected value of the information content (of telling which α-set x belongs to, when x is random in (X, μ)).

Def: Let α, β be partitions of X. Then
$\alpha \vee \beta := \{A \cap B : A \in \alpha, B \in \beta\}$

Theorem 1: $H_\mu(\alpha \vee \beta) \leq H_\mu(\alpha) + H_\mu(\beta)$ with equality iff α, β are independent ($\overset{def}{\Longleftrightarrow} \forall A \in \alpha, B \in \beta : \mu(A \cap B) = \mu(A)\mu(B)$).

We give the proof later. In fact we'll see that
$H_\mu(\alpha \vee \beta) = H_\mu(\alpha) + H_\mu(\beta \mid \alpha)$ "entropy of β given α".

Def (Kolmogorov, Sinai): Let $(X, \mathcal{B}, \mu, T)$ be a ppt.

For any partition α of X (with $H_\mu(\alpha) < \infty$), set

$$h_\mu(T, \alpha) := \lim_{n \to \infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha \right)$$

(metric) entropy of T w.r.t. α

$T^{-i}\alpha := \{ T^{-i}A : A \in \alpha \}$

Now, the metric entropy of T is:

$$h_\mu(T) := \sup\{ h_\mu(T, \alpha) : \alpha \text{ is a partition of } X \text{ with } H_\mu(\alpha) < \infty \}$$

In fact one can restrict the above supremum to finite α; see Problem 39.

Theorem 2: For any partition α of X with $H_\mu(\alpha) < \infty$, the limit defining $h_\mu(T, \alpha)$ exists in $\mathbb{R}_{\geq 0}$, and

$$h_\mu(T, \alpha) = \inf_{n \geq 1} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\alpha \right).$$

Proof: Set $\alpha_n = \bigvee_{i=0}^{n-1} T^{-i}\alpha$ and $a_n = H_\mu(\alpha_n)$. Then

$$a_{n+m} = H_\mu(\alpha_{n+m}) = H_\mu(\alpha_n \vee T^{-n}\alpha_m) \underset{\text{Thm 1}}{\leq} H_\mu(\alpha_n) + H_\mu(T^{-n}\alpha_m)$$

$$= a_n + a_m$$

"obvious", since T preserves μ

$\therefore$ $\boxed{a_{n+m} \leq a_n + a_m, \quad \forall n, m \in \mathbb{Z}^+}$ i.e. (a_n) is subadditive

Hence by Fekete's Subadditive Lemma, — Nice problem to work out the proof!

$\lim_{n \to \infty} \frac{a_n}{n}$ exists and $= \inf_{n \geq 1} \frac{a_n}{n}$! $\square$

OR, overkill: By Kingman's Subadd Erg Thm, with $\# X = 1$!

Example "angle-doubling"

$f : \mathbb{T}^1 \circlearrowleft, \quad f(x) = 2x \mod 1$

$\alpha = \{[0, \tfrac{1}{2}), [\tfrac{1}{2}, 1)\}$

$f^{-1}(\alpha) = \{[0, \tfrac{1}{4}) \cup [\tfrac{1}{2}, \tfrac{3}{4}), \ [\tfrac{1}{4}, \tfrac{1}{2}) \cup [\tfrac{3}{4}, 1)\}$

$\therefore \bigvee_{i=0}^{1} f^{-i}(\alpha) = \{[0, \tfrac{1}{4}), [\tfrac{1}{4}, \tfrac{1}{2}), [\tfrac{1}{2}, \tfrac{3}{4}), [\tfrac{3}{4}, 1)\}$

More generally, $\bigvee_{i=0}^{n-1} f^{-i}(\alpha) = \{[\tfrac{k}{2^n}, \tfrac{k+1}{2^n}) : k \in \mathbb{Z}\}$

$$H_\mu\left(\bigvee_{i=0}^{n-1} f^{-i}(\alpha)\right) = -\log(2^{-n}) = n$$

$\therefore h_\mu(f, \alpha) = 1$

In fact $h_\mu(f) = 1$, as we'll see below.

Example, circle rotation: $f : \mathbb{T}^1 \circlearrowleft, \ f(x) = x + \varphi$,

any fixed $\varphi \in \mathbb{R}$ (irrational or rational).

Take any partition α of $\mathbb{T}^1$ of the form

$\alpha = \{[0, \alpha_1), [\alpha_1, \alpha_2), \dots, [\alpha_{m-1}, 1)\}$, with $0 < \alpha_1 < \dots < \alpha_{m-1} < 1$.

Then α and also every $f^{-i}(\alpha)$ has m break-points

$\Rightarrow \bigvee_{i=0}^{n-1} f^{-i}(\alpha)$ is a partition of $\mathbb{T}^1$ into $\leq nm$ subintervals $\Rightarrow H_\mu\left(\bigvee_{i=0}^{n-1} f^{-i}(\alpha)\right) \leq \log(nm) \Rightarrow h_\mu(f, \alpha) = 0$

In fact $h_\mu(f) = 0$!

Now back to general theory; $(X, \mathcal{B}, \mu, T)$ a ppt.

Notation: $\alpha_m^n = \bigvee_{i=m}^{n} T^{-i}\alpha$ $\quad (m, n \in \mathbb{Z},\ 0 \leq m \leq n)$

If T invertible, we consider α_m^n also when $m, n < 0$

$\alpha_m^\infty = \sigma\left(\bigcup_{i=m}^{\infty} T^{-i}\alpha\right)$ $\quad$ (also for $m = -\infty$)

The σ-algebra generated by $\bigcup_{i=m}^{\infty} T^{-i}\alpha$

Def: For $\mathcal{A}, \mathcal{B} \subset \mathcal{P}(X)$,

$\mathcal{A} \underset{\mu}{\subset} \mathcal{B} \overset{\text{def}}{\Longleftrightarrow} \forall A \in \mathcal{A} : \exists B \in \mathcal{B} : \mu(A \triangle B) = 0.$

$\mathcal{A} \underset{\mu}{=} \mathcal{B} \overset{\text{def}}{\Longleftrightarrow} \left[\mathcal{A} \underset{\mu}{\subset} \mathcal{B} \text{ and } \mathcal{B} \underset{\mu}{\subset} \mathcal{A}\right]$

Def: A partition α of X is called a strong generator (of $(X, \mathcal{B}, \mu, T)$) if $\alpha_0^\infty \underset{\mu}{=} \mathcal{B}$.

this is $\Leftrightarrow \mathcal{B} \underset{\mu}{\subset} \alpha_0^\infty$, since clearly $\alpha_0^\infty \subset \mathcal{B}$.

If T is invertible, then α is called a generator if $\alpha_{-\infty}^\infty \underset{\mu}{=} \mathcal{B}$.

Theorem 3 (Sinai's generator theorem): If α is a strong generator ~~then~~ with $H_\mu(\alpha) < \infty$ then $h_\mu(T) = H_\mu(T, \alpha)$. Also if T is invertible and α is a generator, then again $h_\mu(T) = H_\mu(T, \alpha)$.

Example: Finite Markov Chains

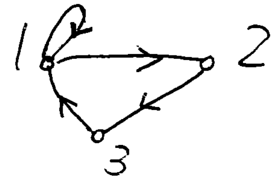
Let S – a finite set.

$A = (a_{ij})_{i,j \in S}$ – a matrix with $a_{ij} \in \{0,1\}$, and no row/column $\equiv 0$.

$$\Sigma_A^+ = \{\underline{x} = (x_0, x_1, \ldots) \in S^{\mathbb{N}} : a_{x_i x_{i+1}} = 1,\ \forall i \in \mathbb{N}\}$$

the subshift of finite type, with alphabet S and transition matrix A.

Ex



$S = \{1,2,3\}$

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Metric on Σ_A^+: $d(\underline{x}, \underline{y}) := 2^{-\min(k:\ x_k \neq y_k)}$.

Map: $\sigma : \Sigma_A^+ \circlearrowleft$, $\sigma(x_0, x_1, x_2, \ldots) = (x_1, x_2, \ldots)$

"left-shift map"

Σ_A^+ is a compact metric space and σ is continuous!

Thus: We are in the standard setting of topological dynamics.

Next: Invariant measures? There are typically many! We'll construct so called Markov Chain measures.

Let $P = (p_{ij})_{i,j \in S}$ be a stochastic matrix, i.e. $p_{ij} \geq 0$ $(\forall i,j)$, $\sum_j p_{ij} = 1$ $(\forall i)$. Assume that P is compatible with A, i.e. $\forall i,j$: $a_{ij} = 0 \Rightarrow p_{ij} = 0$

Let $p = (p_i)_{i \in S}$ be a probability vector (i.e. $p_i \geq 0$, $\sum p_i = 1$)

Given such P, p, we define the Markov (chain) measure μ on $(\Sigma_A^+, \mathcal{B})$ through: $\{\mathcal{B}$ = the Borel σ-algebra$\}$ $\{$in fact a probability measure.$\}$

$$\mu([\underline{a}]) = p_{a_0} P_{a_0 a_1} P_{a_1 a_2} \cdots P_{a_{n-2}a_{n-1}}, \qquad \forall\, \underline{a} = \langle a_0, \ldots, a_{n-1}\rangle \in S^n$$

where $[\underline{a}] := \{\underline{x} \in \Sigma_A^+ : x_i = a_i, \ \forall i \in \{0, 1, \ldots, n-1\}\}$

These $[\underline{a}]$ are called cylinder sets; each cylinder set is open & compact, and the cylinder sets form a (countable) basis for the topology of Σ_A^+. Note that we only prescribe μ on cylinder sets; one verifies that μ is σ-additive on the family of cylinder sets, and then by the Carathéodory Extension Theorem μ extends uniquely to a probability measure on $(\Sigma_A^+, \mathcal{B})$.

$\{$Sarig, Prop 1.8$\}$

Lemma: One has $\sigma_*(\mu) = \mu$ ($\Leftrightarrow (\Sigma_A^+, \mathcal{B}, \mu, \sigma)$ is a ppt) iff p is stationary wrt P, i.e. $pP = p$.
For every P there is at least one stationary p.

Also: μ is ergodic iff P is "irreducible", and mixing iff P is "irreducible & aperiodic". See Sarig Thm 1.2.

<u>Proposition:</u> If $\sigma_*(\mu)=\mu$ then $h_\mu(\sigma) = -\sum_{i,j\in S} p_i P_{ij} \log P_{ij}$.

start

<u>proof</u>: Use the partition $\alpha = \{[a] : a\in S\}$.

cylinder of length 1

This α is a strong generator! Hence by Theorem 3,

$$h_\mu(\sigma) = h_\mu(\sigma,\alpha) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) = \dots$$

compute exactly! Sarig p. 108.

□

<u>Special case; the Bernoulli shift</u>

Take $A=(1)_{i,j\in S}$, p any probability vector,

$P=(P_{ij})$ with $P_{ij}\equiv p_j$. Then the Markov chain measure μ is called <u>Bernoulli measure</u>; note that $X=(X_0,X_1,\dots)$ random in $(S^{\mathbb{N}},\mu)$ means that $X_1,X_2,\dots$ are <u>iid's</u> (with distribution determined by p).

Then get $h_\mu(\sigma) = -\sum_{i\in S} p_i \log p_i$

The exact computation of $H_\mu(\alpha_0^{n-1})$ is <u>very easy</u> in this case!

In particular, the $(\frac{1}{2},\frac{1}{2})$- and $(\frac{1}{3},\frac{1}{3},\frac{1}{3})$- Bernoulli shifts are <u>not isomorphic</u> - this was proved by Sinai!

9.1. **Notes.** .

In this lecture we follow Sarig, [40, Sec. 4.1-4]. We will give proofs of the theorems in the next lecture.

p. 2, Theorem 1: See [40, Prop. 4.3], and also [40, Thm. 4.1].

p. 3, Theorem 2: See [40, Prop. 4.4].

p. 5, Theorem 3: Note that *both* the statements of this theorem are very useful to have. Indeed, if $(X, \mathcal{B}, \mu, T)$ is an invertible ppt with positive entropy, $h_\mu(T) > 0$, then there *does not exist* any *strong* generator with finite entropy (cf. Sarig, [40, the proof of Prop. 4.6]), but there still often exist generators!

pp. 6–7: Here we follow Sarig [40, Sec. 1.5.3-4].

p. 8, the proposition: See [40, Prop. 4.7].

Lecture 10: Measure Theoretic Entropy (II)

Let $(X, \mathcal{B}, \mu)$ be a probability space, and let $\mathcal{F}$ be a σ-subalgebra of $\mathcal{B}$.

Def 1': For $A \in \mathcal{B}$,

a) $\underline{\underline{I_\mu(A|\mathcal{F})}} : X \longrightarrow [0, \infty]$

$$x \longmapsto \underline{-\log \mu(A|\mathcal{F})(x)}$$

Information content of A given $\mathcal{F}$

b) For α a partition of X,

$\underline{\underline{I_\mu(\alpha|\mathcal{F})}} : X \longrightarrow [0, \infty]$

$$x \longmapsto \underline{\sum_{A \in \alpha} I_\mu(A|\mathcal{F})(x)\, 1_A(x)}$$

Information function of α given $\mathcal{F}$

c) $\underline{\underline{H_\mu(\alpha|\mathcal{F})}} := \underline{\int_X I_\mu(\alpha|\mathcal{F})\, d\mu} \in [0, \infty]$

Entropy of α given $\mathcal{F}$

Note: $\mathcal{F} = \{\emptyset, X\} \Rightarrow$ get back the non-conditional notions ($I_\mu(A|\mathcal{F})$ is a constant function, etc).

Convention: $\underline{I_\mu(A|\beta) := I_\mu(A|\sigma(\beta))}$ etc., for β a partition of X.

Note: $\underline{I_\mu(\alpha|\mathcal{F}) < \infty \quad \mu\text{-a.e.}}$ Proof: Given $A \in \alpha$, need to prove $\mu(A|\mathcal{F})(x) > 0$ for μ-a.e. $x \in A$. Let $\underline{N = \{x \in A : \mu(A|\mathcal{F})(x) = 0\}}$

Then $\mu(N) = \int_X 1_A 1_N \, d\mu = \int_X \mu(A|\mathcal{F})(x)\, 1_N(x)\, d\mu(x) = \int_X 0 \, d\mu = 0$. Done!

Some basic formulas:

$$H_\mu(\alpha \mid \mathcal{F}) = \sum_{A\in\alpha} \int_A I_\mu(A\mid\mathcal{F})\, d\mu$$

$$= \sum_{A\in\alpha} \int_A \underbrace{\left(-\log \mu(A\mid\mathcal{F})(x)\right)}_{:=g(x)} d\mu(x)$$

almost "by def", also use monotone conv, since $g \notin L^\infty$.

Then $g: X \to [0,\infty]$, $\mathcal{F}$-m'ble; hence "by def" of $\mu(A\mid\mathcal{F}) = \mathbb{E}(1_A \mid \mathcal{F})$, we have $\underline{\int_A g\, d\mu = \int_X g\cdot \mu(A\mid\mathcal{F})\, d\mu}$

$$= \sum_{A\in\alpha} \int_X \mu(A\mid\mathcal{F})\cdot\left(-\log \mu(A\mid\mathcal{F})\right) d\mu$$

$\sum\int = \int\sum$ by monotone conv

$\Rightarrow$ **Lemma 1:** $$\boxed{H_\mu(\alpha\mid\mathcal{F}) = \int_X \sum_{A\in\alpha} \mu(A\mid\mathcal{F})\left(-\log\mu(A\mid\mathcal{F})\right) d\mu}$$

For β a partition of X and $A \in \mathcal{B}$, we wish to ~~compute $H_\mu(\alpha\mid\beta)$~~ find nice explicit formula for $H_\mu(\alpha|\beta)$.

Convenient notation: For $x \in X$: $\beta(x) := \left[\begin{array}{l}\text{the set } B\in\beta \\ \text{with } x\in B\end{array}\right]$.

Now

$$\mu(A\mid\beta)(x) := \mu(A\mid\sigma(\beta))(x) \overset{\text{if } \mu(\beta(x))>0}{=} \frac{\mu(A\cap\beta(x))}{\mu(\beta(x))} =: \mu(A\mid\beta(x))$$

See Problem 16

New notation (= the classical Bayes' def)

$\because \quad \underline{I_\mu(\alpha|\beta)(x) = -\log \mu(\alpha(x)|\beta(x))} \quad$ for μ-a.e. x.

$\because \quad \underline{\underline{H_\mu(\alpha|\beta)}} = \int_X \left(-\log \mu(\alpha(x)|\beta(x))\right) d\mu(x)$

write as $\sum_{B\in\beta} \sum_{A\in\alpha} \int_{A\cap B}$!

$$= \sum_{B\in\beta} \sum_{A\in\alpha} \underbrace{\int_{A\cap B} \left(-\log \mu(A|B)\right) d\mu}_{= \underbrace{\mu(A\cap B)}_{= \mu(B)\mu(A|B)} \left(-\log \mu(A|B)\right)}$$

$$\cancel{= \sum_{B\in\beta} \mu(B) \sum_{A\in\alpha} \mu(A|B)\left(-\log \mu(A|B)\right)}$$

$\Rightarrow$ $\underline{\underline{\text{Lemma 2}}}$: $\boxed{H_\mu(\alpha|\beta) = \sum_{B\in\beta} \mu(B) \sum_{A\in\alpha} \mu(A|B)\left(-\log \mu(A|B)\right)}$

$$\underline{\underline{= \sum_{B\in\beta} \mu(B)\cdot H_{\mu_{|B}}(\alpha)}}, \quad \text{where } \underline{\underline{\mu_{|B}}} \in P(X) \text{ def by}$$

$$\underline{\mu_{|B}(A) := \frac{\mu(A\cap B)}{\mu(B)}}$$

{numbering continues from Lecture #9}

Theorem 4: $\underline{I_\mu(\alpha \vee \beta \mid \mathcal{F}) = I_\mu(\alpha \mid \mathcal{F}) + I_\mu(\beta \mid \sigma(\mathcal{F} \cup \alpha))}$

{equality in $[0,\infty]$} $\underline{\mu\text{-a.e.}}$

Hence $\underline{H_\mu(\alpha \vee \beta \mid \mathcal{F}) = H_\mu(\alpha \mid \mathcal{F}) + H_\mu(\beta \mid \sigma(\mathcal{F} \cup \alpha))}$, and in particular $\underline{H_\mu(\alpha \vee \beta) = H_\mu(\alpha) + H_\mu(\beta \mid \alpha)}$.

{Note that the formulas ~~for~~ with H_μ are <u>immediate</u> from the first formula (by integrating); hence we only need to prove the first formula!}

{We'll need: For $B \in \beta$}

<u>proof</u>: We first claim that for any $B \in \mathcal{B}$ we have

$$\boxed{\mu(B \mid \sigma(\mathcal{F} \cup \alpha)) = \sum_{A \in \alpha} 1_A \frac{\mu(B \cap A \mid \mathcal{F})}{\mu(A \mid \mathcal{F})} \qquad \mu\text{-a.e.}} \quad \circledast$$

<u>proof of $\circledast$ (outline)</u>: Call the function in the r.h. $\underline{f}$.

Note that f is $\underline{\sigma(\mathcal{F} \cup \alpha)\text{-m'ble}}$ and $\underline{0 \le f \le 1}$ μ-a.e.

(Indeed, $1_{B \cap A} \le 1_A \Rightarrow \mu(B \cap A \mid \mathcal{F}) \le \mu(A \mid \mathcal{F})$ μ-a.e; hence remains to prove $\underline{\mu(A \mid \mathcal{F})(x) > 0 \text{ for } \mu\text{-a.e. } x \in A}$.

~~Let $N := \{x \in A : \mu(A \mid \mathcal{F})(x) = 0\}$. Then $\mu(N) = \int_X 1_A 1_N \, d\mu = \int_X \mu(A \mid \mathcal{F})(x) 1_N(x) \, d\mu(x) = \int_X 0 \, d\mu = 0$. Done.~~ {This was done on p.1!})

It remains to prove that $\forall g \in L^\infty(\sigma(\mathcal{F} \cup \alpha))$:

$\underline{\int_B g \, d\mu = \int_X g f \, d\mu}$. <u>Understanding $\sigma(\mathcal{F} \cup \alpha)$</u>, one finds that it suffices to prove the last identity for $g = 1_A \cdot g_{new}$ with $A \in \alpha$, $g_{new} \in L^\infty(\mathcal{F})$.

Hence let $\underline{A \in \alpha}$ and $\underline{g \in L^\infty(\mathcal{F})}$ be given; we wish to prove $\underline{\int_B 1_A g \, d\mu = \int_X 1_A g f \, d\mu}$. Now:

$$\underline{\underline{\int_X 1_A g f \, d\mu}} = \int_A \underbrace{g \cdot \frac{\mu(B \cap A \mid \mathcal{F})}{\mu(A \mid \mathcal{F})}}_{\in L^\infty(\mathcal{F})!} d\mu =$$

$$= \int_X \mu(A \mid \mathcal{F}) \cdot g \cdot \frac{\mu(B \cap A \mid \mathcal{F})}{\mu(A \mid \mathcal{F})} d\mu = \int_X g \cdot \mu(B \cap A \mid \mathcal{F}) \, d\mu$$

$$= \int_{B \cap A} g \, d\mu = \underline{\underline{\int_B 1_A g \, d\mu}} \; ; \quad \text{done!}$$

$\square$ proof of $\circledast$

$\therefore \; \underline{\underline{I_\mu(\beta \mid \sigma(\mathcal{F} \cup \alpha))}} \overset{\text{By } \circledast,\ \mu\text{-a.e.!}}{=} \sum_{B \in \beta} 1_B \left(-\log \sum_{A \in \alpha} 1_A \frac{\mu(B \cap A \mid \mathcal{F})}{\mu(A \mid \mathcal{F})} \right)$

$$= \sum_{B \in \beta} \sum_{A \in \alpha} 1_{A \cap B} \left(-\log \frac{\mu(A \cap B \mid \mathcal{F})}{\mu(A \mid \mathcal{F})} \right)$$

$$= \sum_{B, A} 1_{A \cap B} \left(-\log \mu(A \cap B \mid \mathcal{F}) \right) - \sum_{B, A} 1_{A \cap B} \left(-\log \mu(A \mid \mathcal{F}) \right)$$

$= \underline{\underline{I_\mu(\alpha \vee \beta \mid \mathcal{F}) - I_\mu(\alpha \mid \mathcal{F})}}$, and all three "$I_\mu$" are $< \infty$ μ-a.e; hence we get the stated equality μ-a.e!

$\square\square$ Theorem 4

DEF: Let α, β be partitions of X.

$\underline{\alpha \leq \beta} \overset{\text{def}}{\iff} \alpha \underset{\mu}{\subset} \sigma(\beta)$

i.e. $\forall A \in \alpha : \exists B \subset \sigma(\beta)$ s.t. $\mu(A \triangle B) = 0$;
cf. Lecture #9, Def. p. 5

$\underline{\alpha = \beta} \overset{\text{def}}{\iff} [\alpha \leq \beta \text{ and } \beta \leq \alpha]$

I would really like to write "$\alpha \underset{\mu}{\leq} \beta$" and "$\alpha \underset{\mu}{=} \beta$", but the notation seems to be fairly standard in the field.

In fact $\underline{\alpha = \beta \iff \forall A \in \alpha : \exists B \in \beta \cup \{\emptyset\} : \mu(A \triangle B) = 0}$ (or $A \leftrightarrow B$!); see Problem 43.

"Monotonicity"

Theorem 5: Let α, β be partitions of X and let $\mathcal{F}_1, \mathcal{F}_2$ be sub-σ-algebras of $\mathcal{B}$.

a) $\underline{\alpha \leq \beta \implies H_\mu(\alpha \mid \mathcal{F}_1) \leq H_\mu(\beta \mid \mathcal{F}_1)}$

b) $\underline{\mathcal{F}_1 \underset{\mu}{\subseteq} \mathcal{F}_2 \implies H_\mu(\alpha \mid \mathcal{F}_1) \geq H_\mu(\alpha \mid \mathcal{F}_2)}$

Note: Thms 4 & 5(**b**) $\implies$ Thm 1 $\boxed{H_\mu(\alpha \vee \beta) \leq H_\mu(\alpha) + H_\mu(\beta)}$

+ "read off condition for equality"!

In fact Thms 4 & 5(b) give, more general result:

Thm 1': $\underline{H_\mu(\alpha \vee \beta \mid \mathcal{F}) \leq H_\mu(\alpha \mid \mathcal{F}) + H_\mu(\beta \mid \mathcal{F})}$

---

proof of Thm 5 (a): Assume $\alpha \leq \beta$. Then $\alpha \vee \beta = \beta$,

and thus $\underline{H_\mu(\beta \mid \mathcal{F})} = H_\mu(\alpha \vee \beta \mid \mathcal{F}) =$

check: The entropy only depends on the partition up to our "="!

Thm 4

$$\overset{\text{Thm 4}}{=} H_\mu(\alpha \mid \mathcal{F}) + H_\mu(\beta \mid \sigma(\mathcal{F} \cup \alpha)) \geq \underline{H_\mu(\alpha \mid \mathcal{F})}; \quad \text{done!}$$

(b): Assume $\mathcal{F}_1 \underset{\mu}{\subset} \mathcal{F}_2$. ~~Take~~ Write $\underline{\varphi(t) = -t \log t}$.

$$\underline{H_\mu(\alpha \mid \mathcal{F}_1)} \overset{\text{Lemma 1}}{=} \int_X \sum_{A \in \alpha} \varphi(\mu(A \mid \mathcal{F}_1)(x))\, d\mu(x)$$

$$= \int_X \sum_{A \in \alpha} \varphi\Big(\mathbb{E}\big(\mathbb{E}(1_A \mid \mathcal{F}_2) \mid \mathcal{F}_1\big)(x)\Big)\, d\mu(x)$$

Basic property of conditioning, when $\mathcal{F}_1 \underset{\mu}{\subset} \mathcal{F}_2$

$$\geq \int_X \sum_{A \in \alpha} \mathbb{E}\Big(\varphi\big(\mathbb{E}(1_A \mid \mathcal{F}_2)\big) \mid \mathcal{F}_1\Big)(x)\, d\mu(x)$$

Jensen's inequality, see below! (Using φ concave.)

$$= \sum_{A \in \alpha} \int_X \varphi\big(\mathbb{E}(1_A \mid \mathcal{F}_2)\big)\, d\mu = \underline{\underline{H_\mu(\alpha \mid \mathcal{F}_2)}} \qquad \square$$

~~Also~~ Above, used "conditional Jensen":

For φ concave: $\mathbb{E}(\varphi \circ f \mid \mathcal{F}) \leq \varphi \circ \mathbb{E}(f \mid \mathcal{F})$ μ-a.e.

(More standard, φ convex: $\mathbb{E}(\varphi \circ f \mid \mathcal{F}) \geq \varphi \circ \mathbb{E}(f \mid \mathcal{F})$)

Note for $\mathcal{F} = \{\emptyset, X\}$ this is standard Jensen between numbers: $\mathbb{E}(\varphi \circ f) \leq \varphi \circ \mathbb{E}(f)$

For general $\mathcal{F}$: Let $\{\mu_x\}_{x \in X}$ be cond. prob for $\mathcal{F}$ (cf. Lecture #4, Thm 1)

Then $\mathbb{E}(\varphi \circ f \mid \mathcal{F})(x) = \int (\varphi \circ f)\, d\mu_x \leq \varphi\left(\int f\, d\mu_x\right) = \left(\varphi \circ \mathbb{E}(f \mid \mathcal{F})\right)(x)$

for μ-a.e. x.

Done!

Also, e.g. Hölder: $\mathbb{E}(f \cdot g) \leq \mathbb{E}(|f|^p)^{\frac{1}{p}} \mathbb{E}(|g|^q)^{\frac{1}{q}}$, $\frac{1}{p} + \frac{1}{q} = 1$

$\Rightarrow \mathbb{E}(fg \mid \mathcal{F}) \leq \mathbb{E}(|f|^p \mid \mathcal{F})^{\frac{1}{p}} \mathbb{E}(|g|^q \mid \mathcal{F})^{\frac{1}{q}}$ μ-a.e.

We'll next prove Thm 3. We leave other important results as exercises: $\boxed{h_\mu(T, \alpha) = H_\mu(\alpha \mid \alpha_1^\infty)}$

(Problem 45)

For μ ergodic:

$$\boxed{\frac{1}{n} I_\mu(\alpha_0^{n-1}) \xrightarrow[n \to \infty]{} h_\mu(T, \alpha) \quad \text{a.e.}}$$

(Problem 46)

Shannon-McMillan-Breiman Thm.

Next, we will recall and prove Thm 3 = Sinai's generator theorem.

Thm 3: If $H_\mu(\alpha) < \infty$ and α is a strong generator, then $h_\mu(T) = h_\mu(T, \alpha)$.

Also if T inv-ble and α is a generator.

Recall: $(X, \mathcal{B}, \mu, T)$ is a ppt.

$$h_\mu(T, \alpha) := \lim_{n\to\infty} \frac{1}{n} H_\mu\Big(\underbrace{\bigvee_{i=0}^{n-1} T^{-i}\alpha}_{:=\alpha_0^{n-1}}\Big)$$

$$h_\mu(T) = \sup\{h_\mu(T, \alpha) : \alpha \text{ with } H_\mu(\alpha) < \infty\}$$

$$\alpha \text{ is a strong generator} \overset{\text{def}}{\iff} \alpha_0^\infty = \sigma\Big(\bigcup_{i=0}^{\infty} T^{-i}\alpha\Big) \underset{\mu}{=} \mathcal{B}.$$

proof of Thm 3: By Problem 39, one may restr. to finite partitions in the def. of $h_\mu(T)$. Hence it suffices to prove that $h_\mu(T, \beta) \leq h_\mu(T, \alpha)$ for any given finite partition β of X. Now:

$$\frac{1}{n} H_\mu(\beta_0^{n-1}) \underset{\text{Thm 4}}{=} \frac{1}{n}\Big(H_\mu(\alpha_0^{n-1} \vee \beta_0^{n-1}) - H_\mu(\alpha_0^{n-1} | \beta_0^{n-1})\Big)$$

$$\underset{\text{Thm 4 again}}{\leq} \frac{1}{n}\Big(H_\mu(\alpha_0^{n-1}) + H_\mu(\beta_0^{n-1} | \alpha_0^{n-1})\Big)$$

$$\underset{\text{Thm 1', p.7}}{\leqslant} \frac{1}{n}\left(H_\mu(\alpha_0^{n-1}) + \sum_{k=0}^{n-1} H_\mu(T^{-k}\beta \mid \alpha_0^{n-1})\right)$$

$$\underset{\text{Thm 5(b)}}{\leqslant} \frac{1}{n}\left(H_\mu(\alpha_0^{n-1}) + \sum_{k=0}^{n-1} \underbrace{H_\mu(T^{-k}\beta \mid T^{-k}\alpha)}_{= H_\mu(\beta\mid\alpha)}\right)$$

$$= \underline{\underline{\frac{1}{n} H_\mu(\alpha_0^{n-1}) + H_\mu(\beta\mid\alpha)}}$$

A useful, general fact!

Hence, letting $n \to \infty$, $\boxed{h_\mu(T,\beta) \leqslant h_\mu(T,\alpha) + H_\mu(\beta\mid\alpha)}$

Mention Rokhlin metric; $d(\alpha,\beta) := H_\mu(\alpha\mid\beta) + H_\mu(\beta\mid\alpha)$

Apply the above with α_0^n in place of α

$$\Rightarrow \boxed{h_\mu(T,\beta) \leqslant \underbrace{h_\mu(T,\alpha_0^n)}_{= h_\mu(T,\alpha)} + H_\mu(\beta\mid\alpha_0^n)}$$

Easy by going into the definition!

$\varphi(t) = -t\log t$

Finally, $\underline{\underline{H_\mu(\beta\mid\alpha_0^n)}} = \int_X I_\mu(\beta\mid\alpha_0^n)\, d\mu$

$$\underset{\text{Lemma 1}}{=} \int_X \sum_{B\in\beta} \varphi\Big(\underbrace{\mu(B\mid\alpha_0^n)(x)}_{\to 1_B \ \mu\text{-a.e. by the Martingale Convergence Theorem. (Probl 37)}}\Big)\, d\mu(x) \ \underline{\underline{\xrightarrow{n\to\infty} 0}}$$

finite sum

Done! $\square\square$

10.1. **Notes.** .

This lecture is a continuation of Lecture #9; we continue to follow Sarig, [40, Sec. 4.1-3].

p. 2–3; Lemmata 1 and 2: Cf. [40, p. 98 (bottom)].

p. 4, Theorem 4: This is [40, Theorem 4.1].

p. 6: For the definition of "$\alpha \leq \beta$" and "$\alpha = \beta$", cf. [40, p. 99 (top)]. Our Theorem 5 is a somewhat generalized version of [40, Prop. 4.2].

pp. 7–8, regarding the conditional Jensen's inequality, cf. [40, Prop. 2.2(3)], and in particular my notes related to that result. (Note that φ is assumed to be convex in [40, Prop. 2.2(3)], whereas our $\varphi(t) = -t \log t$ is concave; hence we get $\leq$ instead of $\geq$ in Jensen's inequality.)

11. Pesin's entropy formula

Lecture #11; Pesin's formula

First some "asides":

- Any mixing finite 2-sided Markov Chain is (measure theoretically) isomorphic to a Bernoulli scheme (Friedman & Ornstein 1970)

- Topological entropy and the Variational Principle – see Sarig Sec. 4.6 (and Problem 47).

Pesin's formula

Let M – a C^∞ compact manifold

$f : M \to M$ a C^1 map.

Lecture #7, Thm. 1

Recall that by Oseledet's Theorem, for any f-inv. $\mu \in P(M)$, for μ-a.e. $x \in M$ there are $\chi_1(x) < \dots < \chi_s(x)$ $(s = s(x))$ and a flag $0 \neq V_x^1 \subsetneq V_x^2 \subsetneq \dots \subsetneq V_x^s = T_x M$ such that

$$\lim_{n\to\infty} \frac{1}{n} \log \|(df^n)(\underline{v})\| = \chi_j(x), \quad \forall \underline{v} \in V_x^j \setminus V_x^{j-1}.$$

> Recall $d(f^n) = (df)^n : T(M) \to T(M)$. Also: $\|\cdot\|$ is "any" norm on $T_x M$ (up to very mild restrictions).
> Assumption in Oseledet's Thm: $\log^+ \|df\| \in L^1$; true!
> — One may have $\chi_1(x) = -\infty$. If f is an immersion then $\log \|df^{\pm 1}\| \in L^1$ and $\chi_1(x) \in \mathbb{R}$.

Theorem (Ruelle's bound & Pesin's formula):

Set $$X(x) = \sum_{\substack{j=1 \\ (\chi_j(x) > 0)}}^{s} \chi_j(x) \cdot \dim H_x^j$$

Thus $\chi : M' \to \mathbb{R}$ where $\mu(M') = 1$.

Then $h_\mu(f) \leqslant \int_M \chi \, d\mu$. If also f is a diffeomorphism and $f \in C^{1+\varepsilon}$, and $\mu \ll Leb$, then $h_\mu(f) = \int_M \chi \, d\mu$.

explain!

Example: Arnold's cat map, $f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \underbrace{\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}}_{A} \begin{pmatrix} x \\ y \end{pmatrix}$ on $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$.

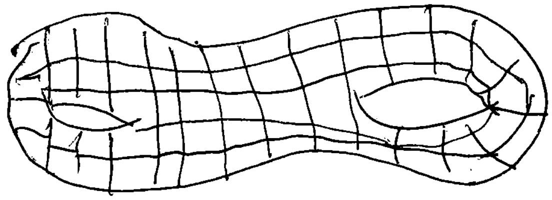
A has eigenvalues $\lambda_{1,2} = \dfrac{3 \pm \sqrt{5}}{2}$; hence

$\chi_1 = \log \dfrac{3-\sqrt{5}}{2}$, $\chi_2 = \log \dfrac{3+\sqrt{5}}{2}$; constant functions on $\mathbb{T}^2$. (Cf. Lecture #7 & Problem 27).

Also $\chi = \chi_2$ (constant), and so Pesin's formula implies $h_m(f) = \chi = \log \dfrac{3+\sqrt{5}}{2}$

m = Lebesgue measure on $\mathbb{T}^2$

"proof" of $h_\mu(f) \leq \int_M \chi \, d\mu$ (Ruelle's bound)

M 

{handle bdry appropriately!}

δ_N = partition of M into nice "cubes" of side $\lesssim \frac{1}{N}$

{for fixed Riemannian metric on M}

Assume $\underline{\mu(\partial S) = 0}$, $\forall S \in \delta_N$ $(\forall N)$

{Nontriv since we need not have $\mu \ll$ Leb. Discuss how Ruelle achieves this.}

{Depends only on M with its Riemannian metric, and (δ_N).}

Lemma: $\exists\, C > 0$ s.t. for any C^1-map $g: M \to M$, $\exists N_0 : \forall N \geq N_0 : \forall S \in \delta_N,\ x \in S$:

$$\#\{S' \in \delta_N : S' \cap g(S) \neq \emptyset\} \leq C \cdot \|(d_x g)^\wedge\|$$

{$(d_x g)^\wedge$ = the collection of all $(d_x g)^{\wedge i} : \Omega^i(T_x M) \circlearrowleft$, $i = 0, 1, \dots, d$ $(d = \dim M)$. Cf. Sarig Sec. 2.6.1.}

"proof": Say "$g: \mathbb{R}^d \to \mathbb{R}^d$" {compact support}

For any η-cube $C \subset \mathbb{R}^d$ (η small) and any $x \in C$, $g(C) \subset$ "ON-box" with sides

{= rectangular parallelepiped}

$$K_d \cdot \max(\lambda_j, 1) \cdot \eta \qquad (j = 1, \dots, d)$$

{a constant which only depends on d}

where $\underline{\lambda_1, \dots, \lambda_d = \text{eigenvalues of } \sqrt{(dg_x)^t (dg_x)}}$

{cf. Problem 48!}

This ON-box can be covered by

$\boxed{\leq \prod_{\substack{j=1 \\ (\lambda_j \geq 1)}}^{d} (2\lambda_j)}$ $\underline{K_d \eta\text{-cubes}}$, and each

$\boxed{\leq 2^d \|(d_x g)^\wedge\|}$ $K_d \eta$-cube intersects $\leq O(1)$

δ_N-sets!

By Sonig Thm 2.9; namely if $0 \leq \lambda_1 \leq \dots \leq \lambda_d$ then $\|(d_x g)^{\wedge i}\| = \prod_{j=d-i+1}^{d} \lambda_j$

Apply with $\eta \asymp \frac{1}{N}$

– done! □

Now use $\underline{\underline{h_\mu(f)}} \underset{\uparrow}{=} \frac{1}{n} h_\mu(f^n) = \underline{\underline{\frac{1}{n} \lim_{N\to\infty} h_\mu(f^n, \delta_N)}}$

Any $n \in \mathbb{Z}^+$; Probl. 49.

Or $N = 2^k$, $k \to \infty$... Proof similar to Sinai's generator theorem.

For any $n, N \in \mathbb{Z}^+$,

$$\underline{\underline{h_\mu(f^n, \delta_N)}} = \lim_{l\to\infty} \frac{1}{l} H_\mu\left(\bigvee_{k=0}^{l-1} f^{-kn}(\delta_N)\right)$$

$$= \lim_{l\to\infty} \frac{1}{l} \sum_{k=0}^{l-1} H_\mu\left(f^{-kn}(\delta_N) \;\Big|\; \bigvee_{j=0}^{k-1} f^{-jn}(\delta_N)\right)$$

$$= \lim_{l\to\infty} \frac{1}{l} \sum_{k=0}^{l-1} \sum_{S \in \bigvee_{j=0}^{k-1} f^{-jn}(\delta_N)} \mu(S) \cdot H_{\mu_{|S}}\left(f^{-kn}(\delta_N)\right)$$

Write $\underline{S[k,x]} := \underline{\text{the}}$ $S \in \bigvee_{j=0}^{k-1} f^{-jn}(\delta_N)$ with $x \in S$.

$\underline{h_{N,n,k}(x)} := H_{\mu_{|S[k,x]}}\left(f^{-kn}(\delta_N)\right)$

$$\underline{\underline{= \lim_{l\to\infty} \frac{1}{l} \sum_{k=0}^{l-1} \int_M h_{N,n,k}(x)\, d\mu(x)}}$$

Now $h_{N,n,k}(x) \leqslant \log \#\{S' \in f^{-kn}(\delta_N) \ : \ S' \cap S[k,x] \neq \emptyset\}$

$$= \underline{\log \#\{S' \in \delta_N \ : \ f^{-kn}(S') \cap S[k,x] \neq \emptyset\}}.$$

Recall $S[k,x] \in \bigvee_{j=0}^{k-1} f^{-jn}(\delta_N)$; take $\underline{S'' \in \delta_N}$ s.t. $S[k,x] \subset f^{-(k-1)n}(S'')$. Set $\underline{y := f^{(k-1)n}(x)} \in S''$.

Note $f^{-kn}(S') \cap S[k,x] \neq \emptyset$

$\Rightarrow f^{-n}(S') \cap S'' \neq \emptyset$

$(\Leftrightarrow \ S' \cap f^{n}(S'') \neq \emptyset$

$$\leqslant \log \#\{S' \in \delta_N \ : \ S' \cap f^n(S'') \neq \emptyset\}$$

Recall $y \in S''$; use $\underline{\text{Lemma}}$ on p. 4!

$$\leqslant \underline{\underline{\log\left(C \cdot \|(d_y f^n)^\wedge\|\right)}}$$

For $N \gg_n 1$

independent of n

Hence for $N \gg_n 1$,

$$\underline{\underline{h_\mu(f^n, \delta_N)}} \leqslant \lim_{\ell \to \infty} \frac{1}{\ell} \sum_{k=0}^{\ell-1} \int_M \log\left(C \|(d_y f^n)^\wedge\|\right) d\mu(x)$$

$y = f^{(k-1)n}(x)$; subst!

$$= \underline{\underline{\log C + \int_M \log \|(d_y f^n)^\wedge\| \, d\mu(y)}}$$

$$\therefore \underline{\underline{h_\mu(f) \leqslant \frac{1}{n} \log C + \int_M \underbrace{\frac{1}{n} \log \|(d_y f^n)^\wedge\|}_{\to\, \chi(y) \text{ for } \mu\text{-a.e. } y \in M} \, d\mu(y)}}, \quad \forall n \in \mathbb{Z}^+$$

Let $n \to \infty$, apply Lebesgue Bounded Convergence

$$\Rightarrow \underline{\underline{h_\mu(f) \leqslant \int_M \chi \, d\mu}}, \qquad \square\square$$

11.1. **Notes.** .

I follow the papers by Ruelle [39] and Mañé [28] [27]. See also my notes to those two papers.

12. Pesin's entropy formula II

Lecture #12; Pesin's formula (cont'd)

We now turn to the <u>bound from below</u>. Review:

M — a C^∞ compact manifold.

$f : M \to M$ a $C^{1+\varepsilon}$ map which is a <u>diffeomorphism</u>

$\mu \in P(M)$, f-invariant.

Assume $\mu \ll \text{Leb}$. (the <u>invertible</u> case; Lect #7, Thm 2)

Recall that by <u>Oseledet's Theorem</u>, for μ-a.e. $x \in M$ there are $\chi_1(x) < \dots < \chi_s(x)$ ($s = s(x)$) and a decomposition $T_x M = H_x^1 \oplus \dots \oplus H_x^s$ such that

$$\lim_{n\to\infty} \frac{1}{n} \log \|(df)^n \underline{v}\| = \chi_j(x), \quad \forall\, \underline{v} \in H_x^j.$$

We'll prove:

$$h_\mu(f) \geq \int_M \chi \, d\mu,$$

where

$$\chi(x) = \sum_{\substack{j=1 \\ (\chi_j(x)>0)}}^{s} \chi_j(x) \cdot \dim H_x^j$$

Recall: In Lecture #11 we proved $h_\mu(f) \leq \int_M \chi \, d\mu$ (Ruelle's bound), under more general assumptions!

Discuss: Glitch in def. of "log" in "Lyapunov theory" and "entropy theory". Now: log = <u>natural log</u>, <u>also</u> in def. of entropy!

We start by giving a general lower bound for $h_\mu(f)$ of <u>geometric</u> nature — which does not involve partitions...

Prop: Let $g: M \to M$ be a m'ble map, and $\mu \in P(M)$, $g_*\mu = \mu$. Let $\rho: M \to (0,1)$ be m'ble, with $\log \rho \in L^1_\mu$.
Set

fixed Riemmanian metric

$$S_n(g,\rho,x) = \{y \in M : d(g^i(x), g^i(y)) \leq \rho(g^i(x)),\ 0 \leq i \leq n\}$$
$$= \bigcap_{i=0}^{n} g^{-i}(B(g^i(x))) \quad \text{with} \quad B(y) := B_{\rho(y)}(y)$$

Ball of radius $\rho(y)$ around y

Let ν be a (σ-)finite Borel measure on M with $\mu \ll \nu$ and set
$$h_\nu(g,\rho,x) = \limsup_{n\to\infty} \frac{1}{n}\left(-\log \nu(S_n(g,\rho,x))\right)$$

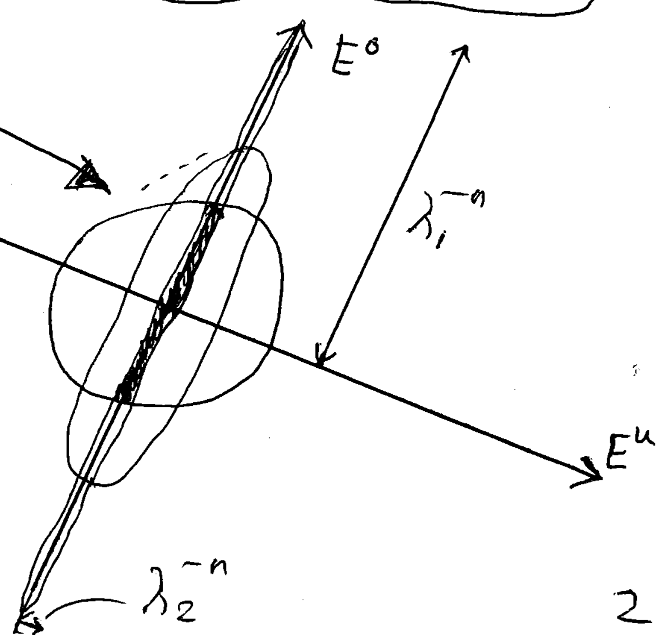
Then $h_\mu(g) \geq \int_M h_\nu(g,\rho,x)^+ \, d\mu(x)$.

Ex: Arnold's cat map, $f: \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$ on $\mathbb{T}^2$.

Take ρ = a small constant, and $\mu = \nu = \text{Leb}$.

$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$; eigenspaces $\mathbb{R}^2 = E^0 \oplus E^u$
eigenvalues $\lambda_1 = \frac{3-\sqrt{5}}{2}$, $\lambda_2 = \frac{3+\sqrt{5}}{2}$

$S_n(f,\rho,x) =$



$\therefore \nu(S_n(f,\rho,x)) \ll \lambda_2^{-n}$

$\therefore h_\mu(f) \geq \log \lambda_2$

proof of Prop: First use $\log\rho \in L^1_\mu$ to see that there is a (m'ble, countable) partition α of M with $H_\mu(\alpha) < \infty$ and diam $\alpha(x) \le \rho(x)$, μ-a.e. $x \in M$

{The α-set which contains x}

Construction: For every small $r>0$, $\exists$ partition $\alpha(r)$ with $\operatorname{diam} A \le r$, $\forall A \in \alpha(r)$ and $\#\alpha(r) \ll r^{-d}$ ($d = \dim M$)

For $n \geq 0$ set $U_n = \{x \in M : e^{-n-1} < \rho(x) \le e^{-n}\}$;

then $M = \coprod_{n=0}^{\infty} U_n$ and ⊛ $\sum_{1}^{\infty} n \cdot \mu(U_n) < \infty$. ← {by using $\log\rho \in L^1$}

Partition each U_n as $U_n = \coprod_{Q \in \alpha(e^{-n-1})} (Q \cap U_n)$.

Let α = the union of these partitions!

Note: $x \in Q \cap U_n \Rightarrow \operatorname{diam}(Q \cap U_n) \le e^{-n-1} < \rho(x)$.

Also: $H_\mu(\alpha) = \sum_{n=0}^{\infty} \sum_{Q \in \alpha(e^{-n-1})} \varphi(\mu(Q \cap U_n))$ $\quad \varphi(t) = -t\log t$

$$\overset{\text{Jensen}}{\le} \sum_{n=0}^{\infty} \mu(U_n)\Big(\underbrace{\log \#\alpha(e^{-n-1})}_{\le O(1) + d(n+1)} - \log \mu(U_n)\Big) \ll \infty.$$

{using ⊛ above}

For this α, use {recall $\alpha_0^n := \bigvee_{j=0}^{n} g^{-j}(\alpha)$}

$$h_\mu(g) \geq h_\mu(g,\alpha) = \int_M \Big(\lim_{n\to\infty} \frac{1}{n} I_\mu(\alpha_0^n)(x)\Big) d\mu(x)$$

{Shannon – McMillan – Breiman Thm see Problem 46}

Hence suffices to prove

$$\lim_{n\to\infty} \frac{1}{n} I_\mu(\alpha_0^n)(x) \geq h_\nu(g, \rho, x) \quad \text{for } \mu\text{-a.e. } x,$$

Note $\underline{\alpha_0^n(x) \subset S_n(g,\rho,x)}$; hence

$$\underline{h_\nu(g,\rho,x) \leq \limsup_{n\to\infty} -\frac{1}{n}\log \nu(\alpha_0^n(x))}, \qquad \text{(A)}$$

How compare?

whereas $$\underline{\lim_{n\to\infty} \frac{1}{n} I_\mu(\alpha_0^n)(x) = \lim_{n\to\infty} -\frac{1}{n}\log \mu(\alpha_0^n(x))}. \qquad \text{(B)}$$

Let $k: M \to [0,\infty)$ be a Radon-Nikodym density of $\mu_{|\alpha_0^\infty}$ w.r.t. $\nu_{|\alpha_0^\infty}$: thus $\underline{k \text{ is } \alpha_0^\infty\text{-m'ble}}$ and

$$\underline{\mu(A) = \int_A k\, d\nu, \quad \forall A \in \alpha_0^\infty.}$$

Then $$\underline{\underline{\lim_{n\to\infty} \frac{\mu(\alpha_0^n(x))}{\nu(\alpha_0^n(x))} = k(x) \qquad \nu\text{-a.e.}}} \qquad \circledast$$

This implies (A) $\leq$ (B) ν-a.e., and

In fact (A) = (B) if $\circledast$ holds and $k(x) > 0$.

<u>the Proposition is proved!</u> □ □ □

<u>proof of $\circledast$</u>: Rescaling $\nu \Rightarrow$ May assume $\underline{\nu \in P(M)}$.

Then $$\frac{\mu(\alpha_0^n(x))}{\nu(\alpha_0^n(x))} = \frac{1}{\nu(\alpha_0^n(x))}\int_{\alpha_0^n(x)} k\, d\nu = \underbrace{\mathbb{E}(k \mid \alpha_0^n)}_{\substack{\text{conditional expectancy with}\\ \text{underlying prob. measure } \nu}}(x)$$

for ν-a.e. x,

and by the Martingale Convergence Theorem, this tends to $\mathbb{E}(k \mid \alpha_0^\infty)(x) = k(x)$ for ν-a.e. x!

Back to $f: M \to M$ a $C^{1+\varepsilon}$-map

Let M' be the "good" subset for Oseledet's Thm (in particular $\mu(M')=1$, $f^{-1}(M')=M'$). For $x \in M'$, set

$$E^u(x) = \bigoplus_{\chi_j(x)>0} H_x^j, \qquad E^0(x) = \bigoplus_{\chi_j(x)\leq 0} H_x^j$$

Then $d f_x(E^u(x)) = E^u(f(x))$, $d f_x(E_0(x)) = E_0(f(x))$ $\forall x \in M'$.

May reduce to $\dim E^u(x) = \text{constant} =: d_u > 0$, $\forall x \in M'$.

{using "Entropy Affine"; Probl. 50}

{if $d_u = 0$ then $\chi \equiv 0 \Rightarrow$ nothing to prove}

Given $\varepsilon > 0$, by Lusin's Theorem there is a compact set $K \subset M'$ with $\mu(K) \geq 1-\varepsilon$ s.t. $E^u(x)$, $E^0(x)$ and $\chi^+(x) := \min\{\chi_j(x) : \chi_j(x) > 0\}$ depend continuously on x for $x \in K$. We may also assume that $\exists\ \lambda > \beta > 1$, $N \in \mathbb{Z}^+$ s.t.

$$\forall x \in K,\ n \geq N : \begin{cases} \left\| (d_x f^n)_{|E^0(x)} \right\| \leq \beta^n \\ \left\| (d_x f^n)(\underline{v}) \right\| \geq \lambda^n \|\underline{v}\| \quad (\forall \underline{v} \in E^u(x)) \\ \log \underbrace{\left| \det (d_x f^n)_{|E^u(x)} \right|}_{\text{expansion factor of } d_u\text{-dim volume}} \geq n(\chi(x) - \varepsilon) \end{cases}$$

(Some steps in the proof: Oseledet's Theorem implies that $\forall x \in M'$: $\limsup_{n\to\infty} \frac{1}{n} \log \|(d_x f^n)_{E^0(x)}\| \leq 0$ (easy!) and $\lim_{n\to\infty} \frac{1}{n} \log \|((d_x f^n)_{|E^u(x)})^{-1}\| = -\chi^+(x)$ (somewhat trickier!), and

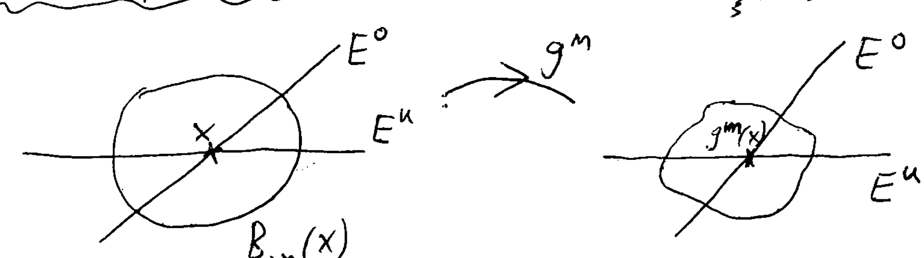
$$\lim_{n\to\infty} \frac{1}{n} \log \left|\det (d_x f^n)_{|E^u(x)}\right| = \chi(x).$$

Once we know this, the rest is basic measure theory.)

Set $\boxed{g := f^N}$, $\bar{\beta} = \beta^N$, $\bar{\lambda} = \lambda^N$. Then

$$\forall x \in K, n \geq 0: \begin{cases} \|(d_x g^n)_{|E^0(x)}\| \leq \bar{\beta}^n & (I) \\ \|(d_x g^n)(\underline{v})\| \geq \bar{\lambda}^n \|\underline{v}\| \quad \forall \underline{v} \in E^u(x) & (II) \\ \log \left|\det (d_x g^n)_{|E^u(x)}\right| \geq n N(\chi(x) - \varepsilon) & (III) \end{cases}$$

Now fix $\xi > 0$ so small that for any $x \in K$ and $m \in \mathbb{Z}^+$ s.t. $g^m(x) \in K$, we have "good control on dispersion" in the ball $B_{\xi^m}(x)$ under g^m!



cf. Mañé's Lemma 4 — crucially uses (I), (II), and $g \in C^{1+\varepsilon}$.

Define $N_K : M \longrightarrow \mathbb{Z}_{\geq 0} \cup \{\infty\}$ by

$$N_K(x) = \begin{cases} \min\{n \geq 1 : g^n(x) \in K\} & \text{if } x \in K \\ 0 & \text{if } x \notin K \end{cases}$$

Then $\int N_K \, d\mu < \infty$. Define $\rho : M \longrightarrow (0,1)$ by

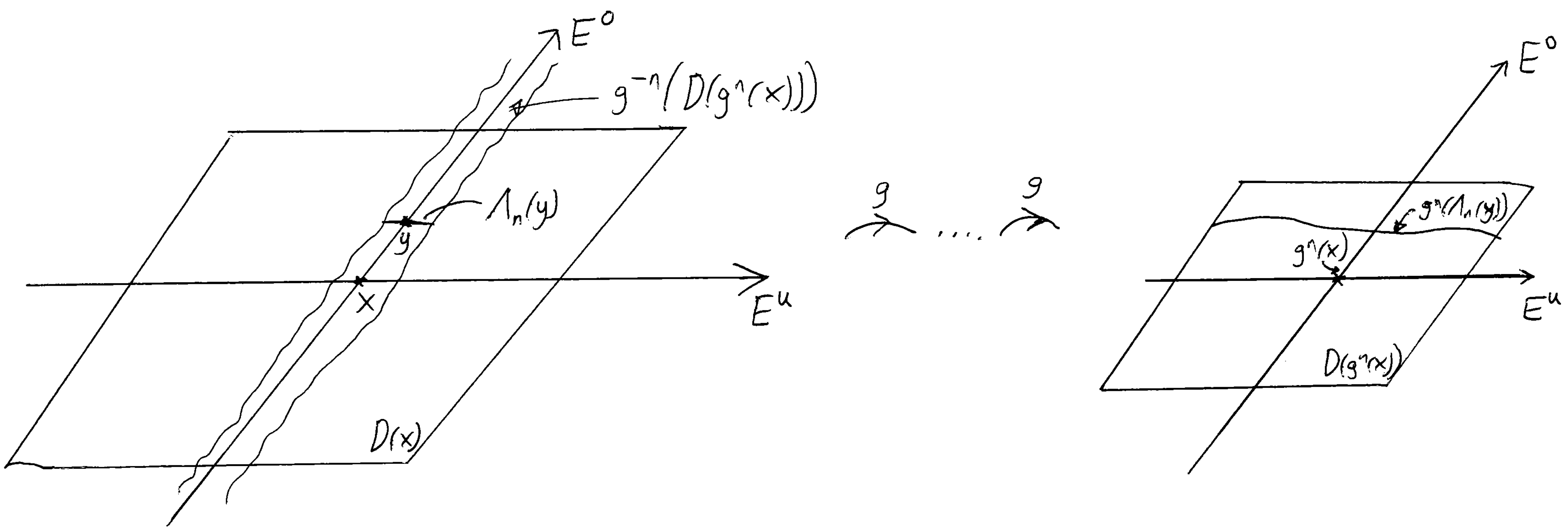
$$\rho(x) = a \cdot \min\left(1, \xi^{N_K(x)}\right)$$

{a suitable, small constant > 0}

Then $\log \rho \in L^1_\mu$!

$D_r(x) = \{x+v+w : v \in E^u(x),\ w \in E^0(x),\ \|v\|, \|w\| < r\}$

$D(x) = D_{\rho(x)/k_1}(x)$ $\quad$ k_1 chosen so that $B_r(x) \subset D_{r/k_1}(x),\ \forall x \in K$



By our choice of ξ, ρ: if $g^n(x) \in K$ then $g^n(\Lambda_n(y))$ is a "$(E^0(g^n(x)), E^u(g^n(x)))$-graph" of small dispersion $\Rightarrow$ $vol(g^n(\Lambda_n(y))) \ll 1$ (bound independent of x, n)

d_u-dim "curved" Lebesgue volume

by (III)

But for each j with $g^j(x) \in K$, the application of g multiplies "that" volume by $\gtrsim e^{N(\chi(x)-\varepsilon)}$

PET: Can arrange this holds for $> 1-\sqrt{\varepsilon}$ of all j's, for $1-\sqrt{\varepsilon}$ of all $x \in K$.

Hence, for every $y \in E^{o}(x)$ with $x + y \in D(x)$:

$$\nu^u(\Lambda_n(y)) \lesssim e^{-nN(\chi(x)-\varepsilon)}$$

d_u-dim Leb. volume in $E^u(x)$

$$\Rightarrow \quad h_\nu(f^N, \rho, x) \geq N(\chi(x) - \varepsilon)$$

for "$1-\sqrt{\varepsilon}$" of all $x \in K$

$$\overset{\text{Prop}}{\Longrightarrow} \quad h_\mu(f) = \frac{1}{N} h_\mu(f^N) \geq \int_M (\chi(x) - \varepsilon)\, d\mu(x).$$

Done! $\square\ \square$

12.1. **Notes.** .

We follow the proof in Mañé [28]. See also my notes to that paper.

On p. 2, the example with Arnold's cat map, note that since $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ is symmetric, the eigenspaces E^0 and E^u are orthogonal and hence $f^{-n}(B(f^n(x)))$ is an ellipse with semi-axes *exactly* equal to $\rho\lambda_1^{-n}$ and $\rho\lambda_2^{-2}$. However the picture is *qualitatively* correct for matrices A with eigenvalues $|\lambda_1| < 1 < \lambda_2$ also if E^0 and E^u not orthogonal; for large n the ellipse $f^{-n}(B(f^n(x)))$ is very long and thin; one semi-axis has length $\asymp |\lambda_1|^{-n}$ and is very nearly parallel with E^0, and the other semi-axis (which is orthogonal to the first) has length $\asymp \lambda_2^{-n}$.

It is also important to note that, for large n, the long thin ellipse $f^{-n}(B(f^n(x)))$ wraps itself around the torus many times, and will certainly (since the cat map f is mixing) visit the ρ-ball $B(x)$ *many* times. Hence to conclude that $S_n(f, \rho, x)$ equals the colored set which I've drawn on p. 2, it is important to use that $S_n(f, \rho, x)$ equals the full intersection $\cap_{i=0}^n f^{-i}(B(f^i(x)))$, and not just "$B(x) \cap f^{-n}(B(f^n(x)))$"; the latter set is much larger and consists of many disconnected parts inside $B(x)$. Also it is important that we *fix ρ sufficiently small.* Indeed, if e.g. $\rho > \sqrt{1/2}$ then $B(x) = \mathbb{T}^2$ for all x and so $S_n(f, \rho, x) = \mathbb{T}^2$ for all x!

In the general setting of Mañé's paper, the above "non-wrapping property" is contained in the statement of [28, Lemma 5] (which in turn makes crucial use of [28, Lemma 4]). Indeed, by definition $g^n(\Lambda_n(y)) \subset D_{\rho(g^n(x))/k_1}(g^n(x))$, i.e. every point in $g^n(\Lambda_n(y))$ can be uniquely expressed as $g^n(x) + y_1 + y_2$ with $y_1 \in E^0(g^n(x))$ and $y_2 \in E^u(g^n(x))$, $\|y_1\|, \|y_2\| < \rho(g^n(x))/k_1$; and now [28, Lemma 5] says that $g^n(\Lambda_n(y))$ is an $(E^0(g^n(x)), E^u(g^n(x)))$*-graph,* which in particular means that for every $y_2 \in E^u(g^n(x))$ there is *at most one* $y_1 \in E^0(g^n(x))$ with $g^n(x) + y_1 + y_2 \in g^n(\Lambda_n(y))$.

Coming back to the case of the torus, it seems that in the case $M = \mathbb{T}^d$ (provided with *any* Riemannian metric, and also for an arbitrary map $f : \mathbb{T}^d \to \mathbb{T}^d$ subject only to the assumptions which Mañé makes on his first page), one can fairly easily follow all of Mañé's proof *directly using the "$\mathbb{R}^d$ mod $\mathbb{Z}^d$" coordinates* on $\mathbb{T}^d$, i.e. without first making a fixed choice of a finite number of coordinate neighborhoods covering $\mathbb{T}^d$. [5] When doing this, there are only a few points in Mañé's proof that require extra considerations. Perhaps the main such point concerns the notion of an (E_1, E_2)*-graph;* this notion was defined on [28, p. 98] when $E = E_1 \oplus E_2$ is a normed linear space but now it seems appropriate to also make a definition of the following kind: for any $x \in \mathbb{T}^d$ and any two linear subspaces $E_1, E_2 \subset \mathbb{R}^d$ satisfying

[5] Actually what we do could be seen as: Around any given $x \in \mathbb{T}^d$, use the coordinate chart $\varphi_x : U_x \to (-\frac{1}{2}, \frac{1}{2})^d \subset \mathbb{R}^d$ where $U_x = x + (-\frac{1}{2}, \frac{1}{2})^d \subset \mathbb{T}^d$ (natural notation), and φ_x is the inverse of the map $y \mapsto x + y$, $(-\frac{1}{2}, \frac{1}{2})^d \to U_x$.

$T_x(\mathbb{T}^d) = \mathbb{R}^d = E_1 \oplus E_2$, a subset $G \subset \mathbb{T}^d$ is called an (E_1, E_2)-*graph* if, setting

$$c_d := \frac{1}{2}\sqrt{\frac{1}{d}},$$

there is an open subset $U \subset E_2 \cap B_{c_d}(0)$ and a C^1-map $\Psi : U \to E_1 \cap B_{c_d}(0)$ such that $G = \{x + \Psi(y_2) + y_2 \ : \ y_2 \in U\}$. [6] Now when proving [28, Lemma 4] there is a little extra discussion needed, to choose ξ sufficiently small so that the "$B_{c_d}(0)$"-containment required in the above definition is guaranteed to hold.

[6]Note that our choice of c_d guarantees that the whole set $(E_1 \cap B_{c_d}(0)) + (E_2 \cap B_{c_d}(0)) \subset \mathbb{R}^d$ is injectively embedded in the torus, i.e. the map $\langle y_1, y_2\rangle \mapsto x + y_1 + y_2$ from $(E_1 \cap B_{c_d}(0)) \times (E_2 \cap B_{c_d}(0))$ to $\mathbb{T}^d$ is injective.

13. IETs; Reuzy-Veech renormalization; Teichmüller flow I

~~Another example: The Rauzy-Veech (-Zorich) cocycle over the space of IETs~~

Lecture #12: IETs, Rauzy-Veech renormalization, Teichmüller flow

IET, precise def.: Let

$\mathcal{A}$ - a finite alphabet, $d = \#\mathcal{A} \geq 2$.

$\lambda = (\lambda_\alpha)_{\alpha \in \mathcal{A}} \in \mathbb{R}_+^{\mathcal{A}}$ $\qquad (\mathbb{R}_+ = (0,\infty))$

$\Sigma_{\mathcal{A}} = \{\pi = (\pi_1, \pi_2) \;:\; \pi_0, \pi_1 \text{ bijections } \mathcal{A} \to \{1, \dots, d\}\}$

For $(\pi, \lambda) \in \Sigma_{\mathcal{A}} \times \mathbb{R}_+^{\mathcal{A}}$, the IET $\boxed{f_{\pi,\lambda} : I^{\pi,\lambda} \to I^{\pi,\lambda}}$

is defined by

$$\underline{I = I^{\pi,\lambda} = [0, |\lambda|)}, \qquad \underline{|\lambda| := \sum_{\alpha \in \mathcal{A}} \lambda_\alpha},$$

$$\underline{I_\alpha = I_\alpha^{\pi,\lambda} = \Big[\sum_{\beta : \pi_0(\beta) < \pi_0(\alpha)} \lambda_\beta \,,\, \sum_{\beta : \pi_0(\beta) \leq \pi_0(\alpha)} \lambda_\beta\Big)}$$

(note $I = \bigsqcup_{\alpha \in \mathcal{A}} I_\alpha$)

$f = f_{\pi,\lambda} : I \to I$ def. by $\boxed{f(x) = x + w_\alpha, \;\forall\, x \in I_\alpha}$

where $\underline{w = (w_\alpha) = \Omega_\pi(\lambda)}$ def. by

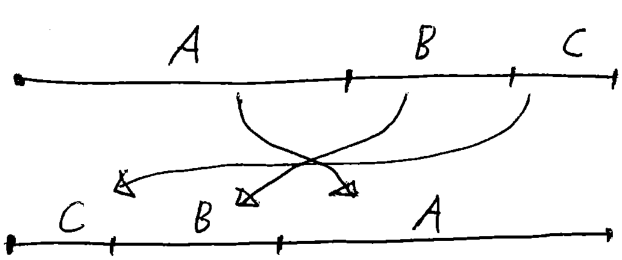
{ w = the translation vector of f. }

$$\underline{w_\alpha := \sum_{\beta : \pi_1(\beta) < \pi_1(\alpha)} \lambda_\beta - \sum_{\beta : \pi_0(\beta) < \pi_0(\alpha)} \lambda_\beta .}$$

Thus Ω_π is a $d \times d$-matrix with entries $0, \pm 1$.

{ It is anti-symmetric; hence it gives rise to a symplectic form on a certain quotient of $\mathbb{R}^{\mathcal{A}}$; cf. Sarig Sec. 10 }

Ex:



$\mathcal{A} = \{A, B, C\}$

$$\pi = \text{"}\begin{pmatrix} A & B & C \\ C & B & A \end{pmatrix}\text{"} = \begin{pmatrix} \pi_0^{-1}(1), \pi_0^{-1}(2), \dots \\ \pi_1^{-1}(1), \pi_1^{-1}(2), \pi_1^{-1}(3) \end{pmatrix}$$

$$\begin{cases} w_A = \lambda_C + \lambda_B \\ w_B = \lambda_C - \lambda_A \\ w_C = -\lambda_A - \lambda_B \end{cases}$$

$$\Omega_\pi = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix} \end{array}$$

~~We'll now def. the Rauzy-Veech induction and the R-V (-Zorich) cocycle; these are very important in the connection IETs $\longleftrightarrow$ translation surfaces and Teichmüller flow, and also for the finer properties of the IET $f_{\pi,\lambda}$ itself: For example, recall that for Lebesgue a.e. $\lambda \in \mathbb{R}_+^{\mathcal{A}}$ (and π irreducible), f is uniquely ergodic (Masur, Veech 1982); hence the orbit of every point $x \in I$ becomes equidistributed (w.r.t. Leb); in particular the counting function $N_\alpha(x,n) = \#\{0 \le k < n : f^k(x) \in I_\alpha\}$ is $\sim \vec{\lambda}_\alpha \cdot n$ (as $n \to \infty$); and Zorich (1997) has proved precise power scaling for this asymptotics, where the precise power is determined by the Lyapunov exponents of the R-V-Z cocycle!~~

Rauzy - Veech induction

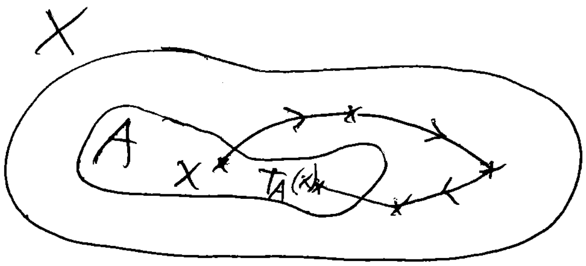
General: For <u>any</u> ppt $(X, \mathcal{B}, \mu, T)$ and any $A \in \mathcal{B}$ with $\mu(A) > 0$, define the <u>induced transformation on A</u> (or <u>the first return map to A</u>) as $(A, \mathcal{B}, \mu_A, T_A)$

where
$$\begin{cases} \varphi_A(x) := \min\{n \geq 1 : T^n(x) \in A\} \\ A_0 := \{x \in A : \varphi_A(x) < \infty\} \\ (\text{then } \mu(A \setminus A_0) = 0 \text{ by } \underline{\text{Poincaré Recurrence}}) \\ \underline{T_A : A_0 \to A}; \quad T_A(x) = T^{\varphi_A(x)}(x). \\ \mathcal{B}_A = \{A \cap E : E \in \mathcal{B}\} = \{E \in \mathcal{B} : E \subset A\} \\ \mu_A(E) = \dfrac{\mu(E)}{\mu(A)} \quad (E \in \mathcal{B}_A) \end{cases}$$

See Sarig 1.6.4

ergodic $\Rightarrow$ ergodic

mixing $\not\Rightarrow$ mixing



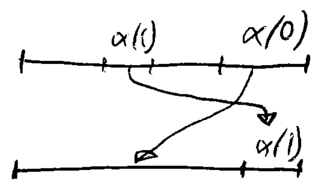
For IET $f = f_{\pi,\lambda} : I \to I$

see Problem 29

The first return map to <u>any</u> subinterval is again an IET.

The R-V is a particularly nice choice!

Set $\underline{\alpha(j)} = \pi_j^{-1}(d) \in \mathcal{A}$ $(j = 0, 1)$



$\underline{\varepsilon = \varepsilon(\pi, \lambda)} =$ "<u>the type of (π, λ)</u>"

$$= \begin{cases} 0 \text{ if } \lambda_{\alpha(0)} > \lambda_{\alpha(1)} \\ 1 \text{ if } \lambda_{\alpha(1)} > \lambda_{\alpha(0)} \end{cases}, \quad (\underline{\text{undef}} \text{ if } \lambda_{\alpha(0)} = \lambda_{\alpha(1)})$$

(We call $I_{\alpha(\varepsilon)}$ <u>the winner</u>, $I_{\alpha(1-\varepsilon)}$ <u>the loser</u>.)

Let $J = \begin{cases} I \setminus I_{\alpha(0)} & \text{if } \varepsilon = 1 \\ I \setminus f(I_{\alpha(1)}) & \text{if } \varepsilon = 0 \end{cases}$

Now $\underline{\underline{[\text{the R-V induction of } f]}}$

$:=$ <u>the first return map of f to J</u> !

$= \underline{\underline{f_{\hat{R}(\pi,\lambda)}}}$,

where $\hat{R} : \{ \Sigma_{\mathcal{A}} \times \mathbb{R}_+^{\mathcal{A}} : \lambda_{\alpha(0)} \neq \lambda_{\alpha(1)} \} \longrightarrow \Sigma_{\mathcal{A}} \times \mathbb{R}_+^{\mathcal{A}}$

$\hat{R}(\pi, \lambda) = (\pi', \lambda')$

$\lambda' = \Theta_{\pi,\lambda}^{*-1}(\lambda)$

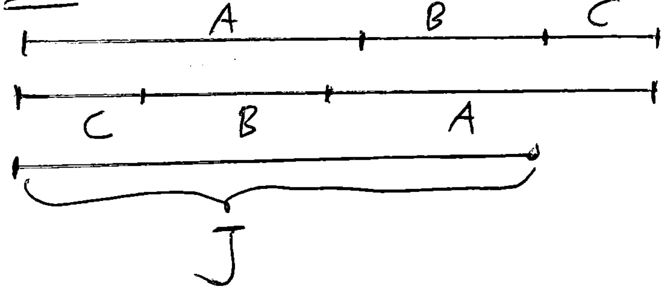
$\boxed{\Theta_{\pi,\lambda} := I + E_{\alpha(1-\varepsilon),\alpha(\varepsilon)}}$

The $\mathcal{A} \times \mathcal{A}$-matrix with all entries $= 0$ except a "1" in position $\alpha(1-\varepsilon), \alpha(\varepsilon)$.

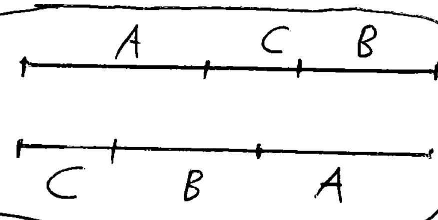
(Also $w' = \Omega_{\pi'}(\lambda') = \Theta_{\pi,\lambda}(w)$!)

$\pi' =$ See Viana p. 16 !

<u>Ex</u>: The IET on p. 10, assuming $\lambda_A > \lambda_C$: $\underline{\underline{\varepsilon = 1}}$



$f_{\hat{R}(\pi,\lambda)}$:



$\pi' = \begin{pmatrix} A & C & B \\ C & B & A \end{pmatrix}$

(cf. Sang's p. 16 (6) with $\alpha(1) = A$, $\alpha(0) = C$, $k =$ [redacted]

$$\begin{cases} \lambda_A' = \lambda_A - \lambda_C \\ \lambda_B' = \lambda_B \\ \lambda_C' = \lambda_C \end{cases}$$

thus $$\underline{\underline{\theta^{*\,-1}_{\pi,\lambda}}} = \begin{array}{c|ccc} & A & B & C \\ \hline A & 1 & 0 & -1 \\ B & 0 & 1 & 0 \\ C & 0 & 0 & 1 \end{array},$$

$$\underline{\underline{\theta_{\pi,\lambda}}} = \begin{array}{c|ccc} & A & B & C \\ \hline A & 1 & 0 & 0 \\ B & 0 & 1 & 0 \\ C & 1 & 0 & 1 \end{array} = \underline{\underline{I + E_{C,A}}}$$

(ok!)

Dynamical system on the space of IET's?

First mod out $\lambda \longleftrightarrow t\cdot\lambda$ $(t \in \mathbb{R}_+)$:

Set $\underline{\Lambda_{\mathcal{A}} = \{\lambda \in \mathbb{R}_+^{\mathcal{A}} : |\lambda| = 1\}}$ ← Concrete version of projective space "$\mathbb{P}\mathbb{R}^{\mathcal{A}}_{+}$"

Let $\underline{\underline{R}} =$ "$\mathbb{P}\hat{R}$", i.e.

$$R : \{\Sigma_{\mathcal{A}} \times \Lambda_{\mathcal{A}} : \lambda_{\alpha(0)} \neq \lambda_{\alpha(1)}\} \longrightarrow \Sigma_{\mathcal{A}} \times \Lambda_{\mathcal{A}},$$

$$R(\pi,\lambda) = \underline{\underline{(\pi',\lambda'')}} \quad \text{where} \begin{cases} (\pi',\lambda') = \hat{R}(\pi,\lambda), \\ \lambda'' = |\lambda'|^{-1}\lambda' \end{cases}$$

Now set

$$X = \{(\pi,\lambda) \in \Sigma_{\mathcal{A}} \times \Lambda_{\mathcal{A}} : R^n(\pi,\lambda) \text{ def. } \forall n \geq 1\}$$

$$= \{(\pi,\lambda) \text{ satisfies the } \underline{\underline{\text{Keane condition}}}\}$$

Viana p.19T & Cor 5.4

Say! Viana Sec 3

Dynamical system on $\Sigma_{\mathcal{A}} \times \mathbb{R}_+^{\mathcal{A}}$?

"the space of IETs"

Set $X = \{(\pi, \lambda) \in \Sigma_{\mathcal{A}} \times \mathbb{R}_+^{\mathcal{A}} : R^n(\pi, \lambda) \text{ def. } \forall n \geq 1\}$

$= \{(\pi, \lambda)$ satisfies the <u>Keane condition</u> $\}$

Viana p. 19T & Cor 5.4

Viana <u>Sec 3</u>: <u>Say</u>!

<u>Note</u>: – X has <u>full measure</u> in $\Sigma_{\mathcal{A}} \times \mathbb{R}_+^{\mathcal{A}}$ wrt "Lebesgue"

– $f_{\pi,\lambda}$ is <u>minimal</u> for every $(\pi, \lambda) \in X$. Sec. 4

(Key result, Masur & Veech): $f_{\pi,\lambda}$ <u>uniquely ergodic</u> for a.a. (π, λ)!)

– If $(\pi, \lambda) \in X$ then π is <u>irreducible</u>, i.e.

$\underbrace{\pi_1 \circ \pi_0^{-1}}(\{1, .., k\}) \neq \{1, .., k\}, \quad \forall k \in \{1, 2, .., d-1\}$.

the <u>monodromy</u> <u>invariant</u>

$$\boxed{\underline{\text{Key dynamical system}}: \quad \hat{R} : X \to X}$$

<u>sloppily</u>: "$\hat{R} : \Sigma_{\mathcal{A}} \times \mathbb{R}_+^{\mathcal{A}} \to \Sigma_{\mathcal{A}} \times \mathbb{R}_+^{\mathcal{A}}$"

Sec. 6

"π' successor of π"

<u>Graph</u> on $\Sigma_{\mathcal{A}}$: $\quad \pi \to \pi'$ if $\exists \lambda', \lambda \in \mathbb{R}_+^{\mathcal{A}}$ s.t. $\hat{R}(\pi, \lambda) = (\pi', \lambda')$.

<u>If $d \geq 3$</u>: $\forall \pi \in \Sigma_{\mathcal{A}}$ (irreducible): indegree(π) = outdegree(π) = 2.

<u>Rauzy class</u> := $\boxed{\text{a } \underline{\text{connected component}} \text{ of this graph}} \subset \Sigma_{\mathcal{A}}$

weak $\Leftrightarrow$ strong; Lem 6.1

Next: Extend the space! {To a space of translation surfaces!}

Given $\pi \in \Sigma_{\mathcal{A}}$ (irreducible), set

$$T_\pi^+ = \left\{ \tau = (\tau_\alpha) \in \mathbb{R}^{\mathcal{A}} : \sum_{\substack{\alpha \\ (\pi_0(\alpha) \le k)}} \tau_\alpha > 0, \ \sum_{\substack{\alpha \\ \pi_1(\alpha) \le k}} \tau_\alpha < 0 \quad \forall k \in \{1, \dots, d-1\} \right\}$$

{an open convex cone}

Viana Sec. 12

For each $\lambda \in \mathbb{R}_{\mathcal{A}}^+$, $\tau \in T_\pi^+$, define a translation surface $M = M(\pi, \lambda, \tau)$

{recall!}

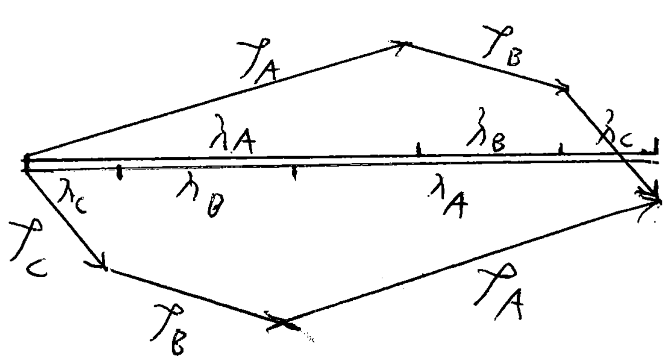
as follows: $\zeta_\alpha := (\lambda_\alpha, \tau_\alpha) \in \mathbb{R}^2$

$$\Gamma = \Gamma(\pi, \lambda, \tau) = \text{polygon } \zeta_{\pi_0^{-1}(1)}, \dots, \zeta_{\pi_0^{-1}(d)}, -\zeta_{\pi_1^{-1}(d)}, \dots, -\zeta_{\pi_1^{-1}(1)}$$

then identify parallel sides in pairs!

For earlier example: $\mathcal{A} = \{A, B, C\}$, $\pi = \begin{pmatrix} A & B & C \\ C & B & A \end{pmatrix}$.

$\tau \in T_\pi^+ \iff \tau_A > 0, \ \tau_A + \tau_B > 0, \ \tau_C < 0, \ \tau_B + \tau_C < 0$

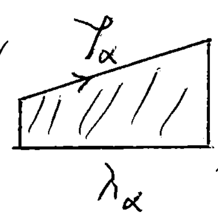


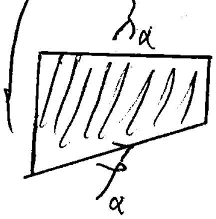
Note: — $I = [0, |\lambda|) \times \{0\}$ is a cross-section for the vertical flow on $M(\pi, \lambda, \tau)$; the corresponding first return map is $f_{\pi,\lambda}$!

— define $\text{type}(\tau) = \begin{cases} 0 & \text{if } \sum_{\mathcal{A}} \tau_\alpha > 0 \\ 1 & \text{if } \sum_{\mathcal{A}} \tau_\alpha < 0 \end{cases}$.

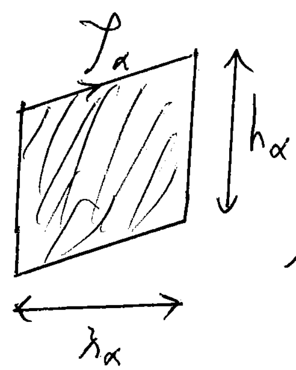
— If $\text{type}(\tau) \neq \text{type}(\pi, \lambda)$ then Γ may intersect itself; need to handle appropriately in the def. of M!

Important formula: $\boxed{\text{Area } M(\pi, \lambda, \tau) = -\lambda \cdot \Omega_\pi(\tau) \qquad \text{def: } h := -\Omega_\pi(\tau)}$

"proof": Area(

) + Area(

) =

= Area(

) $= h_\alpha \zeta_\alpha$

13.1. **Notes.** .

The lecture goes through certain material from Viana, [49, Sec. 1—12].

pp. 1-2, notation; cf. [49, Sec. 1].

p. 3, induced transformation for a general ppt: This is in Sarig, [40, Sec. 1.6.4]. Note that T_A is in general not defined on the whole set A but only on the full measure subset A_0. If we want a "genuine" ppt in the sense that it has been defined in [40] then (as is standard) we can simply pass to the set A_1 consisting of all $x \in A_0$ for which $T_A^n(x)$ is defined for all $n \geq 1$, i.e.

$$A_1 = \bigcap_{n=1}^{\infty} T_A^{-n}(A).$$

This set has full measure in A. Thus: Consider the ppt $(A_1, \mathcal{B}_{A_1}, \mu_{A_1}, T_{A|A_1})$. (See also my notes to [40, Sec. 1.6.4].)

pp. 3–5, the Rauzy-Veech induction map; cf. [49, Sec. 2].

p. 6; cf. [49, Sec. 3–6].

p. 7, the definition of the translation surface $M(\pi, \lambda, \tau)$; cf. [49, Sec. 12].

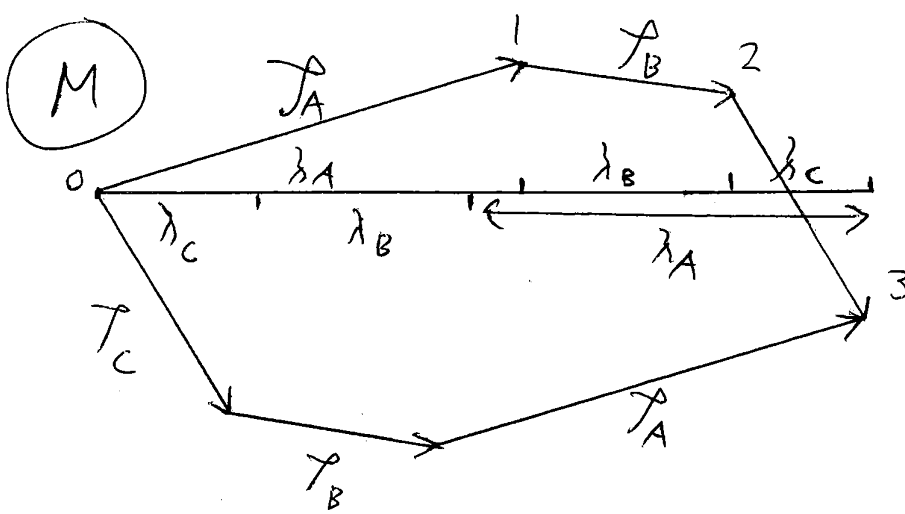
p. 8; the formula for the area of $M(\pi, \lambda, \tau)$ is in [49, p. 54 (48)].

14. IETs; Reuzy-Veech renormalization; Teichmüller flow II

Lecture #14: IETs, Rauzy-Veech, Teichmüller flow

Recall: ~~Given~~ For $\pi \in \Sigma_{\mathcal{A}}$, $\lambda \in \mathbb{R}_+^{\mathcal{A}}$, $\tau \in T_\pi^+$, get a translation surface $\underline{M = M(\pi, \lambda, \tau)}$.

Ex $\mathcal{A} = \{A, B, C\}$, $\pi = \begin{pmatrix} A & B & C \\ C & B & A \end{pmatrix}$



$\left(\zeta_\alpha = (\lambda_\alpha, \tau_\alpha) \in \mathbb{R}^2\right)$

Distinct singular points of M: $1 \sim 3$; angle 2π

$0 \sim 2$; angle 2π

$\therefore\ \kappa = 2,\ m_0 = m_1 = 0$

def: angle $= 2\pi(m_i + 1)$

Viana Sec 14 and 13

g = genus of M

χ = Euler characteristic of M

general formula

$$\boxed{2 - 2g = \chi = \kappa + 1 - d = -\sum_{i=1}^{d} m_i}$$

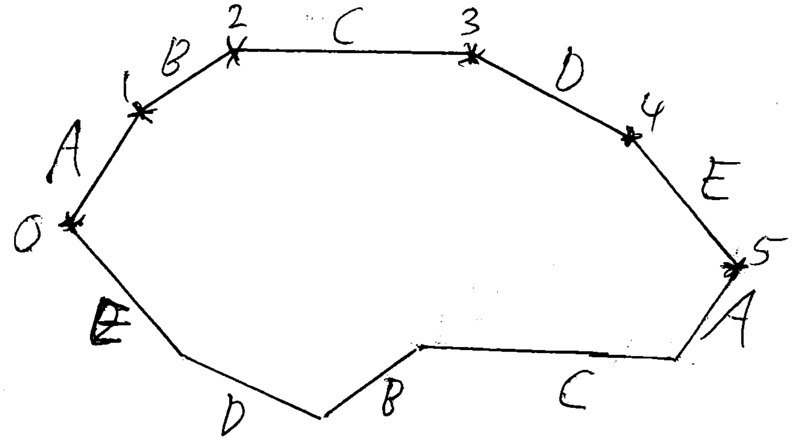
From counting in a triangulation; see Viana p. 41, Fig 15.

From computing total angle in the polygon, in two ways.

$\therefore$ $\underline{g = 1;\ \text{torus}}$

this example is mentioned in Viana p. 44

One more example (Viana p. 46)



$$\pi = \begin{pmatrix} A & B & C & D & E \\ E & D & B & C & A \end{pmatrix}$$

$1 \sim 5 \sim 3 \sim 0 \sim 4$; angle 6π

$2 \sim 2$; angle 2π

$\therefore\ K = 2,\ m_0 = 2,\ m_1 = 0$

$\underline{\underline{g = 2}}$

Computational scheme (Sec 14)

$$p = \pi_1 \circ \pi_0^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 4 & 2 & 1 \end{pmatrix}$$

i.e. $p(1) = 5$, $p(2) = 3$, $p(3) = 4$, $p(4) = 2$, $p(5) = 1$; notation inconsistent with our π-notation!

Set $P = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 0 & 5 & 3 & 4 & 2 & 1 & 6 \end{pmatrix}$,

$$\underline{\underline{\sigma(j) = P^{-1}(P(j)+1) - 1}},$$ (a permutation on $\{0, 1, \dots, d\}$)

$\sigma: 1 \mapsto 5 \mapsto 3 \mapsto 0 \mapsto 4 \mapsto 1, \quad 2 \mapsto 2$!

Get m_i as $\underline{\#(\text{orbit} \cap \{1, 2, \dots, d-1\}) - 1}$!

Alternative model for $M(\pi,\lambda,\tau)$: "zippered rectangles"

Viana Sec. 15

Recall $\underline{h := -\Omega_\pi(\tau) \in \mathbb{R}_+^{\mathcal{A}}}$ (Area $M = \lambda \cdot h$.)

Define
$$R_\alpha^0 = I_\alpha^0 \times [0, h_\alpha]$$
$$R_\alpha^1 = f(I_\alpha^0) \times [-h_\alpha, 0]$$
$$S_\alpha^0 = (\sup I_\alpha) \times \Big[0, \sum_{\beta:\, \pi_0(\beta) \le \pi_0(\alpha)} \tau_\beta\Big]$$
$$S_\alpha^1 = (\sup f(I_\alpha)) \times \Big[\sum_{\beta:\, \pi_1(\beta) \le \pi_1(\alpha)} \tau_\beta,\ 0\Big]$$

Thus $S^0_{\alpha(0)} = S^1_{\alpha(1)}$, but for $\alpha = \alpha(0)$, S_α^0 is a line segment from I to a vertex of $\Gamma(\pi,\lambda,\tau)$, and similarly for $\alpha \neq \alpha(1)$, S_α^1.

good name?

Now set $\underline{\underline{M(\pi,\lambda,\tau,h) := \bigcup_{\alpha\in\mathcal{A}} \bigcup_{\varepsilon=0,1} \left(R_\alpha^\varepsilon \cup S_\alpha^\varepsilon\right)}}$, with

identifications:

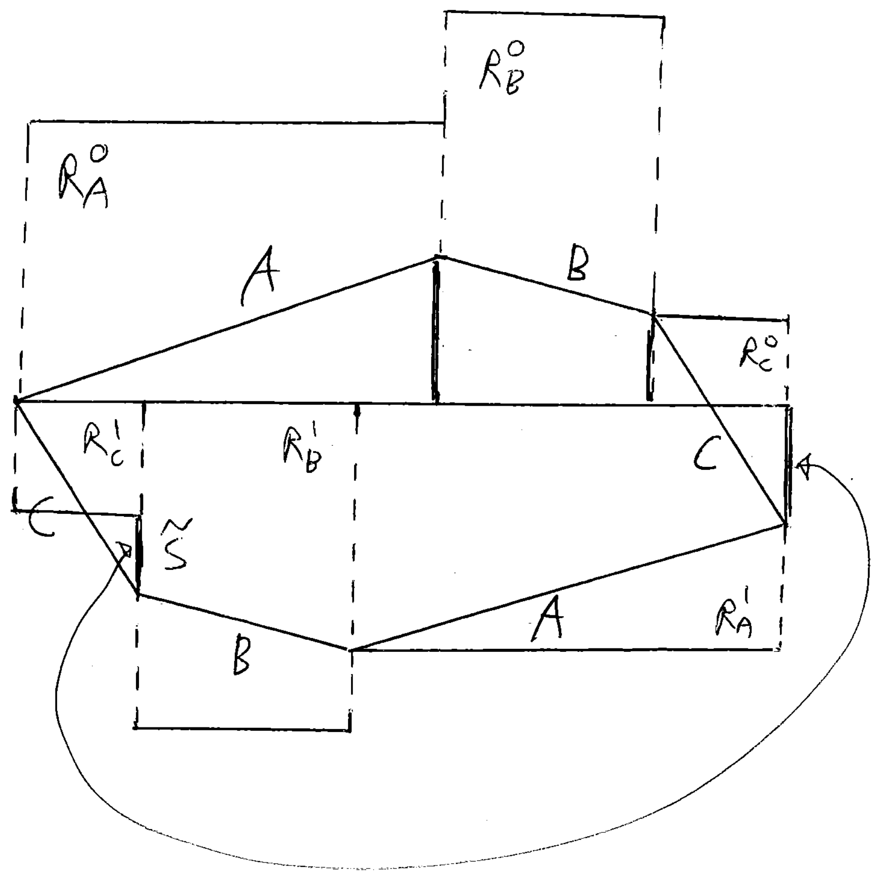
1. $R_\alpha^0 = R_\alpha^1$ via $(x,y) \mapsto (x,y) + (w_\alpha, -h_\alpha)$ $\quad \left\{\begin{array}{l} w = \Omega_\pi(\lambda) \\ h = -\Omega_\pi(\tau) \end{array}\right.$

2. $S^0_{\alpha(0)} = S^1_{\alpha(1)} = \tilde{S}$ $\leftarrow$ $\underline{\text{Def of } \tilde{S}}$: Let $\varepsilon = \text{type}(\tau)$, $\alpha = \alpha(1-\varepsilon)$; then $\underline{S_\alpha^\varepsilon \text{ is longer than } R_\alpha^\varepsilon}$ and $\tilde{S}$ is $\underline{\text{the over-shooting part}}$!

This translation surface $\underline{M(\pi,\lambda,\tau,h)}$ is $\underline{\text{isometric}}$ $\underline{\text{to } M(\pi,\lambda,\tau)}$!

Now $\text{Area}(M) = \lambda \cdot h$ very clear!

Also clear: IET $f_{\pi,\lambda}$ = first return map to I for vertical flow on M, and roof function = "h"



"Invertible Rauzy-Veech induction"

Let $\mathcal{C} \subset \Sigma_{\mathcal{A}}$ a Rauzy class, and set

$$\underline{\underline{\hat{\mathcal{H}} = \hat{\mathcal{H}}(\mathcal{C}) := \{(\pi, \lambda, \tau) : \pi \in \mathcal{C},\ \lambda \in \mathbb{R}^+_{\mathcal{A}},\ \tau \in T^+_\pi\}}}$$

Define

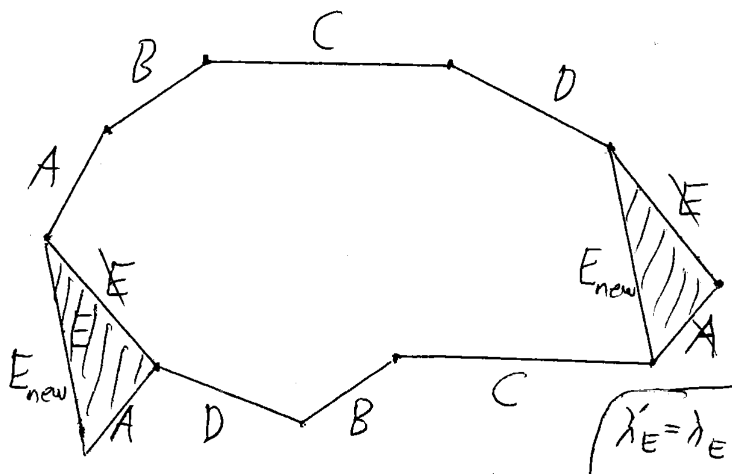
$$\underline{\underline{\hat{\mathcal{R}} : \hat{\mathcal{H}} \cap \{\lambda_{\alpha(0)} \neq \lambda_{\alpha(1)}\} \longrightarrow \hat{\mathcal{H}}}}$$

by $\underline{\hat{\mathcal{R}}(\pi, \lambda, \tau) = (\pi', \lambda', \tau')}$ with $\underline{\underline{\begin{cases} (\pi', \lambda') = \hat{R}(\pi, \lambda), \\ \tau' = \Theta^{*-1}_{\pi,\lambda}(\tau) \end{cases}}}$

{Recall that also $\lambda' = \Theta^{*-1}_{\pi,\lambda}(\lambda)$ }

<u>Geometrically</u>: Cut away the triangle with sides $\tau_{\alpha(0)}, -\tau_{\alpha(1)}$, paste it back on <u>other side labeled by $\alpha(\varepsilon)$</u>, <u>$\varepsilon = \text{type}(\pi, \lambda)$</u>. **Note:** <u>$M(\pi', \lambda', \tau')$ and $M(\pi, \lambda, \tau)$</u> <u>are isometric</u> (by an isometry preserving $\uparrow$, and $\underline{0}$!)

<u>For the example on p. 10:</u>



$\lambda'_E = \lambda_E - \lambda_A$
$\tau'_E = \tau_E - \tau_A$

$\varepsilon = \text{type}(\pi, \lambda) = 0$

Note $\boxed{\text{type}(\tau') = 1 - \varepsilon = 1 - \text{type}(\pi, \lambda)}$ always!

Set $\mathbb{R}^{\mathcal{A}}_{\pi,\varepsilon} := \{\lambda \in \mathbb{R}^{\mathcal{A}}_+ : (\pi,\lambda) \text{ has type } \varepsilon\}$

$T^{\varepsilon}_\pi := \{\tau \in T^+_\pi : \tau \text{ has type } \varepsilon\}$

Then $\hat{R}$ is a bijection from $\{\pi\}\times\mathbb{R}^{\mathcal{A}}_{\pi,\varepsilon}\times T^+_\pi$ onto $\{\pi'\}\times \mathbb{R}^{\mathcal{A}}_+ \times T^{1-\varepsilon}_{\pi'}$, $\forall \pi\in\mathcal{C}$, $\varepsilon\in\{0,1\}$

π' determined by π, ε!

The proof for $\lambda \mapsto \lambda'$ is very simple; just use $\lambda'_{\alpha(\varepsilon)} = \lambda_{\alpha(\varepsilon)}-\lambda_{\alpha(1-\varepsilon)}$, (while $\lambda'_\beta=\lambda_\beta$, $\forall\beta\neq\alpha(\varepsilon)$)! (cf. Sec. 7). For $\tau \mapsto \tau'$ it's also fairly easy; see Lemma 18.1!

$\therefore$ $\hat{R}$ a bijection from $\hat{\mathcal{H}}\cap\{\lambda_{\alpha(0)}\neq\lambda_{\alpha(1)}\}$ onto $\hat{\mathcal{H}}\cap\{\sum_{\mathcal{A}}\tau_\alpha\neq 0\}$

There is a full measure (wrt "Lebesgue") subset $\hat{\mathcal{H}}'\subset\hat{\mathcal{H}}$ s.t. $\hat{R}$ is a bijection $\hat{\mathcal{H}}' \circlearrowleft$!

See my notes to ~~Viana~~ Viana Cor 18.2! Note Viana often later writes $\hat{\mathcal{H}}$ while really meaning $\hat{\mathcal{H}}'$

Four related dynamical systems

Sec. 7; projectivizing $\hat{R}$ – very natural!

Set $\Lambda_{\mathcal{A}} = \{\lambda \in \mathbb{R}_+^{\mathcal{A}} : |\lambda| = 1\}$

$R : C \times \Lambda_{\mathcal{A}} \circlearrowleft$

pedantically: $(C \times \Lambda_{\mathcal{A}}) \cap \{\lambda_{\alpha(0)} \neq \lambda_{\alpha(1)}\}$

$R(\pi, \lambda) =$ " $\hat{R}(\pi, \lambda)$ rescaled"

i.e. multiply [λ new] by $(1 - \lambda_{\alpha(1-\varepsilon)})^{-1}$

For IETs, the map R is of even more fundamental interest than $\hat{R}$.

Key result: $R : C \times \Lambda_{\mathcal{A}} \circlearrowleft$ admits an (infinite!) invariant measure ν, which is $\ll d\pi \times \mathrm{Leb}$. This ν is unique up to scalar mult, and ergodic.

We'll prove in the next lecture. Cf. Viana Thm. 7.2.
This was proved by Masur and Veech (independently) 1982.

Set $\underline{\mathcal{H} = \hat{\mathcal{H}} \cap \{|\lambda| = 1\}}$. Also $\underline{\mathcal{H}_c = \hat{\mathcal{H}} \cap \{|\lambda| = c\}}$, $(c>0)$

Define $\underline{\mathcal{R} : \hat{\mathcal{H}}' \circlearrowleft}$ by

{or $\hat{\mathcal{R}}\left(\mathcal{T}^{t_R(\pi,\lambda)}(\ldots)\right)$}

$$\underline{\mathcal{R}(\pi,\lambda,\tau) = \mathcal{T}^{t_R(\pi,\lambda)}\left(\hat{\mathcal{R}}(\pi,\lambda,\tau)\right)}$$

where $\underline{\mathcal{T}^t(\pi,\lambda,\tau) := (\pi, e^t\lambda, e^{-t}\tau)}$

{$(\mathcal{T}^t)$ is the <u>Teichmüller flow</u> on $\hat{\mathcal{H}}$; cf. Lecture #1!}

{Note: $\mathcal{T}^t$ <u>commutes</u> with $\mathcal{R}$ and $\hat{\mathcal{R}}$}

and $\underline{t_R(\pi,\lambda) = \log\left(\dfrac{|\lambda|}{|\lambda| - \lambda_{\alpha(1-\varepsilon)}}\right)}$

{The <u>Rauzy renormalization time</u>; $t_R : \hat{\mathcal{H}} \to \mathbb{R}_+$, and t_R is invariant under $(\mathcal{T}^t)$}

Note $\mathcal{R}(\mathcal{H}_c) = \mathcal{H}_c$ {bijection of $\mathcal{H}_c' := \mathcal{H}_c \cap \hat{\mathcal{H}}'$ onto itself!}

<u>Summary</u>

$$\begin{tikzcd}
\boxed{\hat{\mathcal{R}} : \mathcal{C}\times\mathbb{R}_+^{\mathcal{A}} \circlearrowleft} & \cdot \arrow[d, "\text{renormalize}"] & \boxed{\hat{\mathcal{R}} : \hat{\mathcal{H}} \circlearrowleft} \\
\boxed{\mathcal{R} : \mathcal{C}\times\Lambda_{\mathcal{A}} \circlearrowleft} \arrow[u, bend left, "\text{2-to-1}"] & \cdot & \boxed{\mathcal{R} : \mathcal{H} \circlearrowleft} \arrow[u, bend right, "\text{bijective}"'] \arrow[ll, "\text{projection } P"']
\end{tikzcd}$$

14.1. **Notes.** .

pp. 1–2: This is a brief survey of some stuff from [49, Sec. 13–14].

pp. 3–4: This is a brief survey of [49, Sec. 15].

pp. 5–6: Here we follow [49, Sec. 18]. Regarding the subset $\hat{\mathcal{H}}' \subset \hat{\mathcal{H}}$ on which $\hat{\mathcal{R}}$ is a bijection, which I mention at the end of p. 6, see my notes to [49, Cor. 18.2].

p. 7: Here we follow [49, Sec. 7].

p. 8: The definitions here are from [49, Sec. 20].

Lecture #15: IETs, Rauzy-Veech, Teichmüller flow

Our main goal today is to prove the following result — which we announced in last lecture.

Theorem 1: $R : \mathcal{C} \times \Lambda_{\mathcal{A}} \circlearrowleft$ admits an (usually infinite!) invariant measure ν, which is $\ll d\pi \times Leb$. This ν is unique up to scalar mult, and ergodic.

Recollection:

$\boxed{\hat{R} : \mathcal{C} \times \mathbb{R}_+^{\mathcal{A}} \circlearrowleft}$ $\qquad$ $\boxed{\hat{\mathcal{R}} : \hat{\mathcal{H}} \circlearrowleft}$

$\boxed{R : \mathcal{C} \times \Lambda_{\mathcal{A}} \circlearrowleft \quad \left(R(\pi,\lambda) = \text{"}\hat{R}(\pi,\lambda) \text{ rescaled"}\right)}$ $\qquad$ $\boxed{\mathcal{R} : \mathcal{H} \circlearrowleft}$

Define $\quad P : \hat{\mathcal{H}} \to \mathcal{C} \times \mathbb{R}_{+}^{\mathcal{A}}, \qquad (\pi, \lambda, \tau) \mapsto (\pi, \lambda)$.

Then $\underline{\hat{R} \circ P = P \circ \hat{\mathcal{R}}}$ and $\underline{R \circ P = P \circ \mathcal{R}}$.

Thus: $\hat{R} : \mathcal{C} \times \mathbb{R}_+^{\mathcal{A}} \circlearrowleft$ is a <u>factor</u> of $\hat{\mathcal{R}} : \hat{\mathcal{H}} \circlearrowleft$, and $R : \mathcal{C} \times \Lambda_{\mathcal{A}} \circlearrowleft$ is a <u>factor</u> of $\mathcal{R} : \mathcal{H} \circlearrowleft$.

Construction of invariant measures

$\hat{m} := d\pi \, d\lambda \, d\tau$ on $\hat{\mathcal{H}}$.

counting

d-dim Lebesgue

recall $\mathcal{T}^t(\pi, \lambda, \tau) = (\pi, e^t\lambda, e^{-t}\tau)$

$\hat{m}$ is invariant under $\mathcal{T}^t$ and under $\hat{\mathcal{R}}$;

since $\lambda_{new} = \Theta^{*-1}(\lambda)$
$\tau_{new} = \Theta^{*-1}(\tau)$,
det $\Theta = 1$

hence under $\mathcal{R}$.

by a computation; see two alternatives in my notes; the crucial fact is that $t_R(\pi, \lambda)$ only depends on the direction of λ!

$m := d\pi \, d_1\lambda \, d\tau$ on $\mathcal{H}$

"Lebsgue" on $\Lambda_{\mathcal{A}}$ when parametrizing $\Lambda_{\mathcal{A}}$ by any family $\{\lambda_\alpha\}_{\mathcal{A}'}$ where $\mathcal{A}' = \mathcal{A} \setminus \{\alpha_0\}$, some $\alpha_0 \in \mathcal{A}$.

m is invariant under $\mathcal{R}$!

Some details: $\mathcal{H}$ is a global cross-section for $(\mathcal{T}^t)$; use this to parametrize $\hat{\mathcal{H}}$ via $\mathcal{H} \times \mathbb{R} \xrightarrow{(bij)} \hat{\mathcal{H}}$,

$(\pi, \tilde{\lambda}, \tau, s) \mapsto \mathcal{T}^s(\pi, \tilde{\lambda}, \tau) = (\pi, e^s\tilde{\lambda}, e^{-s}\tau)$ $\quad \tilde{\lambda} \in \Lambda_{\mathcal{A}}$!

Then $d\hat{m} = dm \, ds$ (since $d\lambda = e^{ds} \cdot d_1\tilde{\lambda} \, ds$, etc.);

also $\mathcal{R}$ and $\mathcal{T}^t$ commute, etc

see my notes to Viana, Lemma 21.1

For $c>0$, let $\hat{m}_c$ = restr. of $\hat{m}$ to $\hat{\mathcal{H}} \cap \{\text{Area } M \le c\}$

m_c = restr. of m to $\mathcal{H} \cap \{\text{Area } M \le c\}$.

Set $\nu = P_*(m_1)$; this is an R-invariant measure on $C \times \Lambda_{\mathcal{A}}$, and $\nu \ll d\pi \times \text{Leb}$.

This fails for "$P_*(m)$"!

{Explicit densities are computed in Ex 21.5 & Sec. 22.}

Also set $\hat{\nu} = P_*(\hat{m}_1)$.

Define: $\boxed{\hat{S} = \hat{S}(C) := \langle \hat{R} \rangle \backslash \hat{\mathcal{H}}}$

{Note: $\hat{S}$ is "almost" a space of isometry classes of translation surfaces!}

Fundamental domain for $\hat{S}$:

$$\{(\pi, \lambda, \tau) \in \hat{\mathcal{H}} \ : \ 0 \le \log|\lambda| < t_R(\pi, \lambda)\} \qquad (*)$$

{Hence we get a concrete model for $\hat{S}$ by taking the ~~the~~ closure of $(*)$ and identifying the boundary part $\log|\lambda| = t_R$ with $\log|\lambda| = 0$ via $\hat{R}$!}

$(\mathcal{T}^t)$ descends to a flow on $\hat{S}$. ← {since $\mathcal{T}^t \hat{R} = \hat{R} \mathcal{T}^t$}

$\mathcal{H}$ embeds injectively in $\hat{S}$; image $=: S$.

Note: R = the first return map of $\mathcal{T}^t$ to $S = \mathcal{H}$.

Indeed, recall def: $R(\pi,\lambda,\tau) = \hat{R}(\mathcal{T}^{t_R(\pi,\lambda)}(\pi,\lambda,\tau))$.
By contrast, $\mathcal{T}^t$ never returns in $\hat{\mathcal{H}}$!

Note: $\hat{m}$ and $\hat{m}_c$ give well-defined measures on $\hat{S}$ namely by "intersecting with a fundamental domain"; cf. Lecture #5.

Theorem 2: $\hat{m}_1(\hat{S}) < \infty$.

Viana proves this in Sec. 23–24; note for a fixed Rauzy class one can in principle check it by a "direct computation".

Recurrence

Def: If $(M, \mathcal{B}, \mu)$ is a measure space (possibly $\mu(M)=\infty$) and $f: M \to M$ is m-ble and non-singular (viz., $\mu(f^{-1}B) = 0 \Leftrightarrow \mu(B) = 0$, $\forall\ B \in \mathcal{B}$), then (f,μ) is called recurrent if for every $E \in \mathcal{B}$: $\mu(\{x \in E : f^n(x) \notin E\ (\forall n \geq 1)\}) = 0$.

alt: conservative

Note: If $\mu(M) < \infty$ and $f_*\mu = \mu$ then (f,μ) is recurrent, by the Poincaré Recurrence Theorem.

Lemma 1 (Viana Lemma 25.1): For any $t > 0$, $(\hat{\mathcal{S}}, \hat{m}, \mathcal{T}^t)$ and $(\hat{S}, \hat{\nu}, T^t)$ are recurrent.

$\hat{S}$ = a projected form of $\hat{\mathcal{S}}$, we don't give the precise def. here!

Also $(\mathcal{H}, m, \mathcal{R})$ and $(C \times \Lambda_A, \nu, R)$ are recurrent.

Proof: Write $\hat{\mathcal{S}}_c := \{(\pi, \lambda, \tau) \in \hat{\mathcal{S}} : \text{Area } M(\pi, \lambda, \tau) \leq c\}$.

Thm 1 & scaling $\Rightarrow \hat{m}(\hat{\mathcal{S}}_c) < \infty$.

Hence $(\hat{\mathcal{S}}_c, \hat{m}, \mathcal{T}^t)$ is recurrent $(\forall c > 0)$.

Using $\hat{\mathcal{S}} = \bigcup_{c=1}^{\infty} \hat{\mathcal{S}}_c \Rightarrow (\hat{\mathcal{S}}, \hat{m}, \mathcal{T}^t)$ is recurrent.

Projecting $\Rightarrow (\hat{S}, \hat{\nu}, T^t)$ is recurrent.

Next, if $(\mathcal{H}, m, \mathcal{R})$ not recurrent then easily...

$\exists E \subset \mathcal{H}$, $m(E) > 0$, $\forall n \geq 1 : \mathcal{R}^n(E) \cap E = \emptyset$.

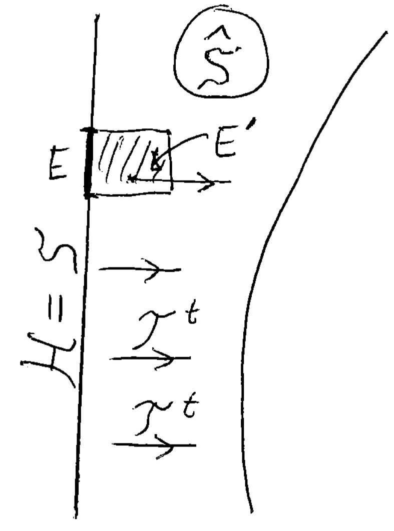
Then set $E' = \{\mathcal{T}^t(x) : x \in E, 0 \leq t < \min(1, t_R(x))\}$

$\hat{m}(E') > 0$.

$\therefore \exists y \in E'$, $n \geq 1$ s.t.

$\mathcal{T}^n(y) \in E'$.

$\Rightarrow$ contradiction!



Similar proof for $(C \times \Lambda_A, \nu, R)$

$\square\square$

If (M, μ, f) recurrent {and non-singular} and $D \overset{\text{m-ble}}{\subset} M$, $\mu(D) > 0$, then the first-return map $f_D : D \to D$ is well-def {on a full measure subset of D.}

{~~see~~ Lecture #13, p. 3 — which I've corrected.}

Lemma 2: In this situation, if also $f_*\mu = \mu$ and $\mu(D) < \infty$, then $(f_D)_*(\mu_D) = \mu_D$!

(Also (f, μ) ergodic $\Rightarrow$ (f_D, μ_D) ergodic) {But not conversely, in general!}

{A partial converse to ~~[illegible]~~}

Lemma 3:

In the above situation, if f_{1D} is a bijection onto a set of full measure in M, and $[\forall N \subset D : \mu(N) = 0 \Rightarrow \mu(f(N)) = 0]$, {i.e. N m-ble and....} then

(f_D, μ_D) ergodic $\Rightarrow$ (f, μ) ergodic

{Note: f_D and f_{1D} are completely different maps!}

For $\pi \in \mathcal{C}$, $N \geq 1$, $\underline{\varepsilon} = (\varepsilon_0, \dots, \varepsilon_{N-1}) \in \{0,1\}^N$, set

$$\underline{\Lambda_{\pi,N,\underline{\varepsilon}} = \{\lambda \in \Lambda_{\mathcal{A}} : \ R^k(\pi,\lambda) \text{ has type } \varepsilon_k, \text{ for } k=0,\dots,N-1\}}$$

Then $\underline{R \text{ maps } \{\pi\} \times \Lambda_{\pi,N,\underline{\varepsilon}} \text{ bijectively onto } \{\pi'\} \times \Lambda_{\pi',N-1,(\varepsilon_1,\dots,\varepsilon_{N-1})}}$

π' determined by π and ε_0

Map: $\lambda \longmapsto \lambda'$

$\lambda' = \Theta^{*-1}_{\pi,\varepsilon_0}(\lambda)$

$\lambda = \Theta^{*}_{\pi,\varepsilon_0}(\lambda')$

<u>proof</u>: For $N=1$ we've noted this before (convention: $\Lambda_{\pi',0,()} = \Lambda_{\mathcal{A}}$!)

For $N>1$ we are restricting the $N=1$ bijection to a <u>subset</u>...

Iterate!

$\Rightarrow$ $\underline{R^N \text{ maps } \{\pi\} \times \Lambda_{\pi,N,\underline{\varepsilon}} \text{ bijectively onto } \{\pi^N\} \times \Lambda_{\mathcal{A}}}$

π^N determined by π and $\underline{\varepsilon}$

Map: $\lambda \mapsto \lambda^N$

$$\lambda = \underbrace{\Theta^*_{\pi,\varepsilon_0}\Theta^*_{\pi',\varepsilon_1}\cdots\Theta^*_{\pi^{N-1},\varepsilon_{N-1}}}_{=:\Theta^{N*}}(\lambda^N)$$

Given $\pi \in C$, can find N and $\underline{\varepsilon}$ s.t. all entries of Θ^{N*} are positive!

Viana Cor. 5.3

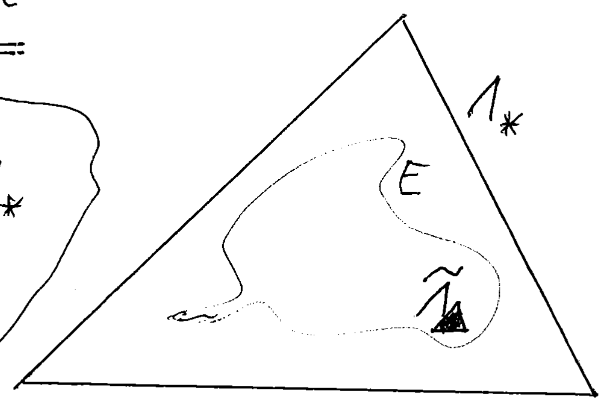
$\Rightarrow \underline{\underline{\Lambda_*}} := \Lambda_{\pi, N, \underline{\varepsilon}} = \Theta^{N*}(\Lambda_{\mathcal{A}})$ is relatively compact in $\Lambda_{\mathcal{A}}$!

identify Λ_* with $\{\pi\}\times\Lambda_*$

$\Rightarrow \nu(\Lambda_*) < \infty$

Set $R_* = (R^N)_{\Lambda_*} : \Lambda_* \to \Lambda_*$.

Now $(\Lambda_*, \nu_{\Lambda_*}, R_*)$ is ergodic



proof: Viana Prop 25.5 & my notes.

Use $\nu_{\Lambda_*} \ll\gg$ Leb. Assume $E \subset \Lambda_*$ R_*-invariant, $\nu_{\Lambda_*}(E) > 0$.

Lebesgue points etc $\Rightarrow$ Can find $\tilde{\Lambda} = \Lambda_{\pi, (k+1)N, \underline{\theta}} \subset \Lambda_*$ s.t. R_*^k for some $k \geq 1$ maps $\tilde{\Lambda}$ bijectively on Λ_*, and $\text{Leb}(\tilde{\Lambda} \setminus E) < 10^{-6} \cdot \text{Leb}(\tilde{\Lambda})$.

But E R_*-inv $\Rightarrow \tilde{\Lambda} \setminus E = \underbrace{\left((R_*^k|_{\tilde{\Lambda}})\right)^{-1}}(\Lambda_* \setminus E)$

unif control on volume scaling...

$\Rightarrow \text{Leb}(\Lambda_* \setminus E) < C \cdot 10^{-6} \cdot \text{Leb}(\Lambda^*)$

unif!

Etc $\Rightarrow \text{Leb}(\Lambda_* \setminus E) = 0 \Rightarrow \nu(\Lambda_* \setminus E) = 0$.

See my notes!

Hence get $(C \times \Lambda_{\mathcal{A}}, \nu, R)$ ergodic, by Lemma 3 + more.

Uniqueness of ν: See Problem 52! □ □ Thm 1 proved!

<u>Remark</u>: Thm 1 (especially the fact that ν is ergodic) is crucial for proving:

<u>For a.e. $(\pi,\lambda) \in C \times \Lambda_{\mathcal{A}}$, $f_{\pi,\lambda} : I \circlearrowleft$ is uniquely ergodic.</u> ("Keane's conjecture")

Viana proves this in Sec. 28-29; the key (beyond Thm 1) is to study the cone $M(\pi,\lambda)$ of finite $f_{\pi,\lambda}$-invariant measures of $I_{\pi,\lambda}$; one proves this set is $\cong \bigcap_{n=1}^{\infty} \Theta_{\pi,\lambda}^{n*}(\mathbb{R}_+^{\mathcal{A}})$.

15.1. **Notes.** .

p. 1: Theorem 1 is [49, Thm. 7.2].

p. 2: The measures $\widehat{m}$ and m are defined in [49, Sec. 21]. (See also my notes to [49, Sec. 21], especially regarding $d_1\lambda$.)

p. 3: Also $\widehat{m}_c$ and m_c are defined in [49, Sec. 21]. The quotient space $\hat{\mathcal{S}}$ is defined in [49, Sec. 20].

p. 4: Theorem 2 is [49, Thm. 24.1]. Viana proves this in [49, Sec. 23–24]. The concept of recurrence is defined on [49, p. 80]; cf. also Aaronson [1, Sec. 1.1]; one can fairly easily prove that a non-singular map is recurrent iff it is *conservative* as defined in [1, p. 15(bottom)].

p. 5: This is [49, Lemma 25.1].

p. 6: The first return map is defined on [49, p. 80(middle)]; cf. also [1, Sec. 1.5]. Lemma 2 on p. 6 is [49, Remark 25.3]. Lemma 3 is a variant of [49, Lemma 25.4]; cf. my notes to Viana's notes.

pp. 7–8: Here we follow [49, p. 82] and then [49, pp. 87–88] (see also my notes about details in Viana's proof of Prop. 25.5). At the bottom of p. 8: Note that in order to conclude that $(C \times \Lambda_{\mathcal{A}}, \nu, R)$ is ergodic, it is not sufficient to use Lemma 3 (p. 6) since $(R^N)_{|\Lambda_*}$ is *not* a bijection onto all of $C \times \Lambda_{\mathcal{A}}$. Viana does not seem to pay sufficient attention to the fact that "$C \times \Lambda_{\mathcal{A}}$ consists of *several* copies of $\Lambda_{\mathcal{A}}$". I have attempted to complete the proof in my notes to Viana's Cor. 27.2. [Brief outline, in the set-up of the lecture: In the construction of Λ_* (p. 8 of the lecture) we can take π arbitrary and then arrange that $\pi^N = \pi$. Then $(R^N)_{|\Lambda_*}$ is a bijection onto $\{\pi\} \times \Lambda_{\mathcal{A}}$, and so by Lemma 3 (p. 6 in the lecture), the fact that $(\Lambda_*, \nu_{\Lambda_*}, R_*)$ is ergodic implies that $(D, \nu_D, (R^N)_D)$ is ergodic, for $D = \{\pi\} \times \Lambda_{\mathcal{A}}$. The fact that there is such an N for every $\pi \in C$ can be shown to imply that $(C \times \Lambda_{\mathcal{A}}, \nu, R)$ is ergodic.]

16. Translation surfaces I

Lecture #16: Translation surfaces

Def: A t.s. is a compact Riemann surface M together with a holomorphic 1-form $\alpha \, (\not\equiv 0)$ on M.

To see, write $\alpha = dz$, or near a zero: $\alpha = z^m dz$

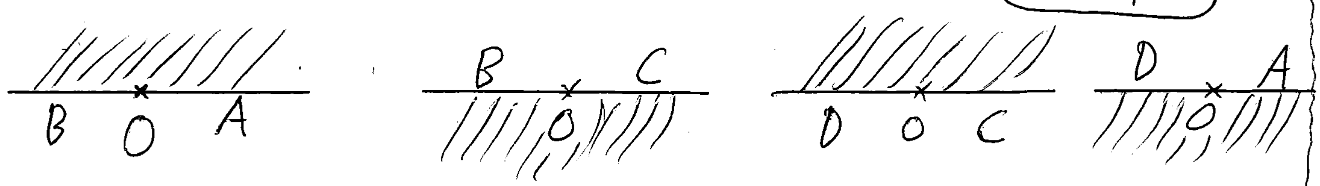
A t.s. is a compact 2-dim mfld M with a flat Riemannian metric having conical singularities $p_1, \dots, p_K$ with angles $2\pi(m_i + 1)$ $(m_i \in \mathbb{Z}^+)$, $i = 1, \dots, K$, together with a parallel unit vector field on $M \setminus \{p_1, \dots, p_K\}$.

the "vertical direction"

2nd version, more explicitly:

A t.s. is a compact 2-dim mfld M with selected points $p_1, \dots, p_K$, provided with a translation atlas on $M \setminus \{p_1, \dots, p_K\}$, that is, an atlas of coordinate neighbourhoods whose transition maps are translations, such that for each p_i there is a neighbourhood $U_i \subset M$ and a homeomorphism of U_i onto an open subset of $2(m_i + 1)$ glued half planes:

ex: $m_i = 1$



taking p_i to O and being an isometry on $U_i \setminus \{p_i\}$.

Def (3rd version, most concrete): A t.s. is a finite set of polygons in $\mathbb{R}^2 = \mathbb{C}$ together with a choice of pairing of parallel sides of equal length that are on "opposite sides". Consider this up to equivalence defined by "cutting in pieces" & "re-gluing".

ex:

$\Leftrightarrow$ torus!

Aside: A more general setting often discussed:
A compact Riemann surface with a quadratic differential
– this corresponds to allowing conical singularities with angles $k_i \pi$ ($k_i \in \mathbb{Z}$) – "half-translation surfaces"

Central topic: Asymptotic behaviour of geodesics on M?

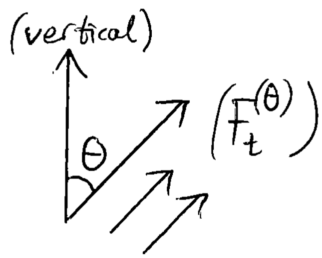
viz. "straight lines"; unclear how to extend through a singular point; there are m_i+1 natural choices!

* Closed geodesics — These come in families of parallel closed geodesics of same length sweeping out an annulus. The bdry of the annulus consists of saddle connections, i.e. straight line segments between two singular points. Counting such families of closed geodesics, and saddle connections, asymptotically wrt length is a much studied topic.

we'll focus mainly on

* Arbitrary (long) geodesics

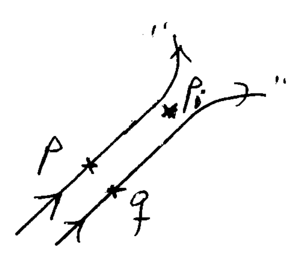
Let $\left(F_t^{(\theta)}\right)_{t\in\mathbb{R}}$ = geodesic flow in direction θ on M.



Convention: If $F_{t_0}^{(\theta)}(p) = \text{sing. } p_i$ then $F_t^{(\theta)}(p)$ undefined for
$$\begin{cases} t \geq t_0 & (\text{if } t_0 \geq 0) \\ t \leq t_0 & (\text{if } t_0 \leq 0) \end{cases}$$

Note: $\left(F_t^{(\theta)}\right)$ is discontinuous

Ex



p, q close

$F_1(p), F_1(q)$ far from each other!

Note: $(F_t^{(\theta)})$ preserves Leb. (the Lebesgue area measure on M).

Theorem 1 (Kerckhoff, Masur, Smillie '86): Given any t.s. M: For a.e. θ, the flow $(F_t^{(\theta)})$ on M is uniquely ergodic.

In Lectures 17–18 we'll study how long geodesics "wrap around the handles of M" asymptotically, i.e. the homology of a long geodesic (when closing it up in a natural way).

Much easier result:

Theorem 2: $(F_t^{(\theta)})$ is minimal iff $\neg \exists$ closed $(F_t^{(\theta)})$-orbit.

Problem 57

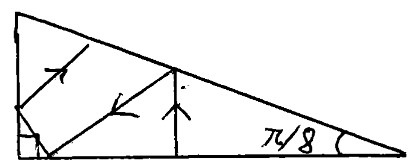
As we mentioned above, if $\exists$ closed $(F_t^{(\theta)})$-orbit then $\exists$ saddle connection in direction θ, and this can only happen for countably many θ (Problem 56). Hence Thm 2 $\Rightarrow$ $(F_t^{(\theta)})$ is minimal for all except countably many θ.

Note: There exist M and θ for which $(F_t^{(\theta)})$ is minimal but not uniquely ergodic.

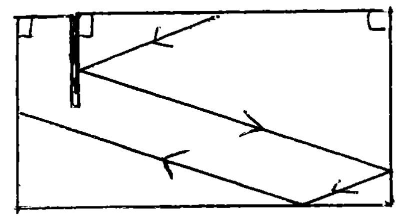
Important application: Rational billiards

Consider a game of billiards in a polygon Q, all of whose angles are rational multiples of π.

ex: 1:



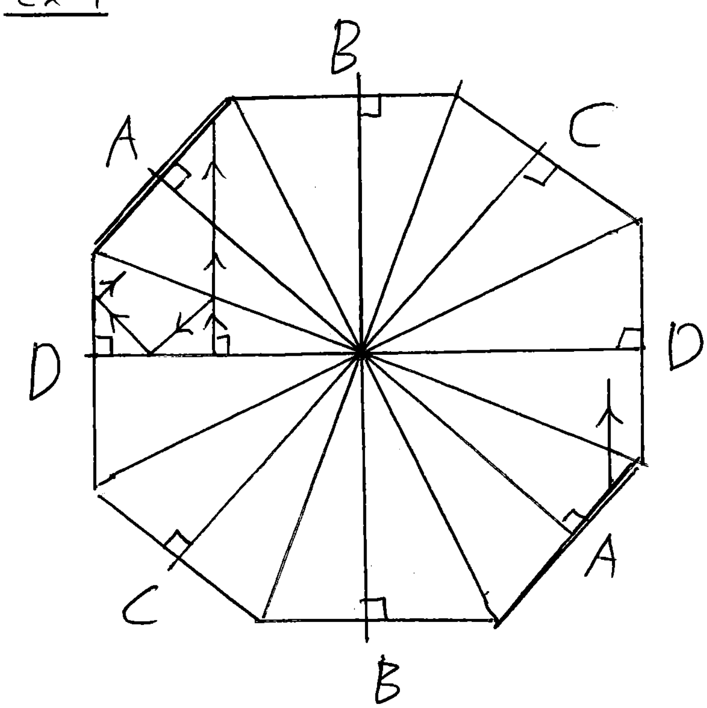
ex: 2:



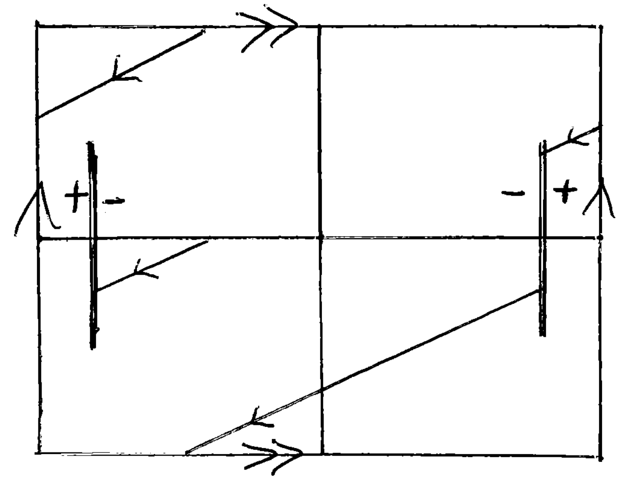
(generalized 7-gon)

By a natural unfolding procedure, this becomes equivalent to geodesic flow on a certain t.s.:

ex 1



ex 2



Billiards in non-rational polygons: Much less is known!
OPEN PROBLEM: Is almost every polygon (or even a positive measure set of polygons) ergodic ?? (wrt. Liouville measure on the unit tangent bundle)

Note Kerckhoff, Masur, Smillie prove that their Thm 1 $\Rightarrow$ any polygon which is "very well approximable by rational polygons" (but is not rational itself) is ergodic!

Also OPEN PROBLEM: $\exists$ closed orbits?
– Not even known for a general obtuse triangle!

For billiards whose walls are (partially) convex curves (thus collisions are dispersive) – very different situation! Strong chaotic features (Bernoulli, entropy > 0, exponential decay of correlations).
Ex: Sinai billiard, Lorentz gas, but also Bunimovich stadium, despite collisions being focusing.

16.1. **Notes.** .

p. 1–2: For presentations of the various equivalent definitions of a translation surface, cf. [49, Sec. 11], but also, e.g. [32, Sec. 1] and [52, Sec. 1].

p. 2: The more general setting with quadratic differentials is for example considered in Veech 1986 [46]; Veech 1990 [47]; Masur 1990 [31]; Veech 1998 [48]; Eskin-Masur 2001 [13].

p. 3: Some key references regarding asymptotics of *closed* geodesics and saddle connections are Masur 1990 [31], Eskin-Masur 2001 [13],

p. 4: For Theorem 1, cf. Kerchhoff-Masur-Smillie 1986 [23]. For the statement at the bottom of the page on existence of M and θ for which $(F_t^{(\theta)})$ is minimal but not uniquely ergodic, cf. [32, Sec. 4] and the references therein.

p. 5: For a description of the unfolding procedure, see [33, Sec. 1.5]. Our ex. 1 is from loc. cit., and our ex. 2 is a somewhat generalized version of the example in [32, Fig. 2 and Thm. 2].

p. 6, on billiards in non-rational polygons: See [17, Question 47] regarding the first open problem. The ergodicity result by Kerchhoff-Masur-Smillie is proved in [23, Sec. 5]. Regarding the second open problem, cf., e.g., [17, Question 46] and [42].

p. 6, on more general billiards: Cf., e.g., the book by Chernov and Markarian, [8].

17. Translation surfaces II

Lecture #17: Translation surfaces II

For any t.s., $2g-2 = \sum_{i=1}^{K} m_i$. ⊛

3 proofs! One of these: Problem 58.

Now take (any) $g \geq 1$ and $1 \leq m_1 \leq \dots \leq m_K$ subject to ⊛ (note $g=1 \Rightarrow K=0$); then there exist many t.s. with such data. Let $\mathcal{A} := \mathcal{A}_g(m_1, \dots, m_K)$ be the moduli space of all such t.s.

Viz., the set of all t.s. with genus g and singularities of order $m_1, \dots, m_K$, up to strong isometry.

Here strong isometry $\stackrel{\text{def}}{=}$ an isometry which also preserves the "vertical up" vector field.

$\mathcal{A}$ has a natural complex orbifold structure,

$\dim_{\mathbb{C}} \mathcal{A} = 2g+K-1$

M_1 near M_2 in $\mathcal{A}$ if M_2 obtains from M_1 by "small deformation".

We've seen $2g+K-1=d$ for a suspension surface of an IET; consistent with $(\lambda, \tau) \in \mathbb{R}^{2d}$!

$\mathcal{A}$ has 1 or 2 or 3 components — Kontsevich & Zorich '03

Background; classical theory

$\mathcal{M}_g$ = the "moduli space of compact Riemann surfaces, genus g.

Teichmüller showed $\mathcal{M}_g = T_g / \Gamma_g$

A complex manifold, $\dim_{\mathbb{C}} \mathcal{M}_g = 3g-3$

the mapping class group; a discrete group of biholomorphisms of T_g

Also, the cotangent space at any $M \in \mathcal{M}_g$ "=" the space of quadratic differentials on M!

naturally

Let C be any <u>nondegenerate</u> Rauzy class

{viz., "all $m_i > 0$" in Viana's Sec. 11 & 14}

"giving" g and $1 \leq m_1 \leq \dots \leq m_K$.

Then $\exists$ obvious map

$$\hat{S}(C) = \langle \hat{R} \rangle \setminus \hat{H}(C) \longrightarrow \mathcal{A}$$

This map is continuous; hence its image is contained in a connected component of $\mathcal{A}$. We'll now give a more explicit understanding...

Summary (of what we'll explain below):

All vertical arrows: ramified (finite) cover!

$$\begin{tikzcd}
\hat{S}(C) \arrow[r, hook, "\text{``}\quad\text{''}"] \arrow[dr] & \mathcal{A}'_C \arrow[d] & \overset{(\text{component of})}{\subset} & \mathcal{A}'_m \arrow[d] \\
& \mathcal{A}_C & \overset{(\text{component of})}{\subset} & \mathcal{A}_{m}
\end{tikzcd}$$

Let m be the order of the "0-vertex" of any $M(\pi, \lambda, \tau)$, $\pi \in C$. {Same for all $\pi \in C$!}

For $M \in \mathcal{A}$, call a geodesic ray r on M an "<u>m-separatrix</u>" if r starts at a singular point of order m and r has direction "$\longrightarrow$ horizontal".

{Clearly there are exactly $\#\{j : m_j = m\} \cdot (m+1)$ m-separatrices on any $M \in \mathcal{A}$!}

Set $\mathcal{A}'_m := \{(M,r) : M \in \mathcal{A} \text{ and } r \text{ is an } m\text{-separatrix on } M\}$

{Note: As $\mathcal{A}$, also $\mathcal{A}'_m$ is a complex orbifold}

The map $\mathcal{A}'_m \to \mathcal{A}$, $(M,r) \mapsto M$ makes $\mathcal{A}'_m$ a ramified cover of $\mathcal{A}$ of degree $\#\{j : m_j = m\} \cdot (m+1)$.

{Ramification points: Any $M \in \mathcal{A}$ which has a ~~strong~~ self-isometry taking some m-separatrix to another m-separatrix.}

Now the obvious map

$$\hat{\mathcal{S}}(C) \longrightarrow \mathcal{A}'_m$$

$$(\pi, \lambda, \tau) \longmapsto \langle M(\pi,\lambda,\tau), \text{"positive } x\text{-axis"} \rangle \qquad ⊛$$

is a homeomorphism of $\hat{\mathcal{S}}(C)$ onto an open subset of $\mathcal{A}'_m$ which equals a component $\mathcal{A}'_C$ of $\mathcal{A}'_m$ minus a countable union of real-analytic submanifolds of (real) codim 2.

{In particular: The image has full measure in $\mathcal{A}'_C$}

{Also: $\zeta_j = \lambda_j + i\tau_j$ $(j = 1, \dots, d)$ are complex analytic local coordinates on $\mathcal{A}'_m$, away from orbifold points!}

{~~proof~~ sketch: The map ⊛ is well-def. since $\hat{\mathcal{R}}$ preserves the t.s.
Def of orbifold structure of $\mathcal{A} \Rightarrow \zeta_1, \dots, \zeta_d$ are local coords
$\Rightarrow$ the map ⊛ is open & continuous.
Injectivity: Veech '82 (Prop 9.1)
Image of ⊛? The argument in Viana '07, Sec. 1.2.3 gives:
a full component up to codim 1. Boissy '12 $\rightsquigarrow$ codim 2!}

{1-1-corr between components of $\mathcal{A}'_m$ and components of $\mathcal{A}$??}

Remark: $\mathcal{A}_C = \mathcal{A}_{C'}$ iff C, C' belong to the same extended Rauzy class!

Cf. Viana 2007, Sec. 1.1.4 & 1.2.3. – However it seems better to define directly for the monodromy invariant $p = \pi_1 \circ \pi_0^{-1}$

Now $(\mathcal{T}^t)$ {the Teichmüller flow, $(\pi, \lambda, \tau) \mapsto (\pi, e^t\lambda, e^{-t}\tau)$}

is well-def on each of $\hat{\mathcal{S}}(C)$, $\mathcal{A}_C'$, $\mathcal{A}_C$ (and $\mathcal{A}_m'$, $\mathcal{A}$).

$\hat{m}$ $(= d\pi\, d\lambda\, d\tau)$ and $\hat{m}_1$ give well-def measures on $\mathcal{A}_C'$ and $\mathcal{A}_C$, invariant under $(\mathcal{T}^t)$.

$\hat{m}_1(\mathcal{A}_C') < \infty \implies \boxed{\hat{m}_1(\mathcal{A}_C) < \infty}$

$\tilde{m}_1$: see notes

Theorem 1: $(\mathcal{A}_C, \tilde{m}_1, (\mathcal{T}^t))$ is ergodic — Veech '86

Note: All $SL_2(\mathbb{R})$ acts on $\mathcal{A}_C$; def: "postcomposition in each chart". This action preserves $\hat{m}$ and $\hat{m}_1$.
(However the $SL_2(\mathbb{R})$-action does not lift to $\mathcal{A}_m'$.
Proof: Consider $SO_2(\mathbb{R}) \subset SL_2(\mathbb{R})$...)

Considering the set of saddle connections (or the annuli of closed geodesics) for a random M in $(\mathcal{A}_C, \hat{m}_1)$ gives a random discrete set in $\mathbb{R}^2$. In fact, this is an $SL_2(\mathbb{R})$-invariant point process in $\mathbb{R}^2$.

Quick (but non-standard) def. of $H_1(M)$

M - any path-connected topological space.

A **loop** (in M) is a continuous map $\sigma : S^1 \to M$.

Now $H_1(M) := L/Z$ where L is the free abelian group generated by all loops in M, and Z is the subgroup of L generated by

$\{\sigma_1 - \sigma_2 \ : \ \sigma_1, \sigma_2 \text{ any two } \underline{\text{homotopic}} \text{ loops}\}$

and

$\{\sigma_1 + \sigma_2 - \underbrace{\sigma_1 \cdot \sigma_2}_{} \ : \ \sigma_1, \sigma_2 \text{ any loops with } \sigma_1(1) = \sigma_2(1)\}$

Concatenation; $\sigma_1 \cdot \sigma_2(z) = \begin{cases} \sigma_1(z^2) & \operatorname{Im} z \geq 0 \\ \sigma_2(z^2) & \operatorname{Im} z \leq 0 \end{cases}$

write $S^1 = \{z \in \mathbb{C} : |z| = 1\}$

Recall the standard def. of homology groups, and prove that the above gives the correct $H_1(M)$. — Problem 59.

Remark: $H_1(M)$ = abelianization of $\pi_1(M)$!

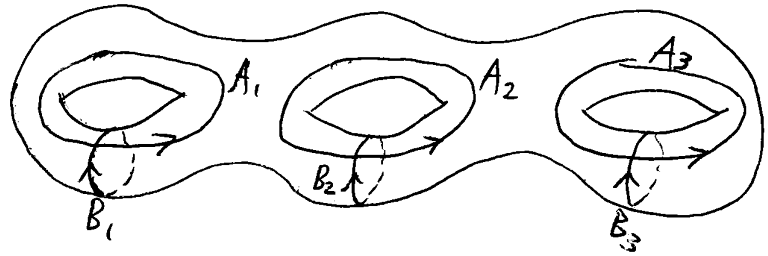
First observations: For $\sigma : S^1 \to M$ constant (viz., a point), $\sigma \in Z$, i.e. $\sigma = 0$ in $H_1(M)$. (Proof: $\sigma = \sigma + \sigma - \sigma \cdot \sigma$!)

Next, for any loop σ, if $\bar{\sigma} = [\sigma \text{ reversed}]$, then $\bar{\sigma} = -\sigma$ in $H_1(M)$.

Fact: If M is a compact (orientable) surface of genus g:

$$H_1(M) \cong \mathbb{Z}^{2g}.$$

"Canonical" basis:



Also: $H_1(M,\mathbb{R}) := H_1(M) \otimes_{\mathbb{Z}} \mathbb{R}$ ← homology with coefficients

$H_1(M,\mathbb{Z}) = H_1(M)$

$H^1(M) = H^1(M,\mathbb{Z}) = H_1(M)^*$ ← cohomology groups

$H^1(M,\mathbb{R}) = H_1(M,\mathbb{R})^*$

de Rham: $H^1(M,\mathbb{R}) = [\text{closed 1-forms}]/[\text{exact 1-forms}]$

by integrating the 1-form around the loop.

For M a t.s., $p \in M$, $\ell > 0$, let

$\gamma(p,\ell) = [\text{the } \uparrow \text{ vertical geodesic of length } \ell \text{ starting at } p]$

If it hits a singular point, continue along the next separatrix in the counter-clockwise direction; thus $\gamma(p,\ell)$ def for all p, ℓ.

$\forall\, p,p' \in M$: Fix a curve $c_{p,p'}$ from p to p', so that $\sup_{p,p'} \text{length}(c_{p,p'}) < \infty$.

Let $[\gamma(p,\ell)]$ = the element in $H_1(M)$ represented by $\gamma(p,\ell) \cdot c_{p,p'}$ where p' = endpoint of $\gamma(p,\ell)$.

Clearly $[\gamma(p,\ell)]$ tells how $\gamma(p,\ell)$ "wraps around the handles of M"!

vertical $\uparrow$ flow; cf. Lecture #16.

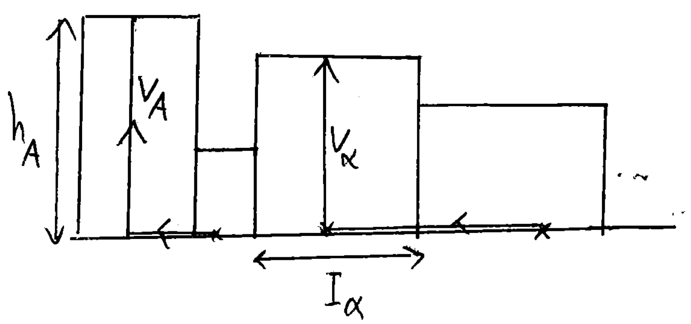
Theorem 2: If $(F_t^{(0)})$ is uniquely ergodic ⊛ then $\exists\, c_1 \in H_1(M,\mathbb{R})$ s.t. $\frac{1}{\ell}[\gamma(p,\ell)] \xrightarrow{\ell\to\infty} c_1$ in $H_1(M,\mathbb{R})$ uniformly over all $p \in M$.

⊛ Recall that by Kerckhoff-Masur-Smillie (Thm 1, Lecture #16) this holds for almost every rotation of the given M.

In next lecture: we'll see more precise asymptotics, coming from Lyapunov exponents of the Rauzy-Veech cocycle.

Explicit formula for c_1

Assume $M = M(\pi,\lambda,\tau,h)$ $\quad (h = -\Omega_{\pi,\lambda}(\tau))$.



$[V_\alpha] \in H_1(M)$ $\quad (\alpha \in \mathcal{A})$

Definition as in the picture.

Then $\underline{\underline{H_1(M) = \sum_{\alpha\in\mathcal{A}} \mathbb{Z}[V_\alpha]}}$

How prove? ... See Problem 60.

Now
$$\boxed{c_1 = \frac{\sum_{\alpha\in\mathcal{A}} \lambda_\alpha [V_\alpha]}{\sum_{\alpha\in\mathcal{A}} h_\alpha \lambda_\alpha}}$$
Area(M)

Idea of proof of Theorem 2:

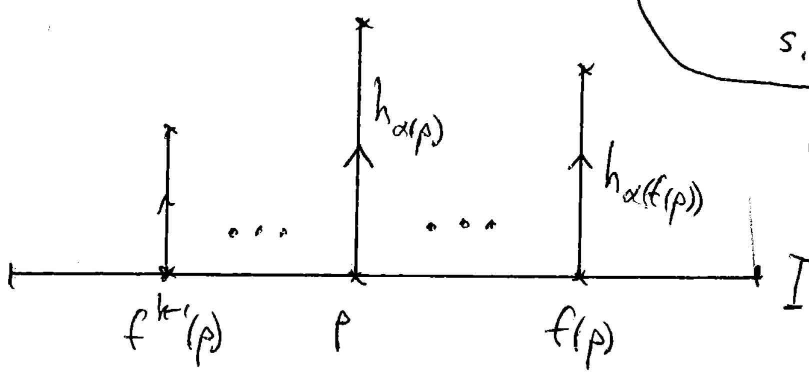
Reduce to the case $p \in I$ and $p' = [\text{endpoint } \gamma(p,\ell)] \in I$.

Say k steps, thus:

$\alpha(x) :=$ the α s.t. $x \in I_\alpha$

$f = f_{\pi,\lambda}$

$\underline{\gamma(p,\ell)} =$



$$\underline{[\gamma(p,\ell)]} = \sum_{j=0}^{k-1} [v_{\alpha(f^j(p))}] = \sum_{\alpha \in \mathcal{A}} \#\{0 \le j < k : f^j(p) \in I_\alpha\} \cdot [v_\alpha]$$

assuming $c_{p,p'} \in I$

$$\sim \underline{\left(\sum_{\alpha \in \mathcal{A}} \frac{\lambda_\alpha}{|\lambda|} \cdot [v_\alpha]\right) \cdot k \quad \text{as } k \to \infty.}$$

Uniformly wrt $p \in I$, since f is uniquely ergodic!

— not completely easy; cf. Problem 10!

Also $\underline{\ell = \sum_{j=0}^{k-1} h_{\alpha(f^j(p))} \sim \left(\sum_{\alpha \in \mathcal{A}} \frac{\lambda_\alpha}{|\lambda|} \cdot h_\alpha\right) \cdot k} \quad$ as $k \to \infty$

$$\therefore \quad \underline{\frac{[\gamma(p,\ell)]}{\ell} \to c_1 \quad \text{as } k \to \infty}$$

$\square$

17.1. **Notes.** .

p. 1: A detailed description of the complex orbifold structure on $\mathcal{A}_g(m_1,\ldots,m_\kappa)$ (and also the more general analogous spaces for quadratic differentials) can be found in Veech 1990, [47]. In fact $\mathcal{A}_g(m_1,\ldots,m_\kappa)$ is obtained from a complex affine *manifold* ("$V(\pi)$" in [47]) of (complex) dimension $2g+\kappa-1$, when taking the quotient by the action of a discrete group of biholomorphisms. The classification of the connected components of $\mathcal{A}_g(m_1,\ldots,m_\kappa)$ was obtained in Kontsevich & Zorich, 2003, [24].

p. 1: Regarding the real/complex analytic theory of the Teichmüller space $\mathcal{T}_g$, cf., e.g., Abikoff [2] and Nag [35].

pp. 2–3: For the statements we make about the map $\hat{\mathcal{S}}(C) \to \mathcal{A}'_m$, cf. Boissy, [6, Lemma 2.1 and Prop. 2.2]; note that the key to prove the injectivity of the map is Veech 1982 [45, Prop. 9.1].

p. 4: Regarding the fact that the components $\mathcal{A}_C$ and $\mathcal{A}_{C'}$ of $\mathcal{A}_g(m_1,\ldots,m_\kappa)$ are equal iff C and C' belong to the same *extended* Rauzy class: See Kontsevich & Zorich, [24, Appendix A] (they attribute this fact to Veech 1982 [45]; but I do not see exactly how to derive the statement from that paper).

p. 4, Theorem 1: See Veech 1986 [46], where much more is proved! In our formulation of Theorem 1 we write $\widetilde{m}_1$ for the natural induced measure *on* $\{M \in \mathcal{A}_C \,:\, \text{area}(M) = 1\}$; recall that by contrast, $\widehat{m}_1$ has support on all $\{M \in \mathcal{A}_C \,:\, \text{area}(M) \leq 1\}$. Here is one way to define $\widetilde{m}_1$, normalized to be a probability measure:

$$\widetilde{m}_1(E) = \frac{\widehat{m}_1(\{\lambda M \,:\, \lambda \in (0,1],\ M \in E\})}{\widehat{m}_1(\mathcal{A}_C)},$$

for any Borel subset $E \subset \mathcal{A}_C \cap \{\text{area} = 1\}$, where λM denotes the t.s. M scaled by λ (thus $\text{area}(\lambda M) = \lambda^2$ for $M \in E$).

p. 4: The perspective of considering the $\text{SL}_2(\mathbb{R})$-invariant point processes in $\mathbb{R}^2$ mentioned here is important in Veech 1998 [48] and Eskin & Masur 2001 [13].

p. 5: For the general definition of the singular homology group $H_n(M)$, cf., e.g., Hatcher, [18, Ch. 2]. The fact that our non-standard definition of $H_1(M)$ is equivalent with the standard one (for M path connected) is essentially seen from the proof of the fact that $H_1(M)$ can be identified with the abelianization of $\pi_1(M)$ [18, Thm. 2A.1]; see Problem 59!

p. 6: Regarding the canonical basis of a compact orientable surface of genus g; cf. [18, Ex. 2A.2].

p. 6: Note that "$H_1(M,\mathbb{R}) := H_1(M) \otimes_{\mathbb{Z}} \mathbb{R}$" is *not* the general definition of "homology with coefficients" [18, pp. 153–]; however it holds for $\mathbb{Z}$ and $\mathbb{R}$ and a general space M; cf. [18, Thm. 3A.3 and Prop. 3A.5(3)]. Also, "$H^1(M) := H_1(M)^*$" (dual $\mathbb{Z}$-module) and "$H^1(M,\mathbb{R}) := H_1(M,\mathbb{R})^*$" (dual

$\mathbb{R}$-module) are *not* the general definitions of the cohomology groups [18, Ch. 3.1]; however these relations are valid for M a compact oriented surface; cf. [18, Thm. 3.2 and Cor. 3.3].

p. 6, regarding the definition of $\gamma(p,\ell)$ when passing through a singular point: It seems to me that we should then continue along the next separatrix in the *counter*-clockwise direction (contrary to what Viana writes on [50, p. 3]); namely in order for the first-return map to I for the vertical flow on $M(\pi,\lambda,\tau,h)$ to be exactly the IET $f_{\pi,\lambda}$. (Recall that by definition, each subinterval $I_\alpha \subset I$ is closed to the left and open to the right.)

p. 7, Theorem 2: This is [50, Theorem A], which is proved in [50, Sec. 3]. (It is not clear to me that the assumption that the vertical flow on M is uniquely ergodic necessarily implies that M has a presentation as a suspension surface $M(\pi,\lambda,\tau)$; hence this may have to be added as an extra assumption in the theorem.)

p. 7: Regarding the statement that $\{[v_\alpha] \,:\, \alpha \in \mathcal{A}\}$ spans $H_1(M)$; cf. Problem 60.

p. 8: This is a brief outline of the argument in [50, Sec. 3].

Lecture #18: Lyapunov exponents of Teichmüller flows

Review from lecture #7: For (X, μ, T) a ppt, a linear cocycle over T is a map $\tilde{T}: X \times \mathbb{R}^d \to X \times \mathbb{R}^d$ satisfying $p_1 \circ \tilde{T} = T \circ p_1$ ($p_1 = \text{proj}: X \times \mathbb{R}^d \to X$) and that $A(x) := \tilde{T}(x, \cdot): \cancel{\{x\}} \times \mathbb{R}^d \to \cancel{\{T(x)\}} \times \mathbb{R}^d$ is in $GL(d, \mathbb{R})$ $(\forall x \in X)$. Then define $A^n(x) := \tilde{T}^n(x, \cdot)$, i.e. $A_n(x) = A(T^{n-1}x) \cdot A(T^{n-2}x) \cdot \ldots \cdot A(x)$.

(Recall $A_{n+m}(x) = A_n(T^m x) A_m(x)$ – the "cocycle identity".)

Oseledet's MET: Assume $\log^+ \|A^{\pm 1}\| \in L^1$. Then for μ-a.e. $x \in X$, there exist $k = k(x) \in \mathbb{Z}^+$, $\lambda_1(x) < \cdots < \lambda_k(x)$ in $\mathbb{R}$ and a flag $\mathbb{R}^d = F_x^1 \supsetneq \cdots \supsetneq F_x^k \supsetneq 0$ s.t. $\forall \underline{v} \in F_x^i \setminus F_x^{i+1}$: $\lim_{n \to \infty} \frac{1}{n} \log \|A_n(x) \underline{v}\| = \lambda_i(x)$.

$F_x^{k+1} := 0$

Also $k, \{\lambda_i\}, \{F^i\}$ are T (resp. $\tilde{T}$)-invariant! The above formulation is a bit imprecise – see Lect#7 for a more precise version; here we are just aiming for a quick review. Note also that we have changed the notation a bit versus #7.

MET-2: If T inv'ble, then $\tilde{T}$ is inv'ble, and $\exists E_x^1, \ldots, E_x^k \subset \mathbb{R}^d$ (invariant!) with $F_x^i = \bigoplus_{j=i}^{k} E_x^j$

Remark 1: If (X', μ', T') is any extension of (X, μ, T) $\{\Leftrightarrow (X, \mu, T)$ is a factor of $(X', \mu', T')\}$

viz., $\exists$ m'ble map $P : X' \to X$ s.t. $P_* \mu' = \mu$

and $P \circ T' = T \circ P$ μ'-a.e.,

then there is a corresponding {"obvious"} extension of $\widetilde{T}$, namely

$$\widetilde{T}' : X' \times \mathbb{R}^d \to X' \times \mathbb{R}^d$$

$$\underline{\widetilde{T}'(x, \underline{v}) = (T'(x), A(P(x))\,\underline{v})} \qquad (\text{thus: } \underline{A'(x) := A(P(x))})$$

Then, if $\log^+ \|A^{\pm 1}\| \in L^1$, MET applied to $\widetilde{T}$

{immediate!} $\Rightarrow$

$$\left\{\begin{array}{l} k(x) := k(P(x)) \\ \lambda_i(x) := \lambda_i(P(x)) \\ F_x^i := F_{P(x)}^i \end{array}\right. \quad \text{"Works" for } \widetilde{T}' \,!$$

{Abuse of notation! Now $k(x)$, $\lambda_i(x)$, F_x^i have meaning both for $x \in X$ and $x \in X'$!}

{i.e., satisfies the conclusions of MET for $\widetilde{T}'$!}

{Here it may be that T' is inv.ble even if T is not. Then "MET-2" applies to $\widetilde{T}'$ and gives the existence of (E_x^i) for $x \in X$; these will in general not "depend on x only via $P(x)$", even though the partial sums $F_x^i = \bigoplus_{j=i}^{k} E_x^j$ do!}

The Rauzy-Veech cocycle

Now take $(X, \mu, T) = (C \times \mathbb{R}_+^{\mathcal{A}}, \hat{\nu}, \hat{R})$

not a ppt; $\hat{\nu}(C \times \mathbb{R}_+^{\mathcal{A}}) = \infty$

Recall: $C \times \mathbb{R}_+^{\mathcal{A}}$ should really be replaced by the full measure subset of (π, λ) satisfying the Keane condition.

Linear cocycle over $\hat{R}$: $F_{\hat{R}} : C \times \mathbb{R}_+^{\mathcal{A}} \times \mathbb{R}^{\mathcal{A}} \circlearrowleft$

$$(\pi, \lambda, \underline{v}) \mapsto \left(\hat{R}(\pi, \lambda), \ \theta_{\pi,\lambda} \cdot \underline{v}\right)$$

we won't use the "A-notation"

Thus $\cancel{A(\pi,\lambda)} = \theta_{\pi,\lambda}$

and $\cancel{A^n(\pi,\lambda)} = \theta^n_{\pi,\lambda} = \theta_{\hat{R}^{n-1}(\pi,\lambda)} \cdot \ldots \cdot \theta_{\pi,\lambda}$

We worked with this $\theta^n_{\pi,\lambda}$ also in Lecture #15!

Recall here: $\theta_{\pi,\lambda} = I + E_{\alpha(1-\varepsilon), \alpha(\varepsilon)}$, and if $(\pi^n, \lambda^n) = \hat{R}^n(\pi, \lambda)$ then $\lambda' = \theta^{*-1}_{\pi,\lambda}(\lambda)$ and (thus) $\lambda^n = \theta^{n\,*-1}_{\pi,\lambda} \cdot \lambda$.

This $\theta_{\pi,\lambda}$ is the R-V cocycle. We'll now give an interpretation of it in terms of the IET $f_{\pi,\lambda}$.

Let $I^n = [0, |\lambda^n|)$ so that IET $f_{\pi^n,\lambda^n} : I^n \circlearrowleft$

Also write $I^n = \bigsqcup_{\alpha \in \mathcal{A}} I^n_\alpha \leftarrow$ the subintervals of continuity of f_{π^n,λ^n}

Note: $f_{\pi^n,\lambda^n} = \left[\text{the first return map of } f_{\pi,\lambda} \text{ to } \underline{\underline{I^n}} \ (\subset I^0)\right]$

This is the def of $\hat{R}$ for $n=1$; for $n \geq 2$ it follows by obvious "induction in steps" property.

Let $r^n = r^n_{\pi,\lambda} : I^n \to \mathbb{Z}^+$ be the <u>first-return time</u>.

(see Problem 32)

Prop 1: $\forall (\pi,\lambda) \in C \times \mathbb{R}_+^{\mathcal{A}}$, $\forall \alpha, \beta \in \mathcal{A}$, $n \geq 1$:

$$\#\{0 \leq j < r^n_{\pi,\lambda}(I^n_\alpha) : f^j_{\pi,\lambda}(I^n_\alpha) \subset I^0_\beta\} = (\Theta^n_{\pi,\lambda})_{\alpha,\beta}$$

To make sense of the above statement, we need:
Lemma: $r^n_{\pi,\lambda}$ is <u>constant</u> on each I^n_α; we write $r^n_{\pi,\lambda}(I^n_\alpha)$ for this constant. Also for any $\alpha \in \mathcal{A}$ and $0 \leq j < r^n_{\pi,\lambda}(I^n_\alpha)$, $f^j(I^n_\alpha)$ is contained in some I^0_β!

Consequence for translation surfaces

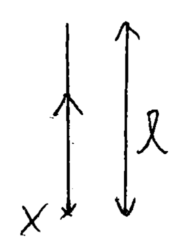
Fix $(\pi, \lambda, \tau) \in \hat{\mathcal{H}}(C)$ {satisfying Keane's condition}

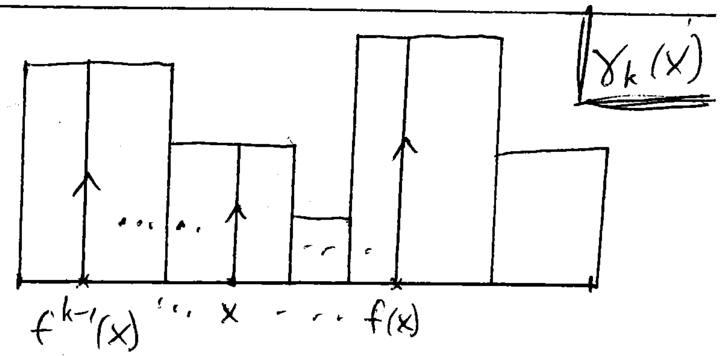
let $M = M(\pi, \lambda, \tau)$.

(cf. Lecture #17, p. 8)

For $x \in M$, $k \in \mathbb{Z}^+$, set $\gamma_k(x) = \gamma\left(x, \sum_{j=0}^{k-1} h_{\alpha(f^j(x))}\right)$

Recall:

$\gamma(x, \ell) =$ 



Corollary 1: $\forall \alpha \in \mathcal{A}$, $x \in I^n_\alpha$, $n \geq 1$: $[\gamma_{r^n(x)}(x)] = \sum_{\beta \in \mathcal{A}} (\Theta^n_{\pi,\lambda})_{\alpha,\beta} \cdot [v_\beta]$

("close up" inside I^0)

Cor. 1 is a first indication of why the Lyapunov exponents which we'll study next can lead to <u>precise asymptotics</u> <u>for $[\gamma(x,\ell)]$</u> as $\ell \to \infty$. (Viana's Thm B.)

Zorich map & cocycle

The "Zorich acceleration" of the R-V induction map

Let ε^j = type of (π^j, λ^j)

$$n = n(\pi,\lambda) = \min\{ j \geq 1 : \varepsilon^j \neq \varepsilon^0 \}$$

$$\hat{Z}(\pi,\lambda) := \hat{R}^n(\pi,\lambda) = (\pi^n, \lambda^n) \qquad \boxed{\hat{Z} : C \times \mathbb{R}_+^{\mathcal{A}} \circlearrowleft}$$

$$Z(\pi,\lambda) := R^n(\pi,\lambda) \qquad \boxed{Z : C \times \Lambda_{\mathcal{A}} \circlearrowleft}$$

$$\hat{\mathcal{Z}}(\pi,\lambda,\tau) = \hat{\mathcal{R}}^n(\pi,\lambda,\tau) \qquad \cancel{\hat{\mathcal{Z}} : \hat{\mathcal{H}}(C) \circlearrowleft}$$

$$\mathcal{Z}(\pi,\lambda,\tau) = \mathcal{R}^n(\pi,\lambda,\tau) \qquad \cancel{\mathcal{Z} : \mathcal{H}(C) \circlearrowleft}$$

<u>Not</u> appropriate domains, since the <u>image</u> $= \{ \text{type}(\pi,\lambda) = \text{type}(\tau) \}$!

<u>Zorich cocycle</u>: $\Gamma_{\pi,\lambda} = \Theta_{\pi,\lambda}^{n(\pi,\lambda)}$

as always, <u>not</u> a power of $\Theta_{\pi,\lambda}$ but the n:th iterate of the R-V cocycle.

$$F_{\hat{Z}} : C \times \mathbb{R}_+^{\mathcal{A}} \times \mathbb{R}^{\mathcal{A}} \circlearrowleft$$

$$F_{\hat{Z}}(\pi,\lambda,\underline{v}) = (\hat{Z}(\pi,\lambda), \Gamma_{\pi,\lambda} \cdot \underline{v}). \quad \text{Similarly: } F_Z, F_{\hat{\mathcal{Z}}}, F_{\mathcal{Z}}.$$

<u>Domains for $\hat{\mathcal{Z}}, \mathcal{Z}$</u>: Set $\hat{\mathcal{Z}}_\varepsilon = \{ (\pi,\lambda,\tau) \in \hat{\mathcal{H}} : \text{type}(\pi,\lambda) = \text{type}(\tau) = \varepsilon \}$

$\hat{\mathcal{Z}}_* = \hat{\mathcal{Z}}_0 \sqcup \hat{\mathcal{Z}}_1$ and $\mathcal{Z}_* = \hat{\mathcal{Z}}_* \cap \mathcal{H}$

Consider $\hat{\mathcal{Z}} : \hat{\mathcal{Z}}_* \circlearrowleft$ and $\mathcal{Z} : \mathcal{Z}_* \circlearrowleft$ These are <u>bijections</u>!

$\underline{Z \text{ preserves } m_1|_{Z_*}}$ since Z is the first-return map of $R : \mathcal{H} \circlearrowleft$ to Z_*.

Cf. Lecture #15, Lemma 2

Hence $\underline{Z \text{ preserves } \mu := P_*(m_1|_{Z_*})}$.

$\boxed{\text{Theorem 1}}$: $\underline{\mu(C \times \Lambda_{\mathcal{A}}) < \infty}$, and $\underline{(C \times \Lambda_{\mathcal{A}}, \mu, Z)}$ is ergodic.

Re-normalize to $\underline{\mu(C \times \Lambda_{\mathcal{A}}) = 1}$.

$\boxed{\text{Theorem 2}}$: $\underline{\log^+ \| \Gamma_{\pi,\lambda}^{\pm 1} \| \in L^1(C \times \Lambda_{\mathcal{A}}, \mu)}$

~~Hence the MET applies, giving that~~

Hence MET applies to the cocycle F_Z, and so there exist Lyapunov exponents $\lambda_1 < \dots < \lambda_k$.

(These are constant on $C \times \Lambda_{\mathcal{A}}$, since $(C \times \Lambda_{\mathcal{A}}, \mu, Z)$ is ergodic.)

These are clearly very important constants, associated to C — or to the connected component $\mathcal{A}_C \subset \mathcal{A}_g(m_1, \dots, m_k)$!

Assume $C \rightsquigarrow g, m_1, \dots, m_k$

$\boxed{\text{Theorem 3}}$: The Lyapunov exponents, with multiplicity, are of the form $\theta_1 \geq \theta_2 \geq \dots \geq \theta_g \geq \underbrace{0 = 0 = \dots = 0}_{k-1} \geq -\theta_g \geq \dots \geq -\theta_1$ for some $\theta_1 \geq \dots \geq \theta_g \geq 0$.

recall $d = 2g + k - 1$

In fact simple; $\theta_1 > \dots > \theta_g > 0$, by Avila–Viana 2007.

Main steps in proof of Thm 3:

Set $H_\pi := \Omega_\pi(\mathbb{R}^{\mathcal{A}}) \subset \mathbb{R}^{\mathcal{A}}$.

Then $\dim H_\pi = 2g$ and Ω_π gives rise to a symplectic form $\omega_\pi : H_\pi \times H_\pi \to \mathbb{R}$

$$\omega_\pi(\Omega_\pi(u), \Omega_\pi(v)) = -u \cdot \Omega_\pi(v).$$

{In fact $H_\pi \cong H^1(M, \mathbb{R})$, and then $\omega_\pi \longleftrightarrow$ the intersection form!}

Now $\Theta_{\pi,\lambda}|H_\pi$ is an isomorphism of symplectic spaces:

$$\Theta_{\pi,\lambda} : \langle H_\pi, \omega_\pi \rangle \xrightarrow{\sim} \langle H_{\pi'}, \omega_{\pi'} \rangle.$$

$\Rightarrow$ $\pm$ symmetry in the Lyapunov spectrum!

{Viana, Prop 2.6.}

Also $H_\pi^\perp = \ker \Omega_\pi$ and $\Theta_{\pi,\lambda}^{*-1}(\ker \Omega_\pi) \subset \ker \Omega_{\pi'}$, and here the action of $\Theta_{\pi,\lambda}$ can be explicitly ("combinatorially") described $\leadsto$ Lyapunov exponents 0.

{Viana, Lemma 5.3}

$\square$

Note: Remark 1 $\Rightarrow$ F_Z has same Lyapunov spectrum!

Also $\pm$ symmetry $\Rightarrow$ F_Z^{*-1} has same —"— !

Theorem 4: The Lyapunov spectrum of the flow ⊛ $(\mathcal{T}^t)$ on $(\mathcal{A}_C, \hat{m}_1)$ (wrt. the derivative cocycle $D\mathcal{T}^t : T(\mathcal{A}_C) \circlearrowleft$) has the form

$$\underbrace{\{\pm 1 \pm \nu_i \;:\; i = 1, \dots, g\}}_{4g} \cup \underbrace{\{1, \dots, 1\}}_{K-1} \cup \underbrace{\{-1, \dots, -1\}}_{K-1}$$

where $\nu_i = \theta_i / \theta_1$.

multisets; we're listing with multiplicity.

⊛ Thus for $\underline{v} \in E_x^i \subset T_x(\mathcal{A}_C)$ corresponding to a Lyapunov exponent λ_i, $\lim_{t \to \pm\infty} \frac{1}{t} \log \|(D\mathcal{T}^t)_x(\underline{v})\| = \lambda_i$

real!

- Discuss $T(\mathcal{A}_C)$; orbifold points....
- Recall $\hat{m}_1$ versus $\tilde{m}_1$ (see Lecture #17, notes for Thm 1); here both work!

Note: Lyapunov exponent 0 has multiplicity 2; this corresponds to an obvious 2-dim subbundle of $T(\mathcal{A}_C)$, namely "$\begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$-directions" (thus: flow direction of $\mathcal{T}^t$, and "trivial scaling direction")

Some steps of the proof:

Use the "a.e. finite-to-one" cover $\hat{S}(C) \to \mathcal{A}_C$

$\Rightarrow$ Suffices to prove corresp. result for $\hat{S}(C)$.

$$\hat{S}(C) = \langle \hat{\mathcal{R}} \rangle \setminus \hat{\mathcal{H}}(C)$$

and $\hat{\mathcal{H}}(C) = \{(\pi, \lambda, \tau) : \pi \in C, \lambda \in \mathbb{R}_+^{\mathcal{A}}, \tau \in T_\pi^+\}$;

locally $\mathbb{R}^{2d}$; hence $T_x(\hat{\mathcal{H}}(C)) = \mathbb{R}^d \times \mathbb{R}^d \quad (\forall x)$

obvious identification

Now study the first return map of $(\mathcal{T}^t)$ to

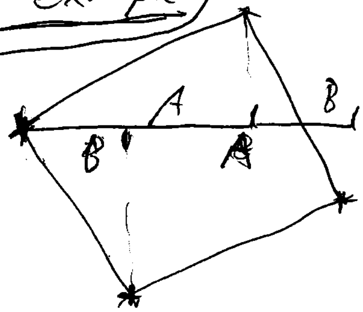
$$Z_* \subset \mathcal{H} \subset \hat{S}(C);$$

one can write out the action of $D\mathcal{T}^t$ completely explicitly, and relate to F_Z, i.e. Thm 3...

$\square$

Special case
"Triv" example

$C = \{\pi\}$ $\quad \pi = \begin{pmatrix} A & B \\ B & A \end{pmatrix}$ $\quad$ degenerate!



$g = 1,\ k = 0$

$m = 0$

only one π

$$\hat{\mathcal{H}} = \{(\lambda, \tau) : \lambda_A, \lambda_B > 0,\ \tau_A > 0 > \tau_B\}$$

$$\mathcal{H} = \{(\lambda, \tau) \in \hat{\mathcal{H}} : \lambda_A + \lambda_B = 1\}$$

$$\rightsquigarrow \quad \{\pm 1 \pm \nu_i\} = \{2, 0, 0, -2\}$$

$\nu_1 = 1$ triv

cf. Lecture 7, p. 8
there "1, 0 −1"
since we considered $\begin{pmatrix} e^{t/2} & \\ & e^{-t/2} \end{pmatrix}$
in place of $\begin{pmatrix} e^{t} & \\ & e^{-t} \end{pmatrix}$.

18.1. **Notes.** .

p. 2: Remark 1 generalizes [50, Remark 2.3].

p. 4: Prop. 1 is [50, Prop. 4.3] and Cor. 1 is [50, Cor. 4.5].

p. 5: Regarding the Zorich map, see [49, Sec. 8 and Sec. 30]. Regarding the Zorich cocycle, see [50, Sec. 4.3].

p. 6: Regarding the claim that $\mathcal{Z}$ preserves $m_{1|Z_*}$ since $\mathcal{Z}$ is the first return map of $\mathcal{R} : \mathcal{H} \to \mathcal{H}$ to Z_*: Note that in fact $m_1(Z_*) < \infty$ (this is equivalent to $\mu(C \times \Lambda_{\mathcal{A}}) < \infty$ in Theorem 1); hence Lemma 2 from Lecture #15 really applies. For Theorem 1, see [49, Prop. 30.2 and Thm. 8.2]. For Theorem 2, see [50, Prop. 4.7]. For Theorem 3, see [50, Prop. 5.1]. Finally, the simplicity statement, $\theta_1 > \cdots > \theta_g > 0$, is (equivalent with) [50, Theorem C]; this is the Zorich-Kontsevich conjecture, which was proved by Avila and Viana in [3].

p. 8: Theorem 4 is [50, Prop. 6.1]; Regarding the remark about Lyapunov exponent 0, cf. [50, Cor. 6.3].

References

1. Jon Aaronson, *An introduction to infinite ergodic theory*, Mathematical Surveys and Monographs, vol. 50, American Mathematical Society, Providence, RI, 1997.
2. William Abikoff, *The real analytic theory of Teichmüller space*, Lecture Notes in Mathematics, vol. 820, Springer, Berlin, 1980.
3. Artur Avila and Marcelo Viana, *Simplicity of Lyapunov spectra: proof of the zorich-kontsevich conjecture*, Acta Math. **198** (2007), no. 1, 1–56. MR 2316268 (2008m:37010)
4. Patrick Billingsley, *Probability and measure*, third ed., Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1995, A Wiley-Interscience Publication.
5. ———, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, 1999, A Wiley-Interscience Publication.
6. Corentin Boissy, *Classification of Rauzy classes in the moduli space of abelian and quadratic differentials*, Discrete Contin. Dyn. Syst. **32** (2012), no. 10, 3433–3457. MR 2945824
7. W. M. Boothby, *An introduction to differentiable manifolds and riemannian geometry*, Academic Press, 1986.
8. Nikolai Chernov and Roberto Markarian, *Chaotic billiards*, Mathematical Surveys and Monographs, vol. 127, American Mathematical Society, Providence, RI, 2006. MR 2229799 (2007f:37050)
9. S. G. Dani, *Invariant measures of horospherical flows on noncompact homogeneous spaces*, Invent. Math. **47** (1978), no. 2, 101–138.
10. ———, *On uniformly distributed orbits of certain horocycle flows*, Ergodic Theory Dynamical Systems **2** (1982), no. 2, 139–158.
11. S. G. Dani and John Smillie, *Uniform distribution of horocycle orbits for Fuchsian groups*, Duke Math. J. **51** (1984), no. 1, 185–194.

12. Manfred Einsiedler and Thomas Ward, *Ergodic theory with a view towards number theory*, Graduate Texts in Mathematics, vol. 259, Springer-Verlag London Ltd., London, 2011.
13. Alex Eskin and Howard Masur, *Asymptotic formulas on flat surfaces*, Ergodic Theory Dynam. Systems **21** (2001), no. 2, 443–478. MR 1827113 (2002g:37049)
14. Gerald B. Folland, *A course in abstract harmonic analysis*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1995. MR 1397028 (98c:43001)
15. ______, *Real analysis*, second ed., Pure and Applied Mathematics (New York), John Wiley & Sons Inc., New York, 1999, Modern techniques and their applications, A Wiley-Interscience Publication. MR 1681462 (2000c:00001)
16. Harry Furstenberg, *The unique ergodicity of the horocycle flow*, Lecture Notes in Math., Vol. 318 (1973), 95–115.
17. A. Gorodnik, *Open problems from the workshop 'emerging applications of measure rigidity'*, skickat okt 2004, 2004.
18. Allen Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge, 2002.
19. Sigurdur Helgason, *Differential geometry, lie groups, and symmetric spaces*, Academic Press, 1978.
20. Olav Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002.
21. Anatole Katok and Boris Hasselblatt, *Introduction to the modern theory of dynamical systems*, Cambridge University Press, Cambridge, 1995, With a supplementary chapter by Katok and Leonardo Mendoza.
22. Alexander S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995. MR 1321597
23. Steven Kerckhoff, Howard Masur, and John Smillie, *Ergodicity of billiard flows and quadratic differentials*, Ann. of Math. (2) **124** (1986), no. 2, 293–311. MR 855297 (88f:58122)
24. Maxim Kontsevich and Anton Zorich, *Connected components of the moduli spaces of Abelian differentials with prescribed singularities*, Invent. Math. **153** (2003), no. 3, 631–678. MR 2000471 (2005b:32030)
25. I. P. Kornfeld, S. V. Fomin, and Ya. G. Sinai, *Ergodic theory*, Springer-Verlag, New York, 1982, Translated from the Russian by A. B. Sosinskii.
26. Erwin Kreyszig, *Introductory functional analysis with applications*, Wiley Classics Library, John Wiley & Sons Inc., New York, 1989.
27. R. Mañé, *Errata to: "A proof of Pesin's formula" [Ergodic Theory Dynamical Systems* **1** *(1981), no. 1, 95–102; MR0627789 (83b:58042)]*, Ergodic Theory Dynam. Systems **3** (1983), no. 1, 159–160. MR 743033 (85f:58064)
28. Ricardo Mañé, *A proof of Pesin's formula*, Ergodic Theory Dynamical Systems **1** (1981), no. 1, 95–102. MR 627789 (83b:58042)
29. A. Manning, *Dynamics of geodesic and horocycle flows on surfaces of constant negative curvature*, Ergodic Theory, Symbolic Dynamics, and Hyperbolic Spaces (T. Bedford, M. Keane, and C. Series, eds.), Oxford University Press, 1991, pp. 71–91.
30. G. A. Margulis, *Discrete subgroups of semisimple lie groups*, Springer-Verlag, 1991.
31. Howard Masur, *The growth rate of trajectories of a quadratic differential*, Ergodic Theory Dynam. Systems **10** (1990), no. 1, 151–176. MR 1053805
32. ______, *Ergodic theory of translation surfaces*, Handbook of dynamical systems. Vol. 1B, Elsevier B. V., Amsterdam, 2006, pp. 527–547. MR 2186247
33. Howard Masur and Serge Tabachnikov, *Rational billiards and flat structures*, Handbook of dynamical systems, Vol. 1A, North-Holland, Amsterdam, 2002, pp. 1015–1089. MR 1928530
34. Dave Witte Morris, *Ratner's theorems on unipotent flows*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 2005.

35. Subhashis Nag, *The complex analytic theory of Teichmüller spaces*, John Wiley & Sons Inc., New York, 1988, A Wiley-Interscience Publication.
36. M. S. Raghunathan, *Discrete subgroups of lie groups*, Springer-Verlag, New York, 1972, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68.
37. Marina Ratner, *On raghunathan's measure conjecture*, Ann. of Math. (2) **134** (1991), no. 3, 545–607.
38. W. Rudin, *Real and complex analysis*, McGraw-Hill, 1987.
39. David Ruelle, *An inequality for the entropy of differentiable maps*, Bol. Soc. Brasil. Mat. **9** (1978), no. 1, 83–87. MR 516310
40. Omri Sarig, *Lecture notes on ergodic theory*, found on www.wisdom.weizmann.ac.il/ sarigo/506/ErgodicNotes.pdf, 2009.
41. P. Sarnak, *Asymptotic behavior of periodic orbits of the horocycle flow and eisenstein series*, Comm. Pure Appl. Math **34** (1981), 719–739.
42. Richard Evan Schwartz, *Obtuse triangular billiards. II. One hundred degrees worth of periodic trajectories*, Experiment. Math. **18** (2009), no. 2, 137–171. MR 2549685
43. Raimond A. Struble, *Metrics in locally compact groups*, Compositio Math. **28** (1974), 217–222. MR 0348037
44. William A. Veech, *Unique ergodicity of horospherical flows*, Amer. J. Math. **99** (1977), no. 4, 827–859.
45. ———, *Gauss measures for transformations on the space of interval exchange maps*, Ann. of Math. (2) **115** (1982), no. 1, 201–242. MR 644019 (83g:28036b)
46. ———, *The Teichmüller geodesic flow*, Ann. of Math. (2) **124** (1986), no. 3, 441–530. MR 866707 (88g:58153)
47. ———, *Moduli spaces of quadratic differentials*, J. Analyse Math. **55** (1990), 117–171. MR 1094714
48. ———, *Siegel measures*, Ann. of Math. (2) **148** (1998), no. 3, 895–944. MR 1670061 (2000k:37028)
49. Marcelo Viana, *Ergodic theory of interval exchange maps*, Rev. Mat. Complut. **19** (2006), no. 1, 7–100. MR 2219821
50. ———, *Lyapunov exponents of Teichmüller flows*, Partially hyperbolic dynamics, laminations, and Teichmüller flow, Fields Inst. Commun., vol. 51, Amer. Math. Soc., Providence, RI, 2007, pp. 139–201. MR 2388695
51. ———, *Lectures on lyapunov exponents*, Cambridge Studies in Advanced Mathematics, vol. 145, Cambridge University Press, Cambridge, 2014. MR 3289050
52. Alex Wright, *Translation surfaces and their orbit closures: An introduction for a broad audience*, EMS Surv. Math. Sci. **2** (2015), no. 1, 63–108. MR 3354955
53. D. Zagier, *Eisenstein series and the riemann zeta function*, Automorphic forms, Representation theory and Arithmetic, 1979, pp. 275–301.