Convergence of entropy along geodesics in negatively-curved groups, SMB theorem, and Rokhlin entropy

December 19, 2022

S. G. Dani birthday conference

Amos Nevo

Based on joint work with

Felix Pogorzelski, Leipzig University

Plan of the talk

- Entropy of general group actions : basic definitions.
- Amenable groups : convergence of information functions.
- Entropy theory beyond amenable groups : main obstruction.
- Bowen's f-invariant, and Seward's definition of Rokhlin entropy.
- Seward's inequality and the definition of finitary entropy.
- Negatively curved groups : orbital entropy and an analog of the Shannon-McMillan-Breiman theorem along geodesics.

Shannon entropy and information function

- Let $(X, \mathcal{B}, \mu)$ be a probability space, and let $\mathcal{P}$ be a partition of X to a disjoint countable union of positive measure sets.

- The Shannon entropy of the partition $\mathcal{P}$ is defined by

$$H(\mathcal{P}) = \sum_{A \in \mathcal{P}} -\mu(A) \log \mu(A) \geq 0\,,$$

- The information function of the partition $\mathcal{P}$ is defined by

$$\mathcal{I}_{\mathcal{P}}(x) = -\log \mu(\mathcal{P}(x))$$

  where $\mathcal{P}(x)$ is the unique atom of $\mathcal{P}$ containing x.

- The Shannon entropy of $\mathcal{P}$ is the space average of its information function : $H(\mathcal{P}) = \int_X \mathcal{I}_{\mathcal{P}}(x) d\mu(x)\,,$

Dynamics of partitions in measure-preserving dynamical system

- The join of two partitions $\mathcal{P}$ and $\mathcal{Q}$ consists of the sets of positive measure in

$$\mathcal{P} \vee \mathcal{Q} = \{A \cap B \,;\, A \in \mathcal{P}\,,\, B \in \mathcal{Q}\}\,.$$

  Entropy is join-subadditive: $H(\mathcal{P} \vee \mathcal{Q}) \leq H(\mathcal{P}) + H(\mathcal{Q})\,.$

- Consider a probability measure preserving action of a countable group Γ on (X, μ),
- For $g : X \to X$, $g\mathcal{P} = \{gA\,;\, A \in \mathcal{P}\}$ is the translate of $\mathcal{P}$.
- $\mathcal{P}$ is called a generating partition if the only Γ-invariant σ-algebra containing it is the full σ-algebra $\mathcal{B}(X)$.
- Let $F_n \subset \Gamma$ be a sequence of finite non-empty sets.
- $\mathcal{P}^{F_n} = \bigvee_{g \in F_n} g\mathcal{P}$ are the successive refinements of $\mathcal{P}$

Convergence of normalized Shannon entropies

- Question 1 : Does the sequence of normalized Shannon entropies of the refined partitions $\mathcal{P}^{F_n}$:

$$\frac{1}{|F_n|} H\left(\bigvee_{g\in F_n} g\mathcal{P}\right)$$

converge ?

- Question 2 : Does the sequence of normalized information functions of the refined partitions $\mathcal{P}^{F_n}$

$$\frac{1}{|F_n|}\mathcal{I}_{\mathcal{P}^{F_n}}(x) = -\frac{1}{|F_n|}\log\mu(\mathcal{P}^{F^n}(x))$$

converge pointwise almost surely ?

Entropy theory for amenable groups

- A sequence $F_n \subset \Gamma$ is said to have the property of Asymptotic Invariance under (left) translations if :

$$\lim_{n \to \infty} \frac{|g \cdot F_n \cap F_n|}{|F_n|} = 1 \ , \ \forall g \in \Gamma \, .$$

- A countable group Γ admitting a sequence of sets F_n with this property is an amenable group (Folner, 1957).
  The sets F_n generalize intervals in the group $\mathbb{Z} = \Gamma$.

- Assume Γ is amenable and acts ergodically on a probability space (X, μ), and that F_n are asymptotically invariant.

Convergence of entropy for amenable groups

- Then sequence of normalized Shannon entropies :

$$\frac{1}{|F_n|} H \left(\bigvee_{g \in F_n} g\mathcal{P} \right)$$

converges to a limit $h(\mathcal{P}, \{F_n\})$, called the entropy of the partition.

- The proof is based on a general subadditive convergence theorem for amenable groups, generalizing the corresponding fact for subadditive sequences, namely that $a_{n+m} \leq a_n + a_m$ implies that a_n/n converges.

- Suitable notions of subadditivity were developed and utilized by Kieffer, Emerson 70's, Ornstein-Weiss, Olagnier 80's, with the most general form given by Lindenstrauss-Weiss 2000.

Kolmogorov-Sinai entropy

- The Kolmogorov-Sinai entropy of the action of an amenable groups is defined by

$$h^{KS}(\Gamma, X) = \sup \{h(\mathcal{P}, \{F_n\}) \,;\, \mathcal{P} \text{ a finite partition of } X\}$$

- It is a fundamental fact that this is indeed an isomorphism invariant of the group action and does not depend of the asymptotically invariant sequence chosen to define it.

- Note that the definition of KS-entropy employs the supremum over the finite partitions.

- If $\phi : (X, \mu) \to (X', \mu')$ is a m.p. equivariant factor map, every partition of X' can be lifted to a partition of X', and so

- $h^{KS}(\Gamma, X) \geq h^{KS}(\Gamma, X')$: KS-entropy is monotone.

Entropy equipartition : Shannon-McMillan-Breiman theorem for amenable groups

- Assume the sequence F_n satisfies the stronger asymptotic invariance property of temperedness.
- $\mathcal{P}^{F_n}(x) =$ unique atom of the partition $\mathcal{P}^{F_n}$ containing x
- Shannon-McMillan-Breiman Theorem : The sequence of normalized information functions associated with the refined partitions $\mathcal{P}^{F_n}$ of a finite partition $\mathcal{P}$, namely

$$\mathcal{I}_{\mathcal{P}^{F_n}}(x)/\left|F_n\right| = -\log \mu(\mathcal{P}^{F_n}(x))/\left|F_n\right|$$

  converges for almost every x to the entropy $h(\mathcal{P}, \{F_n\})$ of the partition $\mathcal{P}$.
- Lindenstrauss '01 (extending Kieffer '75, Ornstein-Weiss 83, Olagnier '85. A different proof was given by Weiss '04).
- This establishes asymptotic "entropy equipartition" for ergodic dynamical systems.

Free groups : Ornstein-Weiss example, '87

- Can entropy theory be extended beyond amenable groups? consider the following example.

- Let $X = \mathbb{Z}_2^{\mathbb{F}_2}$ be the Bernoulli action of $\mathbb{F}_2 = \mathbb{F}_2(a, b)$, with μ the uniform product measure.

- X is a compact Abelian group with Haar measure μ. So is $X \times X = (\mathbb{Z}_2 \times \mathbb{Z}_2)^{\mathbb{F}_2}$ with $\mu \times \mu$.

- Define a map $\Phi : X \to X \times X$ by

$$\Phi(x)(g) = (x(g) - x(ga), x(g) - x(gb))$$

- Φ is a continuous and surjective group homomorphism, mapping Haar measure on X to Haar measure on $X \times X$, and its kernel is just the two constant functions !!

Bernoulli factors with higher base entropy

- Furthermore, Φ is equivariant w.r.t. the left $\mathbb{F}_2$-actions on both spaces, so $X \times X$ is a factor of X !!
- any acceptable entropy theory should give the value $\log k$ for a uniform Bernoulli shift with k symbols,
- but here the factor space $X \times X$ has base-space entropy $\log 4$ while the cover X has base-space entropy $\log 2$.
- so even for Bernoulli actions of free groups such an entropy theory for free groups cannot be monotone.

Entropy theory for free groups : Bowen's f-invariant

- Entropy theory was revolutionized in 2009 by Lewis Bowen, who introduced the f-invariant, a remarkable new isomorphism-invariant for ergodic actions of free groups.

- Define for a partition $\mathcal{P}$ of X :

$$F(\mathcal{P}) = -3H(\mathcal{P}) + H(\mathcal{P} \vee a\mathcal{P}) + H(\mathcal{P} \vee b\mathcal{P})$$

  and

$$f(X, \mu, \mathbb{F}_2) = \inf_{R>0} F(\mathcal{P}^{B_R}),$$

  where B_R is a word-metric ball of radius R, and $\mathcal{P}$ is a generating partition under $\mathbb{F}_2$.

- Note that the restriction to generating partitions eliminates the issue of monotonicity, since partitions lifted from a proper factor space cannot be generating.

Beyond amenable groups : Bowen's sofic entropy

- Bowen proved a number of remarkable results for this new invariant, including the fact that it gives the right value for Bernoulli actions and classifies their isomorphism classes.
- Furthermore, he defined the concept of sofic entropy, denoted $h^{sof}(\Gamma, X)$, for actions of sofic groups and established that it classifies their Bernoulli actions, a vast generalization of Ornstein's celebrated theorem.
- The class of countable sofic groups is truly extensive, and may in fact equal the class of countable groups, but this has not been settled yet.

Beyond amenable groups : Rokhlin entropy

- Another major breakthrough was initiated in 2015 by Brandon Seward, as follows.

- A classical result of Rokhlin, proved in 1967, asserts that Kolmogorov-Sinai entropy for $\mathbb{Z}$-actions is equal to :

$$h^{\mathrm{Rok}} = \inf \{ H(\mathcal{P})\,;\, \mathcal{P} \text{ a countable generating partition of } X \} .$$

- Seward and Tucker-Drob '16 proved that this formula still holds for free ergodic actions of general amenable groups.

- It was proposed by Seward as a general definition of an isomorphism-invariant entropy for p.m.p. actions of all countable groups, which Seward termed Rokhlin entropy.

Finitary entropy and Seward's inequality

- Define the finitary entropy of a partition as follows. Let $Fin(\Gamma)$ denote the finite non-empty subsets of Γ, and set

$$h(\mathcal{P}, \{Fin\}) = \inf_{T \in Fin(\Gamma)} \frac{1}{|T|} H\Big(\bigvee_{g \in T} g\mathcal{P} \Big) .$$

  Define the finitary entropy of an action by

$$h^{\text{fin}}(\Gamma, X) := \inf \{ h(\mathcal{P}, \{Fin\}) \, ; \, \mathcal{P} \text{ countable, generating partition} \}$$

- Seward proved the following remarkable inequality between Rokhlin and finitary entropy. For every generating partition $\mathcal{P}$:

$$h^{\text{Rok}}(\Gamma, X) \leq h(\mathcal{P}, \{Fin\}).$$

  It follows that $h^{fin}(\Gamma, X) = h^{\text{Rok}}(\Gamma, X)$.

Negatively curved groups

- Let $M = \Gamma \setminus \widetilde{M}$ be a closed Riemannian manifold of strict negative curvature, Γ its fundamental group, and $\widetilde{M}$ its universal cover,
- Let SM be the unit tangent bundle, $\widetilde{SM}$ its universal cover,
- $g_t : SM \to SM$ and $\widetilde{g}_t : \widetilde{SM} \to \widetilde{SM}$ denote the geodesic flows,
- Let $\partial\widetilde{M}$ denote the boundary of $\widetilde{M}$,
- Fix an origin $o \in \widetilde{M}$, and a Dirichlet fundamental domain $\widetilde{\mathcal{D}} \subset \widetilde{M}$,
- Let ν denote the unique invariant measure of maximal entropy for the geodesic flow on SM,
- Let η denote the Patterson-Sulivan measure on $\partial\widetilde{M}$ associated with the origin $o \in \widetilde{M}$.

Geodesics in $\widetilde{M}$ and almost geodesics in Γ

- Each point $(\widetilde{p}, v) \in \widetilde{SM}$ determines a unique bi-infinite geodesic starting at $\widetilde{p} \in \widetilde{M}$ (with $v \in \partial\widetilde{M}$). Note that $(\widetilde{p}, v)\, ; \, \widetilde{p} \in \widetilde{\mathcal{D}} \cong SM$, the unit tangent bundle.

- For $t \to \infty$, $\widetilde{g}_t\widetilde{p}$ converges to a point in the boundary $\partial\widetilde{M}$. Along the way, the (one-sided) geodesic $\widetilde{g}_t\widetilde{p}$ passes through a sequence of translates $\gamma_k\widetilde{\mathcal{D}}$ of the fundamental domain.

- For $t \geq 0$ and $(\widetilde{p}, v) \in \widetilde{SM}$, define $\gamma = \gamma_{t,(\widetilde{p},v)} \in \Gamma$ via : $\widetilde{g}_t\widetilde{p} \in \gamma^{-1}\widetilde{\mathcal{D}}$. Namely, $\gamma_{n,(\widetilde{p},v)}$ returns the point $\widetilde{g}_n\widetilde{p}$ to the fundamental domain $\widetilde{\mathcal{D}}$.

- For $(\widetilde{p}, v) \in \widetilde{SM}$ define the finite sets :

$$\mathcal{F}_n(\widetilde{p}, v) = \left\{\gamma^{-1}_{1,(\widetilde{p},v)}, \gamma^{-1}_{2,(\widetilde{p},v)}, \dots, \gamma^{-1}_{n,(\widetilde{p},v)}\right\} \subset \Gamma$$

Convergence along almost geodesics

- $\mathcal{F}_n(\widetilde{p}, v)$ is an almost-geodesic segment in Γ starting at e, w.r.t. the metric on Γ inherited from the embedding $\gamma \mapsto \gamma \cdot o \in \widetilde{M}$.

- Let now (X, μ) be a p.m.p. ergodic action of Γ, and $\mathcal{P}$ a countable partition of finite Shannon entropy.

- Consider the refined partitions, well-defined for ν-a.e. $(\widetilde{p}, v) \in \widetilde{\mathcal{D}}$:

$$\mathcal{P}^{\mathcal{F}_n(\widetilde{p},v)} = \bigvee_{\gamma \in \mathcal{F}_n(\widetilde{p},v)} \gamma\mathcal{P}$$

- and the associated information functions :

$$\mathcal{I}_{\mathcal{P}^{\mathcal{F}_n(\widetilde{p},v)}}(x) = -\log \mu(\mathcal{P}^{\mathcal{F}_n(\widetilde{p},v)}(x)) \,.$$

- We can now state the following convergence results, for ν-almost every almost-geodesic.

Shannon-McMillan-Breiman along almost geodesics

- Theorem A.
  the following limit exists for ν-almost every $(\widetilde{p}, v)$, for μ-almost every $x \in X$:
  $$\lim_{n\to\infty} \frac{1}{n+1} \mathcal{J}\left(\mathcal{P}^{F_n(\widetilde{p},v)}\right)(x)$$

- The value of the limit is $\nu \times \mu$-almost surely constant, which we denote $\mathfrak{h}_{\mathcal{P}}(X)$.

- It is given by the conditional entropy
  $$\mathfrak{h}_{\mathcal{P}}(X) = H\left(\mathcal{P} \mid \bigvee_{k=1}^{\infty} \mathcal{P}^{F_k^1(\widetilde{p},v)}\right),$$
  where $F_k^1((\widetilde{p}, v)) = F_k((\widetilde{p}, v)) \setminus \{e\}$.

Convergence along almost geodesics

- Convergence holds also in the $L^1(SM \times X, \nu \times \mu)$ norm. In particular, for ν-almost $(\widetilde{p}, v)$ the partitions refined along the almost-geodesic segments $F_n(\widetilde{p}, v)$ satisfy

$$\lim_{n\to\infty} \frac{1}{n+1} H\left(\mathcal{P}^{F_n(\widetilde{p},v)}\right)$$

$$= \lim_{n\to\infty} \int_{SM} \frac{1}{n+1} H\left(\mathcal{P}^{F_n(\widetilde{p},v)}\right) d\nu(\widetilde{p}, v) = \mathfrak{h}_{\mathcal{P}}(X)$$

- The set $F(\widetilde{p}, v) = \bigcup_{n\geq 0} \mathcal{F}_n(\widetilde{p}, v)$ is one-sided C-almost geodesic in $\Gamma \cdot o \subset \widetilde{M}$. Denote the limit point of the corresponding geodesic in $\partial\widetilde{M}$ by $F^+(\widetilde{p}, v)$.

- Then the law of $F^+(\widetilde{p}, v)$ belongs to the Patterson-Sullivan measure class on the boundary $\partial\widetilde{M}$.

Orbital Rokhlin entropy

- Assume now that the Γ action is free, and also that $\gamma_{k,(\widetilde{p},v)}$, $k \in \mathbb{Z}$ are distinct (possibly adjusting the speed along the geodesic).

- Define the orbital Rokhlin entropy of (X, μ) :

$$\mathfrak{h}^{orb}(\Gamma \curvearrowright X, \mu) = \inf \{\mathfrak{h}_{\mathcal{P}}(X)\, ;\, \mathcal{P} \text{ a generating partition of } X\}$$

- Theorem B.

$$h^{\text{fin}}\big(\Gamma \curvearrowright X\big) = \mathfrak{h}^{orb}(\Gamma \curvearrowright X) = h^{\text{Rok}}\big(\Gamma \curvearrowright X\big).$$

Naturality of the approximation process

- Corollary. For every $\epsilon > 0$ there exists a generating partition $\mathcal{P}_\epsilon$, such that the Rokhlin entropy has the following dynamical approximation ($\nu \times \mu$-almost surely) :

$$h^{\mathrm{Rok}}(\Gamma \curvearrowright X) \leq \lim_{n\to\infty} \frac{1}{n+1} \mathcal{J}\left(\mathcal{P}_\epsilon^{F_n(\widetilde{p},v)}\right)(x) =$$

$$= H\left(\mathcal{P}_\epsilon \mid \bigvee_{k=1}^{\infty} \mathcal{P}_\epsilon^{F_k^1(\widetilde{p},v)}\right) \leq h^{\mathrm{Rok}}(\Gamma \curvearrowright X) + \epsilon\,.$$

- While the value of the limit $\mathfrak{h}_{\mathcal{P}}(X)$ depends on many choices, the corollary implies that the orbital Rokhlin entropy is independent of these choices : it coincides with the isomorphism-invariant given by the Rokhlin entropy.

Method of proof : Theorem A

- Theorem A is a modest contribution to the theory of skew extensions of p.m.p. flows using a measurable group-valued cocycle, which has long tradition in ergodic theory and entropy.
- The convergence of the integrated information function

$$\int_X \int_{SM} \mathcal{I}_{\mathcal{P}^{\mathcal{F}_n(\widetilde{p},v)}}(x)\, d\mu(x)\, d\nu(\widetilde{p}, v)$$

  was established by Abramov-Rokhlin 1966 for skew products and by Ward-Zhang 1992 for amenable groups.

- Pointwise convergence of information functions was established for independent random transformations by Kifer 1986, Ledrappier-Young 1988, and Bogenschutz 1992, and more generally by Morita 1986, and Zhu 2008.

- Note that the "relative" or "fibrewise" entropy of a skew product over the base, is defined by the authors mentioned above as the supremum over the finite partitions $\mathcal{P}$ of X of $\mathfrak{h}_{\mathcal{P}}(X)$. This follows the classical Kolmogorov-Sinai definition of entropy in the case of a skew product.

- Theorem A utilizes the methods developed in the context of skew products and establishes pointwise almost sure convergence for information functions in somewhat greater generality using somewhat sharper maximal inequalities.

Method of proof : Theorem B

- Theorem B is a consequence of Seward's inequality.
- Indeed, suppose that one finds any collection F_n of increasing finite subsets of Γ, such that the entropy of the refined partitions $\mathcal{P}^{\mathcal{F}_n} = \bigvee_{\gamma \in \mathcal{F}_n} \gamma \mathcal{P}$ converges.
- Then Seward's inequality implies that the infimum of entropy over generating partitions is equal to the Rokhlin entropy of the action.
- In general, nothing is known about the existence of a deterministic sequence of finite sets $F_n \subset \Gamma$ with this property when the group is non-amenable.
- So our approach is based on constructing - explicitly and geometrically - a measurable family of finite sets $F_n(y), y \in Y$ which has favourable convergence properties, almost surely.

Some comments

- Similar results apply to all word-hyperbolic groups. The geodesic flow on the unit tangent bundle of a negatively-curved manifold can be replaced by the construction of a measurable geodesic flow for general hyperbolic groups. Shannon-Macmillan-Breiman theorem for convergence along almost geodesic segments in this case relies on some recent results developed by Bader-Furman.

- The variable measurable family of finite sets $\mathcal{F}_n(\widetilde{p}, v)$ in Γ certainly has good ergodic properties, as far as entropy equipartition is concerned. In fact, in every ergodic action of Γ, averaging on these finite sets gives rise to a mean, maximal and pointwise theorem, almost surely. This is a consequence of the theory developed in much greater generality in previous joint work with L. Bowen. Here it is a consequence of the theory of skew product extensions.

- Theorem A has a more general abstract formulation, which applies to all countable groups Γ. The measurable family $\mathcal{F}_n(y)$ of finite sets in the group arises via a virtual homomorphism from a probability-measure-preserving amenable equivalence relation $\mathcal{R}$ on a probability space (Y, ν) into the group Γ.
- Equivalently, $\mathcal{F}_n(y)$ are the images under a measurable cocycle $\alpha : \mathcal{R} \to \Gamma$, which satisfies suitable ergodicity and injectivity properties. They constitute the image under α of the sequence of finite equivalence classes of a hyper-finite exhaustion $\mathcal{R}_n$ of $\mathcal{R}$. This result was established in previous joint work with F. Pogorzelski.