Concentration of measure (Batumi 2021)

Define the concentration function of a metric measure space (X, d, μ) to be

μ ← Borel probability measure

$$\alpha_X(\varepsilon) := \sup\left\{ \mu(A_\varepsilon^c) \mid A \subseteq X \text{ Borel},\ \mu(A) \geq \tfrac{1}{2} \right\}.$$

Here, $A_\varepsilon = \{ x \in X \mid d(x, A) \leq \varepsilon \}$ is the ε-neighbourhood of A.

We will consider two examples:

- the sphere $S^n \subseteq \mathbb{R}^{n+1}$ equipped with geodesic distance and Haar measure
  μ = unique isometry-invariant Borel prob measure on S^n
  $\leftrightarrow$ Lebesgue measure of subtended wedges

- the symmetric group S_n equipped with the uniform measure and the Hamming distance
  $$d(\sigma, \tau) = \frac{1}{n} \left| \{ i \mid \sigma(i) \neq \tau(i) \} \right|.$$

These both form normal Lévy families: $\exists$ constants $c_1, c_2 > 0$ s.t.

$$\boxed{\alpha_{X_n}(\varepsilon) \leq c_1 \exp(-c_2 n \varepsilon^2)} \leftarrow \text{concentration inequality}$$

SLOGAN : On high dimensional spaces with concentration, Lipschitz functions are approximately constant. (close to their medians with high probability)

Suppose that $f : X \to \mathbb{R}$ is 1- Lipschitz, i.e.

$$|f(x) - f(y)| \leq d(x,y) \quad \forall\, x,y \in X$$

and M_f is a median, i.e.

$$\mu(\underbrace{f \leq M_f}_{A}) \geq \tfrac{1}{2} \quad \text{and} \quad \mu(\underbrace{f \geq M_f}_{B}) \geq \tfrac{1}{2}.$$

Then $\forall\, \varepsilon > 0$, by defn of $\alpha(\varepsilon)$ and since $f \in \mathrm{Lip}'(X)$,

$$\mu(f > M_f + \varepsilon),\ \mu(f < M_f - \varepsilon) \leq \alpha_{X_n}(\varepsilon)$$

so $$\boxed{\mu(|f - M_f| > \varepsilon) \leq 2c_1 \exp(-c_2 n \varepsilon^2)}.$$

Concentration of measure on S^n will follow from the spherical isoperimetric inequality (based on the problem of minimising the perimeter $\liminf_{\varepsilon \searrow 0} \frac{1}{\varepsilon}(\mu(A_\varepsilon) - \mu(A))$ of a closed curve of given area A)

and concentration on S_n will follow from a martingale inequality.

A consequence of the latter is extreme amenability of the group of measure preserving automorphisms of Lebesgue space (also the unitary groups of certain operator algebras).

§ Spherical isoperimetry

d = geodesic metric, μ = Haar measure on S^n

A cap is a set of the form $C = B_r(n)$, where n is the north pole $(0, \dots, 0, 1)$.

<u>Theorem</u> If $A \subseteq S^n$ is a Borel set and $C \subseteq S^n$ is a cap of equal measure, then
$$\mu(A_\varepsilon) \geq \mu(C_\varepsilon) \quad \forall\ \varepsilon > 0.$$

Proof: Approximating by a sufficiently fine net, wma that A is closed. We will use the fact that the set K of nonempty closed subsets of S^n is compact when equipped with the Hausdorff metric
$$\rho(X, Y) = \inf\{r > 0 \mid X \subseteq Y_r,\ Y \subseteq X_r\}.$$

Given $\varphi \in S^n$ with $\langle \varphi, n \rangle > 0$, let

$E_\varphi = \{x \in S^n \mid \langle \varphi, x \rangle = 0\}$ the φ-equator

$K_\varphi^+ = \{x \in S^n \mid \langle \varphi, x \rangle \geq 0\}$ northern hemisphere

$K_\varphi^- = \{x \in S^n \mid \langle \varphi, x \rangle < 0\}$ southern hemisphere

and let P_φ be reflection in E_φ, explicitly:
$$P_\varphi(x) = x - 2\langle \varphi, x \rangle \varphi.$$

Given such a φ and a closed subset $A \subseteq S^n$, let

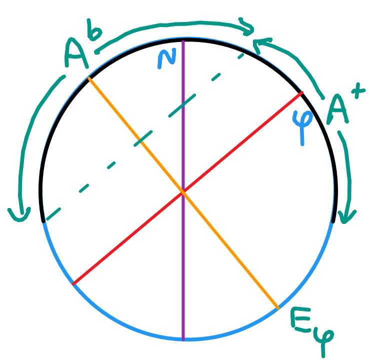
$$A_\varphi^b = \{ x \in A \mid P_\varphi(x) \in A \}$$

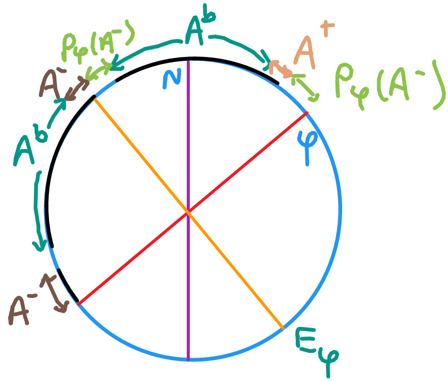
$$A_\varphi^+ = \{ x \in A \cap K_\varphi^+ \mid P_\varphi(x) \notin A \}$$

$$A_\varphi^- = \{ x \in A \cap K_\varphi^- \mid P_\varphi(x) \notin A \}$$

and

$$A_\varphi^* = A_\varphi^b \cup A_\varphi^+ \cup P_\varphi(A_\varphi^-).$$





It is straightforward to check that A_φ^* is a closed subset such that $\mu(A_\varphi^*) = \mu(A)$ and $(A_\varphi^*)_\varepsilon \subseteq (A_\varepsilon)_\varphi^*$, so in particular,

$$\mu\left((A_\varphi^*)_\varepsilon \right) \leq \mu(A_\varepsilon).$$

It is easy to show that
$X := \{ B \in K \mid \mu(B) = \mu(A) \text{ and } \mu(B_\varepsilon) \leq \mu(A_\varepsilon)\ \forall \varepsilon > 0 \}$
is a closed subset of K s.t.
$$B \in X \Rightarrow B^*_\varphi \in X \ \forall \varphi \in S^n \text{ with } \langle \varphi, n \rangle > 0.$$

Let Y be the smallest closed subset of K s.t.
$A \in Y$ and $B \in Y \Rightarrow B^*_\varphi \in Y \ \forall \varphi \in S^n$ with $\langle \varphi, n \rangle > 0$.

Then $Y \subseteq X$, so for every $B \in Y$ and $\varepsilon > 0$,
$$\mu(B) = \mu(A) \text{ and } \mu(B_\varepsilon) \leq \mu(A_\varepsilon).$$

By compactness of Y, there exists $B_0 \in Y$ s.t. $\mu(B \cap C)$ attains its maximum on Y at $B = B_0$.
The Theorem will be proved once we show that $C \subseteq B_0$.

Suppose $C \not\subseteq B_0$. Then $\exists\ x \in S^n$, $\varepsilon > 0$ s.t. $B_\varepsilon(x) \subseteq C \setminus B_0$.
Hence, $\mu(C \setminus B_0) > 0$, so $\mu(B_0 \setminus C) > 0$ too.
(otherwise, $\mu(B_0) = \mu(B_0 \cap C) < \mu(C) = \mu(A) = \mu(B_0)$)
Cover $B_0 \setminus C$ by finitely many balls of radius $\frac{\varepsilon}{3}$ and choose one, say $B_{\frac{\varepsilon}{3}}(y)$, s.t. $\mu\left(B_{\frac{\varepsilon}{3}}(y) \cap B_0 \setminus C\right) > 0$.

Note that $d(x,y) \geq \frac{2\varepsilon}{3}$. (otherwise $B_\varepsilon(x) \cap B_0 \neq \emptyset$)
Let $\varphi = \dfrac{x-y}{\|x-y\|}$, so that $P_\varphi(y) = x$ and (so) $\langle \varphi, n \rangle > 0$.
Then $P_\varphi\left(B_{\frac{\varepsilon}{3}}(y) \cap B_0 \setminus C\right) \subseteq C$, so $\mu\left((B_0)^*_\varphi \cap C\right) > \mu(B_0 \cap C)$. ※ $\square$

Corollary The concentration function on S^n satisfies

$$\alpha_{S^n}(\varepsilon) \leq \sqrt{\frac{\pi}{8}} \exp\left(-\frac{n-1}{2}\varepsilon^2\right).$$

Proof: Integration! $\square$

Other examples of isoperimetry:

- In $\mathbb{R}^n$, equipped with Euclidean distance and Lebesgue measure λ, the isoperimetric problem is solved by the the ball. See: Brunn-Minkowski inequality

Equipped instead with Gaussian measure

$$\gamma_n(A) = (2\pi)^{-\frac{n}{2}} \int_A \exp\left(-\frac{\|x\|^2}{2}\right) d\lambda(x),$$

the problem is solved by the half-space

$$H_s = \{ x \mid x_n \leq s \}.$$

$$\Rightarrow \alpha_{(\mathbb{R}^n, \gamma_n)}(\varepsilon) \leq \frac{1}{\varepsilon\sqrt{2\pi}} \exp\left(-\frac{\varepsilon^2}{2}\right).$$ Gaussian isoperimetry

- On the hypercube $Q_n = \{0,1\}^n$ equipped with uniform measure and Hamming distance $\|x-y\|_1 = |\{i \mid x_i \neq y_i\}|$, the problem is solved by initial segments in the reverse lexicographic order. This is Harper's Theorem.

Therefore, $\alpha_{Q_n}(\varepsilon) \leq \exp\left(-\frac{2}{n}\varepsilon^2\right)$.

To see this, first note that an initial segment I of measure $\geq \frac{1}{2}$ must contain an element $x \in Q_n$ of length $\ell(x) := \|x\|_1 \geq \frac{n}{2}$ (if n is even) or the largest x of length $\frac{n-1}{2}$ (if n is odd).

Otherwise, $|I| < 2^{n-1} = \frac{1}{2}|Q_n|$.

Then, since I is an initial segment, $B_{\lceil \frac{n-1}{2} \rceil}(0) \subseteq I$.

So for any $\varepsilon > 0$, $I_\varepsilon^c \subseteq \{x \in Q_n \mid \ell(x) > \frac{n-1}{2} + \varepsilon\}$.

By isoperimetry,

$$\begin{aligned}
\alpha_{Q_n}(\varepsilon) &= \mu(I_\varepsilon^c) \\
&\leq \mu\left(\ell > \frac{n-1}{2} + \varepsilon\right) \\
&= \text{probability of } > \tfrac{n-1}{2} + \varepsilon \text{ heads when} \\
&\quad \text{a fair coin is tossed } n \text{ times} \\
&= \mathbb{P}(s_n \geq 2\varepsilon + 1) \leq \mathbb{P}(s_n \geq 2\varepsilon),
\end{aligned}$$

where $s_n = \sum_{i=1}^n \varepsilon_i$ is the sum of n i.i.d. Bernoulli r.v.s with $\mathbb{P}(\varepsilon_i = 1) = \mathbb{P}(\varepsilon_i = -1) = \frac{1}{2}$.

Each ε_i is sub-Gaussian: $\mathbb{E}(e^{t\varepsilon_i}) = \cosh t \leq e^{\frac{1}{2}t^2}$, hence so is s_n: $\mathbb{E}(e^{ts_n}) \leq e^{\frac{1}{2}nt^2}$ by independence.

It follows from Markov's inequality with $t = \frac{2\varepsilon}{n}$ that

$$\mathbb{P}(s_n \geq 2\varepsilon) \leq e^{-2\varepsilon t}\,\mathbb{E}(e^{ts_n}) \leq \exp\left(\frac{-2\varepsilon^2}{n}\right).$$

§ The symmetric group

- Martingales

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space, $\mathcal{G}$ a sub-σ-algebra of $\mathcal{F}$ and $f \in L^1(\Omega, \mathcal{F}, \mu)$.

Then $\nu(A) := \int_A f \, d\mu$, $A \in \mathcal{G}$, defines a measure on $\mathcal{G}$ that is absolutely continuous wrt μ.

There is therefore a unique $h \in L^1(\Omega, \mathcal{G}, \mu)$, the Radon-Nikodym derivative $\frac{d\nu}{d\mu}$, s.t.

$$\int_A f \, d\mu = \int_A h \, d\mu \quad \forall A \in \mathcal{G}.$$

The function h is called the conditional expectation of f wrt $\mathcal{G}$, written $h = \mathbb{E}(f \mid \mathcal{G})$.

The operator $f \mapsto \mathbb{E}(f, \mathcal{G})$ is a positive linear map of norm one on all L^p spaces $(1 \leq p \leq \infty)$ s.t.

- $\mathbb{E}(\mathbb{E}(f \mid \mathcal{G}) \mid \mathcal{G}') = \mathbb{E}(f \mid \mathcal{G}') \quad \forall \mathcal{G}' \leq \mathcal{G}$
- $\mathbb{E}(f \cdot g \mid \mathcal{G}) = g \cdot \mathbb{E}(f \mid \mathcal{G}) \quad \forall g \in L^\infty(\Omega, \mathcal{G}, \mu)$
- $\mathbb{E}(f \mid \mathcal{G}) = \mathbb{E} f = \int f \, d\mu$ if $\mathcal{G} = \{\emptyset, \Omega\}$.

A martingale wrt a sequence of σ-algebras $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots \subseteq \mathcal{F}$ is a sequence $f_1, f_2, \dots$ of functions $f_i \in L^1(\Omega, \mathcal{F}_i, \mu)$ s.t. $\mathbb{E}(f_{i+1} | \mathcal{F}_i) = f_i \quad \forall i$.

Special case : Ω finite, μ uniform measure,

- $\{\Omega_i\}_{i=1}^k$ a sequence of partitions s.t. Ω_{i+1} refines $\Omega_i \ \forall i$,
- $\mathcal{F}_i = \sigma$-algebra generated by Ω_i,
- atoms of $\mathcal{F}_i$ are sets of minimal +ve measure

$\mu(A) > 0$ and $B \in \mathcal{F}_i$ s.t. $B \subseteq A$, $\mu(B) < \mu(A) \Rightarrow \mu(B) = 0$

$\rightsquigarrow$ For a function f on Ω, $\mathbb{E}(f, \mathcal{F}_i)$ is the function which is constant on atoms of Ω_i, the constant being the average of f on the atom.

- $\mathcal{F}_i = \mathcal{P}(\Omega) \Rightarrow$ atoms are singletons, $\mathbb{E}(f | \mathcal{F}_i) = f$
- $\mathcal{F}_i = \{\emptyset, \Omega\} \Rightarrow \Omega$ is an atom, $\mathbb{E}(f | \mathcal{F}_i) = \mathbb{E}f = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} f(\omega)$
- $\mathcal{F}_i = \sigma(\Omega_i) \Rightarrow$ atoms are elements of the partition Ω_i

Using the basic properties of the conditional expectation together with the inequality $e^x \leq x + \exp(x^2)$, one proves:

<u>Lemma</u> Let $f \in L^\infty(\Omega, \mathcal{F}, \mu)$ and

$$\{\phi, \Omega\} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_n = \mathcal{F}.$$

Write $d_j = \mathbb{E}(f \mid \mathcal{F}_j) - \mathbb{E}(f \mid \mathcal{F}_{j-1})$ for $1 \leq i \leq n$.

Then for every $c > 0$,

$$\mu(|f - \mathbb{E}f| \geq c) \leq 2\exp\left(-\frac{c^2}{4\sum \|d_j\|_\infty^2}\right).$$

<u>Theorem</u> (Maurey) The family of symmetric groups S_n, with uniform measure and Hamming distance

$$d(\sigma, \tau) = \frac{1}{n} \left|\{i \mid \sigma(i) \neq \tau(i)\}\right|,$$

is a normal Lévy family with constants $c_1 = 2$, $c_2 = \frac{1}{64}$, i.e. $\forall A \subseteq S_n$ $\forall \varepsilon > 0$

$$\mu(A_\varepsilon^c) \leq 2\exp\left(-\frac{n}{64}\varepsilon^2\right).$$

Proof (sketch): For each $j \in \{1, \dots, n\}$, let Ω_j be the partition $\{A_{i_1, \dots, i_j} \mid 1 \leq i_1, \dots, i_j \leq n \text{ distinct}\}$,

where $A_{i_1, \dots, i_j} = \{\pi \in S_n \mid \pi(1) = i_1, \dots, \pi(j) = i_j\}$,

and let $\mathcal{F}_j = \sigma(\Omega_j)$, so

$$\{\phi, S_n\} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_n = \mathcal{P}(S_n).$$

Key fact: For any atom $A = A_{i_1,\dots,i_j} \in \mathcal{F}_j$ and any two atoms
$$B = A_{i_1,\dots,i_j,r}\,,\quad C = A_{i_1,\dots,i_j,s} \in \mathcal{F}_{j+1}$$
contained in A, there is a bijection
$\varphi : B \to C$ s.t. $d(b, \varphi(b)) \le \frac{2}{n}$ $\forall\, b \in B$.

Namely: $\varphi(\pi) = \rho \circ \pi$, where $\rho = (r\ s)$.
$\leadsto$ $\varphi(\pi)(i) = \pi(i)$ for all i except possibly $i = j+1$ and $i = \pi^{-1}(s)$

From this, deduce that for any 1-Lipschitz function f on S_n, the martingale $f_j = \mathbb{E}(f \mid \mathcal{F}_j)$ satisfies $\|d_j\|_\infty \le \frac{2}{n}$ for $j = 1, \dots, n$.

Lemma $\Rightarrow$ $\mu\left(|f - \mathbb{E}f| \ge c\right) \le 2\exp\left(-\frac{n}{16}c^2\right)$ $(*)$

for any such f and any $c > 0$.

This in particular applies to $f = d(\cdot, A)$ $\forall\, A \subseteq S_n$.
Taking $c = 4\left(\frac{1}{n}\log 4\right)^{\frac{1}{2}}$, we get
$$\mu\left(\left| d(\cdot, A) - \mathbb{E}\, d(\cdot, A)\right| < 4\left(\frac{\log 4}{n}\right)^{\frac{1}{2}}\right) > \frac{1}{2}.$$

Fix A with $\mu(A) \ge \frac{1}{2}$. Then $\mu\left(d(\cdot, A) = 0\right) \ge \frac{1}{2}$, so
$\exists\, \pi \in S_n$ s.t. $d(\pi, A) = 0$ and $\left| d(\pi, A) - \mathbb{E}\, d(\cdot, A)\right| < 4\left(\frac{\log 4}{n}\right)^{\frac{1}{2}}$.

That is, $\mathbb{E}\, d(\cdot, A) < 4\left(\frac{\log 4}{n}\right)^{\frac{1}{2}}$.

(∗) $\Rightarrow \mu\left(d(\cdot, A) \geq c + 4\left(\frac{\log 4}{n}\right)^{\frac{1}{2}}\right) \leq 2\exp\left(-\frac{n}{16}c^2\right)$

for any $c > 0$

$\Rightarrow \mu(A_\varepsilon^c) = \mu\left(d(\cdot, A) > \varepsilon\right) \leq 2\exp\left(-\frac{n}{64}\varepsilon^2\right)$

for any $\varepsilon > 8\left(\frac{\log 4}{n}\right)^{\frac{1}{2}}$.

$\left(\text{taking } c = 4\left(\frac{\log 4}{n}\right)^{\frac{1}{2}} < \frac{\varepsilon}{2}\right)$

For $\varepsilon \leq 8\left(\frac{\log 4}{n}\right)^{\frac{1}{2}}$,

$\mu(A_\varepsilon^c) \leq \mu(A^c) \leq \frac{1}{2} \leq 2\exp\left(-\frac{n}{64}\varepsilon^2\right)$,

so the inequality holds for all $\varepsilon > 0$. $\square$

§ Extreme amenability

A topological group G is amenable if
- every continuous affine action of G on a compact convex set has a fixed point

or equivalently if
- there is a left invariant mean on the space $C_r^b(G)$ of bounded right uniformly continuous functions on G.

right uniformity : coarsest compatible with the topology s.t. each $g \mapsto gh$ is uniformly continuous

$m : C_r^b(G) \to \mathbb{C}$ positive, linear, unital,
$m({}^g f) = m(f) \quad \forall f \in C_r^b(G),\ g \in G$
(where ${}^g f(x) = f(g^{-1}x)$)

If G is locally compact, amenability is also equivalent to the existence of an invariant mean on the C^*-algebra $L^\infty(G)$ of measurable functions $G \to \mathbb{C}$ that are essentially bounded wrt Haar measure.

+ many, many more equivalent conditions

Examples
- compact and abelian groups
- the unitary group of an injective von Neumann algebra with the "ultraweak" topology (de la Harpe)

but not
- any G containing $\mathbb{F}_2$ as a closed subgroup.

G is extremely amenable if

- every continuous action of G on a compact space has a fixed point

or equivalently if

- $C^b_r(G)$ has a left invariant multiplicative mean.

Theorem (Veech) Any locally compact group admits a free action on a compact space, so is NOT extremely amenable.

Using concentration of measure, Gromov + Milman proved:

Theorem The unitary group of an infinite dimensional Hilbert space is extremely amenable under the strong operator topology.

SOT : $T_j \to T$ iff $\|T_j x - Tx\| \to 0 \ \forall x \in H$

WOT : $T_j \to T$ iff $\langle T_j x, y \rangle \to \langle Tx, y \rangle \ \forall x, y \in H$

SOT = WOT on $\mathcal{U}(H)$

I will illustrate the idea with the proof of the following.

Theorem (Giordano-Pestov) The group $\mathrm{Aut}(X,\mu)_w$ of all measure-preserving automorphisms of a standard nonatomic (sigma-) finite measure space is extremely amenable under the weak topology.

In the finite case, we may assume up to isomorphism that (X, μ) is $[0,1]$ with Lebesgue measure.

$\mathrm{Aut}(X,\mu)$ is then the group of (equivalence classes of — i.e. up to sets of measure zero) invertible maps $T : [0,1] \to [0,1]$ s.t. $\lambda(T^{-1}E) = \lambda(E)$ $\forall$ measurable E.

Each such T induces a unitary operator

$$U_T : L^2([0,1],\lambda) \to L^2([0,1],\lambda)$$
$$f \mapsto f \circ T^{-1}.$$

The weak topology on $\mathrm{Aut}([0,1],\lambda)$ is the restriction of the WOT (= SOT). Equivalently, $T_j \to T$ iff $\lambda(T_j E \,\Delta\, TE) \to 0$ $\forall$ measurable E.

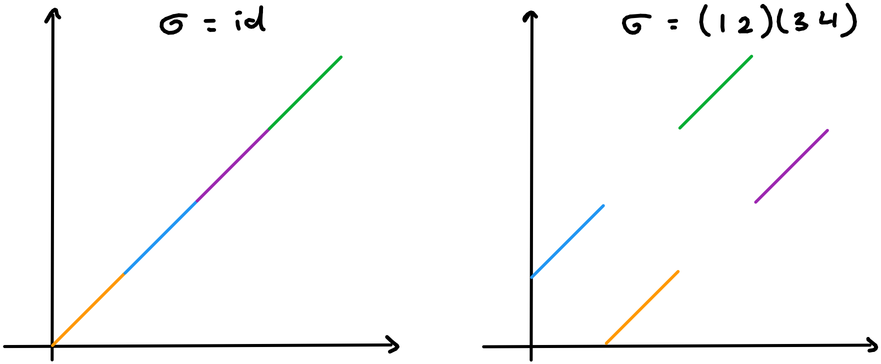
The uniform topology on $\mathrm{Aut}([0,1],\lambda)$ is induced by the left invariant metric

$$d(\sigma,\tau) = \lambda\{x \in [0,1] \mid \sigma(x) \neq \tau(x)\}.$$

It is strictly finer than the weak topology.

For each $n \in \mathbb{N}$, the symmetric group S_{2^n} embeds into $\mathrm{Aut}([0,1],\lambda)$ via interval exchange transformations of the dyadic intervals $\left(\frac{k}{2^n}, \frac{k+1}{2^n}\right)$, $k = 0,1,\dots,2^n-1$.

The above metric d restricts to Hamming distance.



<u>Theorem</u> (Weak approximation)

The increasing union $\bigcup_{n\in\mathbb{N}} S_{2^n}$ is dense in $\mathrm{Aut}([0,1],\lambda)$.

See e.g. Halmos's Lectures on Ergodic Theory.

It suffices to prove extreme amenability of $G = \bigcup_{n\in\mathbb{N}} S_{2^n}$.

Denote Haar measure on S_{2^n} by μ_n.

$C_r^b(G)$ is a commutative unital C^*-algebra whose spectrum is the Samuel compactification of G so its state space is weak* compact.

$\Rightarrow$ wma the states $\varphi_n : C_r^b(G) \to \mathbb{C}$, $\varphi_n(f) = \int_G f \, d\mu_n$ converge weak* to a state φ.

We will show that φ is multiplicative and G-invariant.

By Maurey's Theorem, for every $f \in C_r^b(G)$, $n \in \mathbb{N}$ and $\varepsilon > 0$,

$$\mu_n\left(|f - M_n(f)| > \varepsilon \right) \leq 4 \exp\left(-\frac{2^n}{64} L_f^2 \varepsilon^2 \right),$$

where $M_n(f)$ is a μ_n median of f and where, by uniform continuity, $L_f > 0$ is s.t. f varies by at most ε on entourages of width at most $L_f \varepsilon$.

$$\Rightarrow \int_G |f - M_n(f)| \, d\mu_n \to 0 \quad \forall f \in C_r^b(G).$$

Hence, for every $f, g \in C_r^b(G)$,

$$\begin{aligned} \left| \int fg \, d\mu_n - \int f \, d\mu_n \int g \, d\mu_n \right| \leq & \int |f - M_n(f)| \, |g| \, d\mu_n \\ & + |M_n(f)| \int |g - M_n(g)| \, d\mu_n \\ & + |M_n(g)| \int |M_n(f) - f| \, d\mu_n \\ & + \int |f| \, d\mu_n \int |M_n(g) - g| \, d\mu_n \\ & \to 0 \quad \text{as } n \to \infty, \end{aligned}$$

so $\varphi(fg) = \varphi(f)\varphi(g)$.

Now let $g \in G$, say $g \in S_{2^m}$.

For every $n \geq m$, since μ_n is left G invariant,
$$\int_G {}^g f \, d\mu_n = \int_G f \, d\mu_n .$$

So, $\varphi({}^g f) = \varphi(f)$. $\square$

<u>Remark</u> · (Gromov - Milman)
More <u>generally</u>, every Lévy group
$G = \overline{\bigcup_{i \in I} K_i}$, K_i compact s.t. the Haar
measures on K_i concentrate wrt the right
uniformity, is extremely amenable.

· (Giordano - Pestov)
With the uniform topology, $\mathrm{Aut}([0,1], \lambda)$
is NOT amenable.

§ Anosov diffeomorphisms $\subseteq$ Axiom A systems $\subseteq$ Smale spaces

when irreducible

Let $f: X \to X$ be a measure preserving transformation of a probability space (X, Σ, μ).

The dynamical system $f: X \circlearrowleft$ is measured via observables : functions φ from the state space X to, say, $\mathbb{R}$ that may be required to be Lipschitz/ continuous / integrable.

states $x_0, x_1 = fx_0, \ldots, x_n = fx_{n-1}, \ldots \rightsquigarrow (\varphi(x_n))_{n=0}^{\infty}$

The finite time averages

$$A_n \varphi(x) = \frac{1}{n}\left(\varphi(x) + \varphi(f(x)) + \cdots + \varphi(f^{n-1}x)\right)$$

may be thought of as statistical estimators of the spatial average $\int_X \varphi \, d\mu$.

This is especially true if μ is ergodic wrt f, i.e.

$$E \in \Sigma, \ \mu(f^{-1}E \Delta E) = 0 \Rightarrow \mu(E) = 0 \text{ or } 1.$$

equivalently : every f-invariant (up to measure 0) measurable function $X \to \mathbb{R}$ is a.e. constant

In this case, Birkhoff's ergodic theorem implies that $\lim\limits_{n\to\infty} A_n \varphi(x) \to \int_X \varphi \, d\mu$ almost surely.

Remarks

- If X is compact, there is $G \in \Sigma$ of full measure s.t. for every continuous φ,
$$A_n \varphi(x) \to \int \varphi \, d\mu \quad \forall x \in G$$
i.e. there is a common set of generic points for all continuous observables.

- If μ is not ergodic wrt f, then the convergence is to the conditional expectation of φ wrt the σ-algebra of invariant sets.

Q: How fast does the finite time estimate converge to the spatial average?

Not much can be said without imposing some regularity on f. Convergence for $x \in G$ can be arbitrarily slow.

Let M be a compact Riemannian manifold. A diffeomorphism $f: M \to M$ is called an Anosov diffeomorphism if the tangent space splits into Df-invariant sub-bundles $TM = E^s \oplus E^u$ s.t. $Df|_{E^u}$ is uniformly expanding

and $Df|_{E^s}$ is uniformly contracting.

This means : $(Df)_x E^s(x) = E^s(f(x))$,

$$(Df)_x E^u(x) = E^u(f(x))$$

and there are constants $C > 0$, $0 < \lambda < 1$ s.t. for every $x \in M$ and $n \geq 0$,

$$\| D(f^n)_x \xi \| \leq C \lambda^n \|\xi\| \quad \forall\ \xi \in E^s(x)$$

$$\| D(f^{-n})_x \eta \| \leq C \lambda^{-n} \|\eta\| \quad \forall\ \eta \in E^u(x).$$

The stable subspace $E^s(x)$ is tangent at x to

$$W^s(x) = \{ y \in M \mid d(f^n x, f^n y) \to 0 \text{ as } n \to \infty \}.$$

necessarily exponentially fast

immersed submanifold of M called the stable manifold at x

The unstable manifold $W^u(x)$ is defined similarly.

$n \to -n$

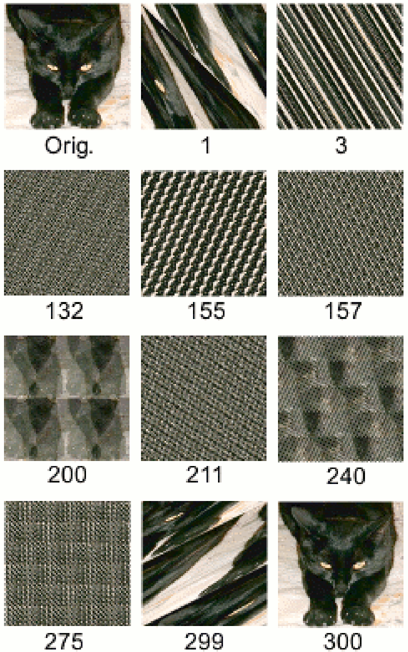
Example Arnold's cat map

M is the torus $\mathbb{R}^2 / \mathbb{Z}^2$ and $f: M \to M$ is defined by $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \bmod 1$.

hyperbolic toral automorphism

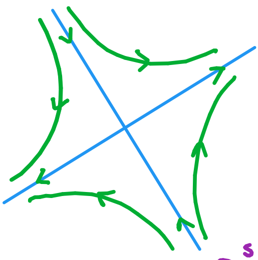
The matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ has eigenvalues $\lambda = \frac{1}{2}(3 + \sqrt{5}) > 1$ and $\frac{1}{\lambda} < 1$.

$(Df)_{\bar{x}} = A$ expands by a factor of λ along the λ-evector $\left(\frac{1}{2}(1+\sqrt{5}), 1\right)$ and contracts by λ along $\left(\frac{1}{2}(1-\sqrt{5}), 1\right)$.



By Claudio Rocchini - Own Work (It's not proper Arnold's cat but my black cat, due copyright restrictions), CC BY 2.5, https://commons.wikimedia.org/w/index.php?curid=1350710

Points in $\mathbb{Q}^2/\mathbb{Z}^2$, i.e. those with rational coordinates, have periodic orbits.



$E^u(\bar{x})$ is tangent to the geodesic through $\bar{x}$ $/\!/ \left(\frac{1}{2}(1+\sqrt{5}), 1\right)$

$E^s(\bar{x})$ is tangent to the geodesic through $\bar{x}$ $/\!/ \left(\frac{1}{2}(1-\sqrt{5}), 1\right)$

$\rightsquigarrow$ every (un)stable manifold is dense

An Anosov diffeomorphism $f: M \to M$ is
irreducible $(\forall \text{ open } \emptyset \neq U, V \subseteq M)(\exists\, n \in \mathbb{N})(f^n U \cap V \neq \emptyset)$

iff f is mixing $(\forall \text{ open } \emptyset \neq U, V \subseteq M)(\exists\, n \in \mathbb{N})(\forall m \geq n)$
$(f^m U \cap V \neq \emptyset)$
iff every $x \in M$ is non-wandering
$(\forall \text{ open } U \subseteq M \text{ containing } x)(\exists n \in \mathbb{N})(f^n U \cap U \neq \emptyset)$
iff every stable manifold is dense in M
iff every unstable manifold is dense in M.

Theorem (Anosov) A C^2 Anosov diffeomorphism that preserves a smooth measure

$$\mu(A) = \int_A \underbrace{q(x)}_{\text{cts}} \, d\underbrace{m}_{\text{Riemannian volume}}(x)$$

is ergodic.

Birkhoff $\Rightarrow$ $\quad \frac{1}{n} \sum_{i=0}^{n-1} \varphi(f^i x) \to \int_M \varphi \, d\mu$ a.s. $\forall$ cts φ

Even if f does not preserve volume, as long as $f: M \circlearrowleft$ is irreducible and C^2, there exists a unique f-invariant measure μ, called the Sinai-Ruelle-Bowen measure of f, s.t.

$$\left\{ x \in M \;\middle|\; A_n \varphi(x) \to \int_M \varphi \, d\mu \;\; \forall \text{ cts observable } \varphi \right\}$$

has full volume.

Using martingale techniques, Chazottes + Goeuzel demonstrate exponential concentration inequalities for the SRB measures of such Anosov systems.

(more generally, for a dynamical system modelled by a uniform Young tower with exponential tails)

Theorem (CG) There is a constant $C > 0$ s.t. $\forall n \in \mathbb{N}$ and every function $K(x_0, \dots, x_{n-1})$ that is 1-Lipschitz in each coordinate,

$$\int_M e^{K(x, fx, \dots, f^{n-1}x) - \int K(y, fy, \dots, f^{n-1}y)\, d\mu(y)}\, d\mu(x) \le e^{Cn}.$$

Corollary For any $\varepsilon > 0$ and any 1-Lipschitz $\varphi : M \to \mathbb{R}$,

$$\mu\left\{x \in M : \left| A_n \varphi(x) - \int_M \varphi\, d\mu \right| > \varepsilon \right\} \le 2 \exp\left(-\frac{n\varepsilon^2}{4C}\right).$$

Proof : Take $K = \varphi(x_0) + \dots + \varphi(x_n)$ and apply Chernoff's bounding trick, i.e. a similar application of Markov's inequality used in the context of sub-Gaussian random variables (p.7). $\square$

Conclusion : For C^2 Anosov diffeomorphisms, the finite time averages of a Lipschitz observable converge in SRB measure to its spatial average exponentially fast.

Batumi 2021: references

Bhishan Jacelon

[BS02] M. Brin and G. Stuck. *Introduction to dynamical systems*. Cambridge University Press, Cambridge, 2002.

[CCS05] J.-R. Chazottes, P. Collet, and B. Schmitt. Devroye inequality for a class of non-uniformly hyperbolic dynamical systems. *Nonlinearity*, 18(5):2323–2340, 2005.

[Cha15] J.-R. Chazottes. Fluctuations of observables in dynamical systems: from limit theorems to concentration inequalities. In *Nonlinear dynamics new directions*, volume 11 of *Nonlinear Syst. Complex.*, pages 47–85. Springer, Cham, 2015.

[GM83] M. Gromov and V. D. Milman. A topological application of the isoperimetric inequality. *Amer. J. Math.*, 105(4):843–854, 1983.

[GP07] T. Giordano and V. Pestov. Some extremely amenable groups related to operator algebras and ergodic theory. *J. Inst. Math. Jussieu*, 6(2):279–315, 2007.

[Hal60] P. R. Halmos. *Lectures on ergodic theory*. Chelsea Publishing Co., New York, 1960.

[Mau79] B. Maurey. Construction de suites symétriques. *C. R. Acad. Sci. Paris Sér. A-B*, 288(14):A679–A681, 1979.

[MS86] V. D. Milman and G. Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.

[Run02] V. Runde. *Lectures on amenability*, volume 1774 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002.

[You98] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math. (2)*, 147(3):585–650, 1998.

[You02] L.-S. Young. What are SRB measures, and which dynamical systems have them? volume 108, pages 733–754. 2002. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays.