CHAPTER III TRANSFER OPERATORS

III.1 STUDYING DYNAMICAL SYSTEMS WITH DENSITIES

3.1 IN PRACTICE

Q: How do we get μ, if only $T: X \to X$ is known?

Trajectory simulation:

Empirical measure: $\mu_N := \frac{1}{N} \sum_{k=0}^{N-1} \delta_{T^k x} \longrightarrow \mu$

weak convergence towards an ergodic measure (cf. ergodic decomposition & BET)

Q: How fast is this convergence?

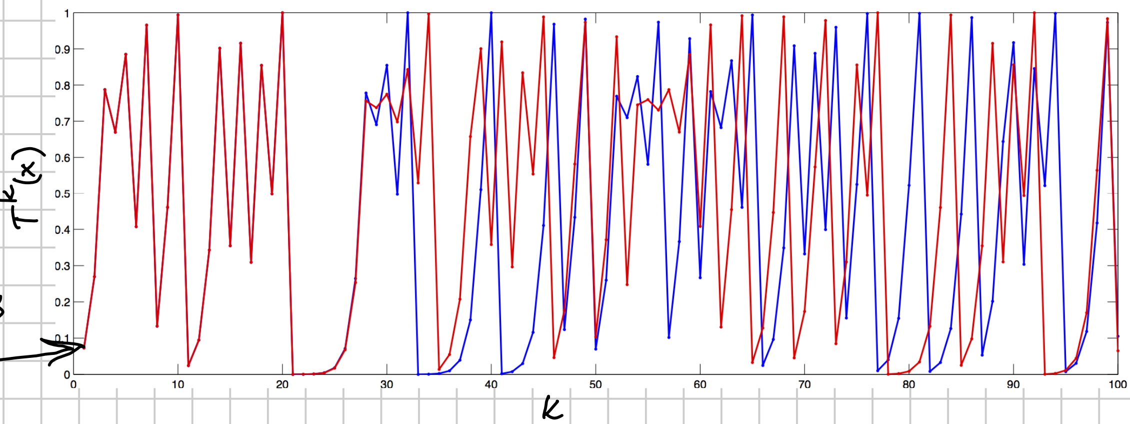
→ What is the dynamical behavior on subdominant time-scales?

3.2 LONG TRAJECTORIES VS DENSITIES

- Round-off errors accumulate ⇝ faith in the results? (cf. angle doubling exercise)
- Large condition number ⇒ close initial conditions quickly diverge ("unpredictability", "chaos")

$T: [0,1] \to [0,1]$

$x \mapsto 4x(1-x)$



Initial conditions 10^{-10} apart

- Gibbs' insight: For chaotic systems it's easier to predict the evolution of densities of a large collection of initial conditions, than individual trajectories.

Recall from exercises: Sensitivity of trajectories not reflected if histograms of different trajectories are plotted.

$\rightsquigarrow$ promising approach: consider dynamics on ensemble of states

$\shortparallel$ many states described by a distribution

$$\underset{\substack{\shortparallel \\ \text{one state with} \\ \text{prob. distribution}}}{\underline{\text{Random variable}}}: \quad Y \longmapsto T(Y)$$

$$\begin{array}{ccc} ? & & ? \\ L^1 \ni f_0 & \longmapsto & f_1 \in L^1 \end{array}$$

(transfer operator (linear, bounded, see below)

Comparison:

1) Non-linear, (possibly) low-dimensional dynamics vs linear ∞-dim. operator

2) One long trajectory vs many (∞) short trajectories

In applications, trajectories and operators are often used hand-in-hand:

(i) Molecular dynamics: $T: X \to X$, where $\dim(X) \gg 1$, hence it is unthinkable to approximate the full operator

$\rightsquigarrow$ Long simulations are used to set up a discretized operator

(ii) Data-driven approaches: Sometimes the dynamics isn't known fully, just through some (few) realizations of the system.

$\rightsquigarrow$ Set-up under uncertainty / smoothness assumptions on densities

3.3 FROBENIUS-PERRON OPERATOR

Def A: $(X, \mathcal{B}, \mu)$ prob. space, $T: X \to X$ measurable map. T is called non-singular if for any $A \in \mathcal{B}$ $\mu(A) = 0$ implies that $\mu(T^{-1}A) = 0$ too; or equivalently: $\mu \circ T^{-1} \ll \mu$

In words: one cannot create mass out of nothing by taking pre-images; or: positive measure cannot disappear under forward iteration.

Example:

1) $T:[0,1]\to[0,1]$, $T(x)=1/2$ is singular wrt Lebesgue
   is non-singular wrt $\delta_{1/2}$

2) $T:[0,1]\to[0,1]$, $T(x)=4x(1-x)$ is non-singular wrt Lebesgue

How does the dynamics propagate (prob.) densities?

A distribution of infinitely many initial conditions $\Longleftrightarrow$ one initial random variable with the same density. We work with the second.

$Y\sim f_0\in L^1(X,\mu)$, i.e. $\mathbb{P}[Y\in A]=\int_A f_0\,d\mu \quad \forall A\in\mathcal{B}$

$T(Y)\sim f_1=?$

$$\mathbb{P}[T(Y)\in A]=\mathbb{P}[Y\in T^{-1}(A)]=\int_{T^{-1}(A)} f_0\,d\mu \overset{!}{=} \int_A f_1\,d\mu$$

Observe: $\nu:\mathcal{B}\to[0,1]$ with $\nu(A)=\int_{T^{-1}(A)} f_0\,d\mu$ is a prob. measure,

and if T is non-singular, then $\nu\ll\mu$

$\xRightarrow[\text{Nikodym}]{\text{Radon-}} \exists_1\ 0\le f_1\in L^1(X,\mu)$ s.t. $\nu(A)=\int_A f_1\,d\mu$

Def B: Let T be a non-singular transformation on a prob. space $(X,\mathcal{B},\mu)$. The Frobenius-Perron operator (FPO) $P:L^1(X,\mu)\to L^1(X,\mu)$ is the unique operator satisfying

$$\int_A Pf\,d\mu=\int_{T^{-1}(A)} f\,d\mu \qquad \forall A\in\mathcal{B},\ f\in L^1.$$

(Note: uniqueness follows from the Radon-Nikodym thm, applied to the positive & negative parts, f^+ and f^-, respectively)

<u>Example</u>: Circle doubling: $T: S^1 \to S^1$, $T(x) = 2x \mod 1$

$I \subseteq S^1$ interval $\Rightarrow$ $T^{-1}(I) = \frac{1}{2}I \,\dot\cup\, \{\frac{1}{2}I + \frac{1}{2}\}$

$$\int_{T^{-1}I} f(x)dx = \int_{\frac{1}{2}I} f(x)dx + \int_{\{\frac{1}{2}I+\frac{1}{2}\}} f(x)dx$$

subst: $x := y/2$ $\qquad$ $x := y/2 + 1/2$

$$= \frac{1}{2}\int_I f(y/2)\,dy + \frac{1}{2}\int_I f(y/2 + 1/2)\,dy$$

$$= \int_I \underbrace{\tfrac{1}{2}\left(f(y/2) + f(y/2+1/2)\right)}_{=Pf} dy \overset{!}{=} \int_I Pf(y)\,dy$$

$$\Rightarrow \qquad \underline{\underline{= Pf}}$$

3.4 PROPERTIES OF THE FPO

<u>PropA</u>: For a non-singular transformation $T: X \to X$, the FPO satisfies

(a) $P(\alpha f + g) = \alpha Pf + Pg \quad \forall f, g \in L^1, \alpha \in \mathbb{C}$ (linearity)

(b) $f \geq 0 \Rightarrow Pf \geq 0$ (positivity)

(c) $\int Pf\, d\mu = \int f\, d\mu$ (integral preserving)

(d) $\|Pf\|_{L^1} \leq \|f\|_{L^1}$ (contraction)

(e) If $S: X \to X$ is another non-sing. map, then

$$P_{T\circ S} = P_T \circ P_S$$

<u>Remark</u>: An operator satisfying (a), (b), (d) is called a Markov operator

<u>Proof</u>:

(a) For $A \in \mathcal{B}$:

$$\int_A P(\alpha f + g)\,d\mu = \int_{T^{-1}A} \alpha f + g\, d\mu = \alpha\int_{T^{-1}A} f\,d\mu + \int_{T^{-1}A} g\,d\mu =$$

$$= \alpha\int_A Pf\,d\mu + \int_A Pg\,d\mu \quad \Rightarrow \text{claim}$$

(b) For $A \in \mathcal{B}$, if $f \geq 0$:

$\int_A Pf\,d\mu = \int_{T^{-1}A} f\,d\mu \geq 0$. Since A arbitrary $\Rightarrow Pf \geq 0$ a.e.

(c) With $A = X$:

$$\int_X Pf\,d\mu = \int_{\underbrace{T^{-1}X}_{=X}} f\,d\mu = \int_X f\,d\mu$$

(d) Let $f = f^+ - f^-$, both $f^+, f^- \geq 0$. Then

$$\|Pf\|_{L^1} = \int |Pf|\,d\mu = \int |Pf^+ - Pf^-|\,d\mu \leq \int |Pf^+| + |Pf^-|\,d\mu$$

$$\overset{(b)}{=} \int Pf^+ + Pf^-\,d\mu \overset{(c)}{=} \int f^+ + f^-\,d\mu = \int |f|\,d\mu = \|f\|_{L^1}$$

(e) $T \circ S$ non-singular: $(T \circ S)^{-1}(A) = S^{-1} \circ T^{-1}(A)$

$$\mu(A) = 0 \underset{T\text{ n-s.}}{\Rightarrow} \mu(T^{-1}A) = 0 \underset{S\text{ n-s.}}{\Rightarrow} \mu(S^{-1} \circ T^{-1}A) = 0$$

$\Rightarrow P_{T \circ S}$ well-defined

For $A \in \mathcal{B}$, $f \in L^1$

$$\int_A P_{T\circ S} f\,d\mu = \int_{S^{-1}\circ T^{-1}(A)} f\,d\mu = \int_{T^{-1}A} P_S f\,d\mu = \int_A P_T(P_S f)\,d\mu \qquad \blacksquare$$

<u>Prop B</u>: If $(X, \mathcal{B}, \mu, T)$ is a ppt, then

$$Pf \circ T = \mathbb{E}(f \mid T^{-1}\mathcal{B}) \quad \text{a.e.}$$

<u>Proof</u>: $Pf \circ T$ is clearly $T^{-1}\mathcal{B}$-measurable. We have for $A = T^{-1}B \in T^{-1}\mathcal{B}$

$$\int_A Pf \circ T\,d\mu = \int_{T^{-1}B} Pf \circ T\,d\mu \underset{\mu\text{ inv.}}{=} \int_B Pf\,d\mu = \int_{T^{-1}B} f\,d\mu$$

$$= \int_A f\,d\mu \qquad \blacksquare$$

<u>Corollary A</u>: If $X \subseteq \mathbb{R}^d$ open and $T: X \to X$ is a Lebesgue-preserving homeomorphism (T invertible and both T and T^{-1} are continuous), then $Pf = f \circ T^{-1}$ a.e.

<u>Proof</u>: T, T^{-1} continuous $\Rightarrow T^{-1}\mathcal{B} = \mathcal{B}$ (Borel) $\Rightarrow \mathbb{E}(f \mid T^{-1}\mathcal{B}) = f$, and Prop B yields the claim $\blacksquare$

Corollary B: If $(X, \mathcal{B}, \mu, T)$ is a ppt, then $\mathcal{P}$ is a contraction on $L^p(X,\mu)$ for every $1 \leq p \leq \infty$.

Proof: For $1 \leq p < \infty$:

μ inv.

$$\|Pf\|_{L^p}^p = \int |Pf|^p \, d\mu = \int |Pf \circ T|^p \, d\mu$$

Prop B

$$= \int \left| \mathbb{E}(f \mid T^{-1}\mathcal{B}) \right|^p d\mu \leq \int \mathbb{E}(|f|^p \mid T^{-1}\mathcal{B}) \, d\mu$$

$x \mapsto |x|^p$ convex on $\mathbb{R}_+$ & Exercise 1 (Sheet 7)

$$= \int |f|^p \, d\mu = \|f\|_{L^p}^p$$

For $p = \infty$

$$\|Pf\|_{L^\infty} = \|Pf \circ T\|_{L^\infty} = \|\mathbb{E}(f \mid T^{-1}\mathcal{B})\|_{L^\infty} \leq \|f\|_{L^\infty}$$

μ inv $\Rightarrow \mu(T(X)) = \mu(X) = 1$ ■ Ex 1 (Sheet 7)

3.5 ERGODICITY AND MIXING

Def: The operator $\mathcal{U} : L^\infty(X,\mu) \to L^\infty(X,\mu)$, $f \mapsto f \circ T$ is called the *Koopman operator* associated with T.

Prop A: The Koopman and Frobenius-Perron operators are adjoint, i.e. $\forall f \in L^1, g \in L^\infty$

$$\int Pf \cdot g \, d\mu = \int f \cdot \mathcal{U}g \, d\mu,$$

we also write $\langle Pf, g \rangle = \langle f, \mathcal{U}g \rangle$

Proof: For $f \in L^1$, and A m'ble:

$$\int f \, \mathcal{U}\chi_A \, d\mu = \int f \cdot \chi_A \circ T \, d\mu = \int f \cdot \chi_{T^{-1}A} \, d\mu = \int_{T^{-1}A} f \, d\mu$$

$$= \int_A Pf \, d\mu = \int Pf \cdot \chi_A \, d\mu$$

Since characteristic functions span L^∞, the claim follows ■

Let $\mathcal{D} = \mathcal{D}(X,\mathcal{B},\mu) = \{ f \in L^1(X,\mathcal{B},\mu) \mid f \geq 0, \|f\|_{L^1} = 1 \}$ denote the set of densities.

Prop B: Let $(X,\mathcal{B},\mu,T)$ be a non-sing. ppt. Then

(a) T is ergodic if and only if

$$\frac{1}{n}\sum_{k=0}^{n-1} \langle P^k f, g\rangle \xrightarrow[n\to\infty]{} \langle \mathbb{1}, g\rangle = \int g\, d\mu \qquad \forall f \in \mathcal{D}, g \in L^\infty$$

(b) T is (strongly) mixing if and only if

$$\langle P^n f, g\rangle \xrightarrow[n\to\infty]{} \langle \mathbb{1}, g\rangle \qquad \forall f \in \mathcal{D}, g \in L^\infty$$

Proof: Recall characterizations 2.2 and 1.15 of ergodicity and mixing, respectively:

$$\text{ergodicity} \iff \frac{1}{n}\sum_{k}^{n-1} \mu(A \cap T^{-k}B) \xrightarrow[n\to\infty]{} \mu(A)\mu(B) \qquad (1)$$

$$\text{mixing} \iff \mu(A \cap T^{-n}B) \xrightarrow[n\to\infty]{} \mu(A)\mu(B) \qquad (2)$$

These are equivalent with

$$\frac{1}{n}\sum_{k}^{n-1} \langle f, U^k g\rangle \xrightarrow[n\to\infty]{} \int f\,d\mu \cdot \int g\, d\mu \qquad (1,1)$$

$$\langle f, U^n g\rangle \xrightarrow[n\to\infty]{} \int f\, d\mu \cdot \int g\, d\mu \qquad (2,1)$$

if one takes $f = \chi_A$, $g = \chi_B$, and hence with (1,1) resp. (2,1) holding $\forall f \in L^1, g \in L^\infty$ (characteristic functions span L^1, L^∞).

Use the adjoint property (Prop A) and restrict to $f \in \mathcal{D}$ to obtain the claim. ∎

III.2 ABSOLUTELY CONTINUOUS INVARIANT MEASURES

3.6 ACIMS AND THE FPO

Recall: $\nu \ll \mu \;:\Leftrightarrow\; \mu(A)=0 \Rightarrow \nu(A)=0$

Prop: Let $(X,\mathcal{B},\mu)$ be a prob. space and $T: X\to X$ non-singular. Let $f^* \in D(X,\mathcal{B},\mu)$, and $\nu = f^*\cdot\mu$ (meaning $\frac{d\nu}{d\mu}=f^*$). Then

$$Pf^* = f^* \iff \nu \text{ is } T\text{-invariant, i.e. } \nu = \nu\circ T^{-1}$$

Proof: For any $A\in\mathcal{B}$:

$$\nu(A) = \nu(T^{-1}A) \iff \int_A f^*\,d\mu = \int_{T^{-1}A} f^*\,d\mu = \int_A Pf^*\,d\mu \quad \blacksquare$$

Why is this interesting?

Recall $T:[0,1]\to[0,1]$, $T(x)=\sqrt{x}$, for which δ_0 is an ergodic measure, but not one that we "observe", since $T^n x \to 1$ for all x, except $x=0$.

What is wrong with that?

In our perception, it is "natural" that significant sets have nonzero Lebesgue measure, and $\{0\}$ doesn't.

Or: δ_0 is not absolutely continuous wrt Lebesgue

If μ represents a measure which we consider to be "natural", and wrt this we find $Pf^*=f^*$, then BET gives (suppose ν ergodic)

$$\frac{1}{n}\sum_{k=0}^{n-1} h\circ T^k(x) \xrightarrow[n\to\infty]{} \int h\,d\nu = \int h f^*\,d\mu$$

for ν-a.e. $x\in X$, which means for a set of x with positive μ-meas.

$\Rightarrow$ we can observe ν!

3.7 PIECEWISE MONOTONIC TRANSFORMATIONS

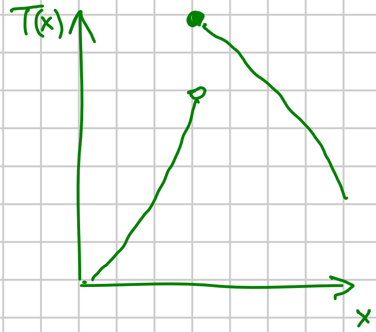
From now on in this section: $I=[a,b]$, $a,b\in\mathbb{R}$.

Def: The map $T: I\longrightarrow I$ is called piecewise monotonic, if there is a partition $a=x_0<x_1<\dots<x_q=b$ of I and some $r\geq 1$ s.t.

1) $T\big|_{(x_{i-1},x_i)}\in C^r$, $i=1,\dots,q$, which can be extended to a C^r function on $[x_{i-1},x_i]$; and

2) $|T'(x)|>0$ on (x_{i-1},x_i)

If, in addition, $|T'(x)|\geq\alpha>1$ $\forall x\in(x_{i-1},x_i)$, $i=1,\dots,q$, T is called expanding.



Prop: $T: I\to I$ pw monotonic, then

$$Pf(x)=\sum_{z\in T^{-1}(x)}\frac{f(z)}{|T'(z)|}$$

Proof: Exercise

3.8 FUNCTIONS OF BOUNDED VARIATION (see [BG])

Def A: For $a=x_0<x_1<x_2<\dots<x_m=b$, $m\geq 1$, we define a partition $\mathcal{P}=\{I_i=[x_{i-1},x_i)\mid i=1,\dots,m\}$ of $[a,b]$. $\{x_0,x_1,\dots,x_m\}$ are the endpoints of $\mathcal{P}$, and one also writes $\mathcal{P}=\mathcal{P}\{x_0,x_1,\dots,x_m\}$

Def B: Let $f: I\to\mathbb{R}$. If there is a $M>0$ s.t.

$$\sum_{i=1}^{m}|f(x_i)-f(x_{i-1})|<M$$

for all partitions $\mathcal{P}=\mathcal{P}\{x_0,x_1,\dots,x_m\}$, then f is said to be of bounded variation (BV). If f is of BV, then

$$V_I f=V_{[a,b]}f=\sup_{\mathcal{P}}\sum_{i=1}^{m}|f(x_i)-f(x_{i-1})|$$

is called the variation of f.

Thm A:

(a) If f is of BV on I, then f is bounded on I; in fact

$$|f(x)| \leq f(a) + V_I f \qquad \forall x \in I.$$

(b) Let f be of BV and s.t. $\|f\|_{L^1} < \infty$. Then

$$|f(x)| \leq V_I f + \frac{\|f\|_{L^1}}{b-a} \qquad \forall x \in I.$$

Thm B:

(a) f, g of BV $\Rightarrow$ $f \pm g$ of BV, and $V_I(f \pm g) \leq V_I f + V_I g$

(b) $V_I(f \cdot g) \leq V_I f \cdot \sup_{x \in I} |g(x)| + V_I g \cdot \sup_{x \in I} |f(x)|$

(c) If $c \in (a,b)$, then $V_{[a,b]} f = V_{[a,c]} f + V_{[c,b]} f$

(d) $f \in C^1(a,b) \cap C^0[a,b]$, then

$$V_{[a,b]} f = \int_a^b |f'(x)|\, dx.$$

Thm C (Helly's selection theorem)

Let $F = \{f\}$ be a family of functions on I. If

- $|f(x)| \leq K \quad \forall f \in F, x \in I$, and
- $V_I f \leq K \quad \forall f \in F$,

then there is a sequence $\{f_n\} \subseteq F$ converging at every point of $[a,b]$ to a function f^* of BV, also $f_n \longrightarrow f^*$ in L^1, and $V_I f^* \leq K$.

Def C: Let

$$BV([a,b]) = \left\{ f \in L^1(I, Leb) \,\middle|\, \inf_{g = f \text{ a.e.}} V_{[a,b]} f < \infty \right\}$$

with the norm

$$\|f\|_{BV} = \|f\|_{L^1} + \inf_{g = f \text{ a.e.}} V_{[a,b]} f$$

(This is a Banach space)

Thm D:

(a) BV is dense in L^1

(b) BV is compact in L^1 (i.e. every bounded sequence in BV has a subsequence which is strongly convergent in L^1)

(c) If $V_I f_n \leq K$ $\forall n$ and $f_n \to f$ in L^1, then $V_I f \leq K$

3.9 A CONTRACTION PROPERTY OF THE FPO

Let $\mathcal{T}(I)$ denote the class of transformations $T: I \to I$ s.t.

(i) T is pw expanding; that is there is a $\mathcal{P} = \{I_1, \dots I_q\}$ with $|T'(x)| \geq \alpha > 1$ $\forall x \in I_i$

(ii) $g := \frac{1}{|T'|} \in BV(I)$

For $n \geq 1$ define $\mathcal{P}^{(n)}$ as the partition "generated by T^n"; i.e.

$$\mathcal{P}^{(n)} := \bigvee_{k=0}^{n-1} T^{-k}(\mathcal{P}) = \{I_{i_0} \cap T^{-1} I_{i_1} \cap \dots \cap T^{-n+1} I_{i_{n-1}} \,|\, I_{i_j} \in \mathcal{P}\}$$

Note: T^n is pw expanding on $\mathcal{P}^{(n)}$

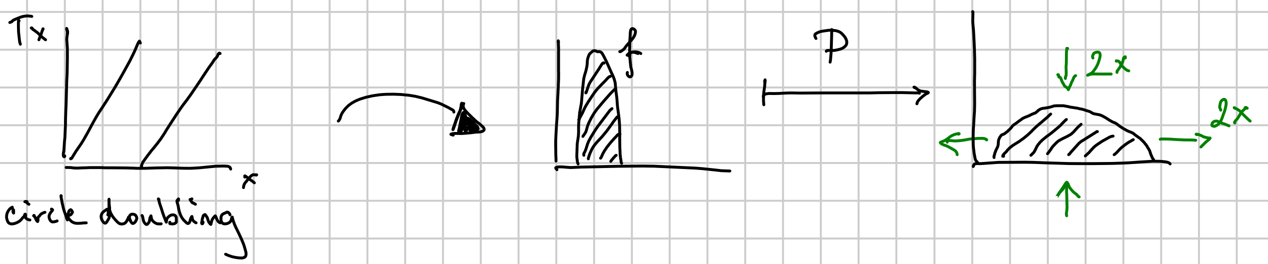
Lemma A: Let $T \in \mathcal{T}(I)$, $\delta := \min_{i=1,\dots,q} m(I_i)$. For any $f \in BV(I)$,

"Lasota–Yorke inequality"

$$V_I Pf \leq A\, V_I f + B \int_I |f|\, dm,$$

where $A = \frac{2}{\alpha} + \max_{i=1,\dots,q} V_{I_i} g$, $\underbrace{B = \frac{2}{\alpha\delta} + \frac{1}{\delta} \max_{i=1,\dots,q} V_{I_i} g}_{= \frac{1}{\delta} A}$, and $\inf |T'| \geq \alpha > 1$.

Idea (proof: [BG, Lemma 5.2.1])



here: $V_I Pf = \frac{1}{2} V_I f$

$\frac{1}{|T'|}$ "

"P is kind of a contraction in the BV-norm"

Lemma B: $T \in \mathcal{J}(I)$, then for $m \geq 1$

$$W_m := \max_{J \in \mathcal{P}^{(m)}} V_J g_m \leq \frac{m}{\alpha^{m-1}} W_1,$$

where $g_m = |(T^m)'|^{-1} = |T'(T^{m-1}) \cdot \ldots \cdot T'(T) \cdot T'|^{-1}$.

Proof: $m = 1$: obvious.

Assume claim holds for some $m \in \mathbb{N}$. Let $J \in \mathcal{P}^{(m)}$, and $y_0 < y_1 < \ldots < y_k \in J$. Then

$$\sum_{j=0}^{k-1} |g_{m+1}(y_{j+1}) - g_{m+1}(y_j)| = \sum_{j=0}^{k-1} |g_m(Ty_{j+1}) g(y_{j+1}) - g_m(Ty_j) g(y_j)|$$

extend with $\pm g_m(Ty_j) g(y_{j+1})$

$$\leq \sum_{j=0}^{k-1} \left|\left(g_m(Ty_{j+1}) - g_m(Ty_j)\right) g(y_{j+1})\right| + \left|g_m(Ty_j)\left(g(y_{j+1}) - g(y_j)\right)\right|$$

$$\leq \sup_{y \in J} g(y) \; W_m + \sup_{y \in J} |g_m(Ty)| \; W_1$$

$$\leq \frac{1}{\alpha} W_m + \frac{1}{\alpha^m} W_1 \underset{\text{induction}}{\leq} \frac{m+1}{\alpha^m} W_1$$

∎

Lemma C: Let $T \in \mathcal{J}(I)$. Then there exist constants $0 < r < 1$, $C, R > 0$ s.t. for any $f \in BV(I)$ and $n \geq 1$

$$\|P^n f\|_{BV} \leq C r^n \|f\|_{BV} + R \|f\|_{L^1}$$

Proof: Let $\alpha_m = (\sup_I g_m)^{-1}$, $W_m = \max_{J \in \mathcal{P}^{(m)}} V_J g_m$, $\delta_m = \min_{J \in \mathcal{P}^{(m)}} m(J)$

$$g_m = \frac{1}{|(T^m)'|} = \frac{1}{|T'(T^{m-1}) \cdot \ldots \cdot T'|} \underset{|T'| \geq \alpha}{\leq} \alpha^{-m} \Rightarrow \alpha_m \geq \alpha^m$$

Lemma B $\Rightarrow$ $W_m \leq m \alpha^{-m+1} W_1$. Since $\alpha > 1$, there is a $k \geq 1$ s.t.

$$r_k := \frac{2}{\alpha_k} + W_k < 1$$

Fix such a k, and let $R_k = \frac{r_k}{\delta_k}$, $C_1 = \max\{r_0, \ldots, r_{k-1}\}$,

$$C_2 = \max\left\{\frac{r_0}{\delta_0}, \dots, \frac{r_{k-1}}{\delta_{k-1}}\right\}.$$

Let $n = jk + i$, $j \geq 0$, $0 \leq i \leq k-1$. Since $P_{T^n} = (P_{T^k})^j P_{T^i}$, we have from Lemma A:

$$\begin{aligned}
V_I P_{T^n} f &= V_I P_{T^k}^j P_{T^i} f = V_I P_{T^k}\left(P_{T^k}^{j-1} P_{T^i} f\right) \\
&\leq r_k V_I P_{T^k}^{j-1} P_{T^i} f + R_k \|f\|_{L^1} \\
&\ \vdots \\
&\leq r_k^j V_I P_{T^i} f + \left(r_k^{j-1} + \dots + r_k + 1\right) R_k \|f\|_{L^1} \\
&\overset{r_k<1}{\leq} C_1 r_k^j V_I f + \left(C_2 + \frac{1}{1-r_k}\right) R_k \|f\|_{L^1}
\end{aligned}$$

Then:

$$\begin{aligned}
\|P_{T^n} f\|_{BV} &= \|P_{T^n} f\|_{L^1} + \inf_{h = P_{T^n} f,\, h \in BV} V_I h \\
&\overset{\text{Lemma A}}{=} \|f\|_{L^1} + \inf_{\bar{f} = f,\, \bar{f} \in BV} V_I P_{T^n} \bar{f} \\
&\leq \|f\|_{L^1} + C_1 r_k^j V_I f + \left(C_2 + \frac{1}{1-r_k}\right) R_k \|f\|_{L^1} \\
&\overset{V_I f \leq \|f\|_{BV}}{\leq} \|f\|_{L^1} + C_1 r_k^j \|f\|_{BV} + \left(C_2 + \frac{1}{1-r_k}\right) R_k \|f\|_{L^1}
\end{aligned}$$

With $r = (r_k)^{1/k}$, $C = C_1 r^{-(k-1)}$, $R = R_k\left(C_2 + \frac{1}{1-r_k}\right) + 1$ we obtain

$$\|P_{T^n} f\|_{BV} \leq C r^n \|f\|_{BV} + R \|f\|_{L^1}$$

■

3.10 EXISTENCE OF ACIMS

Thm: Let $T \in \mathcal{J}(I)$. Then it admits an ACIM whose density is of BV.

Proof: Let $\mathbb{1} = \chi_I$. Then by Lemma 3.9C

$$\|P^n \mathbb{1}\|_{BV} \leq (Cr^n + R)(b-a) \leq (C+R)(b-a)$$

and hence also $\left\{ f_n = \frac{1}{n} \sum_{k=0}^{n-1} P^k \mathbb{1} \right\}$ is a bounded seq. in BV.

Thus, by Helly's thm there is a subseq. $\{f_{m_\ell}\}_{\ell\in\mathbb{N}}$, s.t. $f_{m_\ell} \xrightarrow[\ell\to\infty]{} f^*$ in L^1. Then

$$\|Pf^* - f^*\|_{L^1} \le \underbrace{\|Pf^* - Pf_{m_\ell}\|_{L^1}}_{\to 0\ (\ell\to\infty)} + \|Pf_{m_\ell} - f_{m_\ell}\|_{L^1} + \underbrace{\|f_{m_\ell} - f^*\|_{L^1}}_{\to 0\ (\ell\to\infty)}$$

(P contraction on L^1)

Middle term:

$$\|Pf_{m_\ell} - f_{m_\ell}\|_{L^1} = \frac{1}{m_\ell}\|P^{m_\ell}\mathbb{1} - \mathbb{1}\|_{L^1} \le \frac{2}{m_\ell} \to 0$$

$\Rightarrow Pf^* = f^*$ ■

Remarks:

1) Lemma 3.9C implies quasi-compactness of $P: L^1 \to L^1$

↳ $\exists \rho > 0$ s.t. P has only finitely many eigenvalues $> \rho$, and all those have finite multiplicity ("quasi a matrix")

$\Downarrow$

(i) [GB, 7.2]

$T \in \mathcal{T}(I)$ has finitely many ACIMS $f_1, \dots, f_m \in BV$

(ii) [GB, 8.3]

"strong" ⇒ "weak"

Let $T \in \mathcal{T}(I)$, suppose the ACIM f is unique, and T is (weakly) mixing

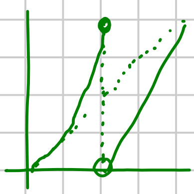
$\Rightarrow \exists D > 0$, $0 < r < 1$ s.t. $\forall g \in BV$

$$\|P^n g - (\textstyle\int_I g\,dm) f\|_{BV} \le D r^n \|g\|_{BV}$$

Exponential convergence! (Compare BET: no statement about convergence speed)

2) No ACIM:

$$T(x) = \begin{cases} \frac{x}{1-x} & 0 \le x \le 1/2 \\ 2x-1 & 1/2 < x \le 1 \end{cases}$$



$T'(0) = 1$!

Conditions of Thm "almost" met

3) Higher dim: much harder problem.

III.3 NUMERICAL APPROXIMATION OF TRANSFER OPERATORS

3.11 ABSTRACT SETTING

Q: How to approximate fixed points of P (or eigenfunctions, in general)?

$P: L^1 \to L^1$: linear operator

$V_n \subseteq L^1$: subspace with $\dim V_n = n \in \mathbb{N}$

$\pi_n: L^1 \to V_n$: linear projection, i.e. $\pi_n \circ \pi_n = \pi_n$

Solve eigenproblem for $P_n: V_n \to V_n$, $P_n = \pi_n \circ P$, instead!

Q: How to choose V_n / its basis?

Examples:

- "hat functions" $\rightsquigarrow$ finite element methods (local support $\Rightarrow$ sparse stiffness mat.)
- polynomials of different order $\rightsquigarrow$ spectral methods (fast convergence for "smooth problems", global support of basis fcns $\Rightarrow$ full matrix)

3.12 ULAM'S METHOD (ULAM 1960)

Partition of X: $\mathcal{P}_n = \{B_1, \dots, B_n\}$, where

usually rectangles, i.e. "boxes"

m = Lebesgue

- $B_i \in \mathcal{B}$
- $m(B_i \cap B_j) = 0$
- $\bigcup_{i=1}^{n} B_i = X$

Let $\chi_i := \chi_{B_i}$, and define projection

$$\pi_n f = \sum_{i=1}^{n} c_i \frac{\chi_i}{m(B_i)}, \quad c_i = \int_{B_i} f \, dm \qquad \Rightarrow V_n = \operatorname{span}\{\chi_1, \dots, \chi_n\}$$

Discretized operator: $P_n := \pi_n P$

Matrix representation w.r.t. basis $\{\chi_1/m(B_1), \dots, \chi_n/m(B_n)\}$ also denoted by P_n:

For $c \in \mathbb{R}^n$ write $c = f \in V_n$ iff $f = \sum_i c_i \frac{\chi_i}{m(B_i)}$

Find $P_m \in \mathbb{R}^{m\times m}$ s.t. $\Pi_m P f = c^T P_m$

$$\Leftrightarrow \quad \Pi_m P\left(\frac{\chi_i}{m(B_i)}\right) = \sum_{j=1}^{m} P_{m,ij} \frac{\chi_j}{m(B_j)} \quad \forall i = 1,\dots,m$$

$$\parallel$$

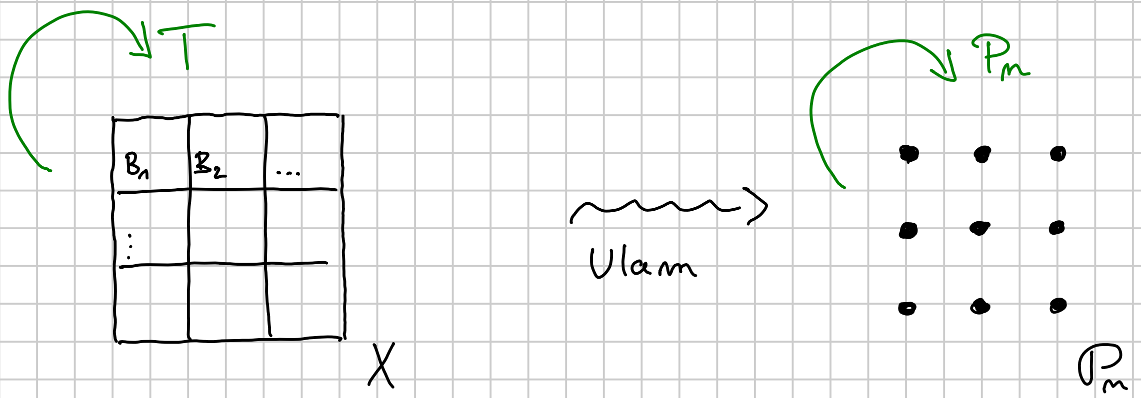
$$\sum_{j=1}^{m} \left(\int_{B_j} P\frac{\chi_i}{m(B_i)}\, dm\right)\frac{\chi_j}{m(B_j)}$$

$$P_{m,ij} = \int_{B_j} P\left(\frac{\chi_i}{m(B_i)}\right) dm = \frac{1}{m(B_i)}\int_{T^{-1}(B_j)} \chi_i \, dm = \boxed{\frac{m(B_i \cap T^{-1}B_j)}{m(B_i)}}$$

3.13 PROBABILISTIC INTERPRETATION

$$P_{m,ij} = \frac{m(B_i \cap T^{-1}B_j)}{m(B_i)} \Rightarrow P_{m,ij} = \mathbb{P}\left[T(x) \in B_j \,\middle|\, x \sim \text{unif}(B_i)\right]$$

"probability that a uniformly sampled state from B_i will jump into B_j"



Markov chain on the partition elements

P_m stochastic matrix $\overset{1.20}{\Rightarrow}$ $\exists c^*$: stationary probability vector

$$\Rightarrow f = \sum_i c_i^* \frac{\chi_i}{m(B_i)} \text{ stationary density of } P_m$$

Prop: P_m always has an invariant density. Even if P doesn't.

3.14 COMPUTATIONAL ASPECTS

Recall $P_{n,ij} = \frac{1}{m(B_i)} \int \underbrace{\chi_{B_i} \cdot \chi_{T^{-1}B_j}}_{} \, dm$

discontinuous functions, high order quadrature doesn't help

$\rightarrow$ Monte-Carlo sampling:

Sample uniformly from B_i : $x_1, x_2, \dots, x_N$

$$\text{Approximate } P_{n,ij} \approx \hat{P}_{n,ij} = \frac{1}{N} \sum_{k=1}^{N} \chi_j(x_k)$$

Properties:

- $\hat{P}_n \to P_n$ as $N^{-1/2}$
- curse of dimension (if $\dim X \gg 1$, it is very expensive (exponentially in $\dim X$) to resolve every coordinate dimension)
- Ulam's method / P_n preserves properties of the Markov operator P: linearity, positivity, integral-preserving (cf. 3.4)
- $\hat{P}_n$ preserves the same properties
- Up to numerical errors: If $P_{n,ij} = 0$, then $\hat{P}_{n,ij} = 0$
- $B_i \cap T^{-1}B_j \neq \emptyset \iff TB_i \cap B_j \neq \emptyset$

  If T Lipschitz continuous, and boxes are "local" (i.e. $\operatorname{diam} B_i$ small), then P_n and $\hat{P}_n$ are sparse
- Computation of $\hat{P}_n$ parallelizable

3.15 CONVERGENCE: LI'S PROOF (1976)

<u>Thm</u>: Let $T: I \to I$ be piecewise C^2, and s.t. $\inf |T'| > 1$. Suppose P has a unique inv. density h. Let $\mathcal{P}_n = \{B_1^{(n)}, \dots, B_n^{(n)}\}$ be a partition of I s.t. $\max_{1 \le i \le n} m(B_i^{(n)}) \to 0$ $(n \to \infty)$, and f_n a fixed density of $P_n^k = \Pi_n \circ P^k$, k suff. large (depends only on T). Then $f_n^* := \frac{1}{k} \sum_{i=0}^{k-1} P^i f_n \xrightarrow{L^1} h$ as $n \to \infty$.

Pf (sketch):

1) $\Pi_n g \to g$ $(n\to\infty)$ for every $g \in L^1$

2) Lasota-Yorke inequality for $k \geq 1$ sufficiently large:
$$V_I P^k g \leq \alpha V_I g + \beta \|g\|_{L^1}, \quad \alpha < 1$$

3) Π_n decreases variation, i.e.
$$V_I \Pi_n g \leq V_I g$$

$\Rightarrow$ 4) The sequence $\{V_I f_n\}$ is bounded

$\Downarrow$ Helly (Thm 3.8C)

5) $\exists$ conv. subseq. $\{f_{n_j}\}$, $f_{n_j} \xrightarrow{L^1} f$

6) $$\|f - P^k f\|_{L^1} \leq \|f - f_{n_j}\|_{L^1} + \underbrace{\|f_{n_j} - P^k_{n_j} f_{n_j}\|_{L^1}}_{=0 \text{ per def.}} + \|P^k_{n_j} f_{n_j} - P^k_{n_j} f\|_{L^1} + \|P^k_{n_j} f - P^k f\|_{L^1}$$

1st and 3rd terms $\to 0$ because $f_{n_j} \to f$, and $\|P^k_{n_j}\|_{L^1} \leq 1$,

4th term $\to 0$ by 1)

$\Rightarrow P^k f = f$

7) $f^* = \frac{1}{k} \sum_{j=0}^{k-1} P^j f$ is fixed point of P, hence $f^* = h$ and $f^*_{n_j} \xrightarrow{L^1} h$ $(j \to \infty)$

8) So far $f_{n_j} \to f$. Assume $\exists$ subsequence of $\{f_n\}$, say $\{f_{n_\ell}\}$, s.t. $|f^*_{n_\ell} - f^*| > \varepsilon$ $\forall n_\ell$ for some $\varepsilon > 0$. Then apply 1)–6) to $\{f_{n_\ell}\}$, and obtain subsequence $f^*_{n_{\ell_q}} \to \tilde{f}^*$ with $P\tilde{f}^* = \tilde{f}^*$, a contradiction. ∎

Remark: $f^*_n = \frac{1}{k} \sum_{i=0}^{k-1} P^i f_n$ is not accessible since we have only finite-dimensional approximations P^j_n $(j \geq 0)$ of the FPO P.

Indeed, one can show that $\frac{1}{k} \sum_{i=0}^{k-1} (\Pi_n P)^i f_n = \frac{1}{k} \sum_{i=0}^{k-1} (P^1_n)^i f_n \to h$

as $n \to \infty$, too.

Cf. Handout #2.

3.16 REMARKS

1) If there is a subsequence $\{f_{m_i}\}$ and a prob. measure μ s.t. $\forall g \in C_b^0(X)$

$$\int f_{m_i} g \xrightarrow[i\to\infty]{} \int g \, d\mu \text{, then } \mu \text{ is an invariant measure}$$

[Froyland 1996]

2) Li's result can be extended to specific multidimensional systems, but rigorous results are scarce

3) P_m^T is a discretization of the Koopman operator:

$$(P_m^T)_{ij} = P_{m,ji} = \frac{m(B_j \cap T^{-1}B_i)}{m(B_j)} = \frac{1}{m(B_j)} \int_{B_j} \chi_{T^{-1}B_i} \, dm$$

$$= \frac{1}{m(B_j)} \int_{B_j} \chi_i \circ T \, dm = \frac{1}{m(B_j)} \int_{B_j} \mathcal{U}\chi_i \, dm = \mathbb{E}\left[\mathcal{U}\chi_i(x) \,\middle|\, x \sim \operatorname{unif}(B_j)\right],$$

or, equivalently, if $\mathbb{R}^n \ni c = f \in L^\infty$, then $P_m c = \mathbb{E}\left[\mathcal{U}f \,\middle|\, \sigma(\mathcal{P}_m)\right]$

4) In the last 20 years Ulam's method (and other discretizations) have been used to approximate further eigenfunctions of the FPO (associated with eigenvals $|\lambda| < 1$)

$\rightsquigarrow$ almost-invariant- / persistent- / metastable sets

Dynamical meaning of $Pf = \lambda f$:

- $\lambda = 1$: dominant behavior for $t \to \infty$ (cf. BET)
- $|\lambda| < 1$: dominant behavior on time-scales $t < \infty$

Applications:
- atmospheric science
- molecular dynamics
- oceanography