UiO : Department of Mathematics
University of Oslo

Dynamical Entropy

Gaute Schwartz

Master's Thesis, Spring 2020

This master's thesis is submitted under the master's programme *Mathematics*, with programme option *Mathematics*, at the Department of Mathematics, University of Oslo. The scope of the thesis is 60 credits.

The front page depicts a section of the root system of the exceptional Lie group E_8, projected into the plane. Lie groups were invented by the Norwegian mathematician Sophus Lie (1842–1899) to express symmetries in differential equations and today they play a central role in various parts of mathematics.

Abstract

Given a group action $\mathbb{Z} \curvearrowright (X, \mu)$ on a measure space one can associate a numerical quantity to it, its entropy, which measures how chaotic that group action is. In this thesis we will examine various generalizations of such a group action and explore how a notion of entropy can then be defined. First we will replace $\mathbb{Z}$ by a general amenable group, then a sofic group and finally replace (X, μ) by the automorphism group of a C*-algebra. To reassure ourselves that the definitions of entropy are good we will verify various natural properties.

An important recurring example will be the Bernoulli Shift on $(\{1, \ldots, n\}^G, \nu^{\otimes G})$. It is well known that when $G = \mathbb{Z}$ its entropy is $H_\nu(1, \ldots, n)$ and this serves as a benchmark when generalizing entropy; we should expect this result to also hold when G is amenable and even sofic. We should also expect that the entropy of an operator algebraic analouge of the Bernoulli Shift, a shift action on an infinite tensor product $B^{\otimes G}$, somehow only depends on B. This will be proven and much of the theory will be developed for this purpose.

Acknowledgements

First and foremost I want to thank Sergey Neshveyev for great supervision, for giving quick and thorough feedback and carefully walking me through difficult proofs.

I am indebted to the lecturers at UiO in general, for making my courses interesting. A special thanks goes out to Sergey Neshveyev, Nadia Larsen and Tom Lindstrøm for excellent teaching and making me interested in analysis and operator algebra.

I also want to thank my good friend Richard Patrono, for assisting me, on many, many occasions, in the technical matters of LaTex. You usually knew the magic fix when my tex didn't compile.

Finally I want to thank all fellow students whom I've collaborated with over the past few years. You have made my learning process enjoyable and been a great source for ideas and mathematical critique.

Contents

Abstract **i**

Acknowledgements **iii**

Contents **v**

1 Introduction **1**
- 1.1 Introduction 1
- 1.2 Outline of the Thesis 1
- 1.3 Preliminary notions 2

2 Groups and Classical Entropy **5**
- 2.1 Amenable groups 5
- 2.2 Sofic groups 13
- 2.3 Classical amenable entropy 15
- 2.4 Classical sofic entropy 20

3 Dynamical Entropy **29**
- 3.1 Completely positive maps and conditional expectations 29
- 3.2 State entropy 35
- 3.3 Mutual entropy of channels 44
- 3.4 Entropy of dynamical systems 55

4 Examples and an Alternative Definition of Entropy **65**
- 4.1 Non-commutative Bernoulli Shifts 65
- 4.2 Type 1 W*-dynamical systems 68
- 4.3 An alternative definition of entropy via stationary couplings . 72

Bibliography **77**

CHAPTER 1

Introduction

sec:intro

1.1 Introduction

Very roughly speaking the entropy of a group action on a measure space is a non-negative number which measures how chaotic that group action is. It is not obvious how to define such a quantity and to get a reasonable definition we will need to restrict ourselves to groups with certain properties; first we restrict ourselves to amenable groups and later to sofic groups. Amenable groups are quite common, for example any abelian group is amenable. Sofic groups are even more common and though it isn't proven that every group is sofic, not a single example of a non-sofic group is known. The measure spaces on which our group act are assumed to be standard Borel spaces.

It turns out however, that one can extend the notion of entropy to a non-commutative setting, namely where certain groups act on the Automorphism group of a C*-algebra. Here it is even trickier to come up with a reasonable definition and prove useful properties. This is especially so when the group is sofic.

1.2 Outline of the Thesis

In this master thesis we begin with an exploration of group actions on measure spaces in chapter 2. Here we first develop a theory of amenable and sofic groups and then define a notion of entropy for amenable and sofic group actions, respectively, on measure spaces.

In chapter 3 we begin with some general theory about completely positive entropy, then define a notion of entropy for states on a C*-algebras and gradually turn to defining the entropy of a group action $G \to \mathrm{Aut}(A)$ with respect to some state ϕ. This is under the assumption that G is amenable.

In chapter 4 we compute our operator algebraic notion of entropy on a few examples. We also mention an alternative definition of entropy.

1.3 Preliminary notions

There are a few notions that will be used throughout the thesis, so we find it best to discuss them here. In this thesis, G, will always denote a countable, discrete group, i.e. we do not endow G with any topology. A will denote a C*-algebra and M a von Neumann algebra. $S(A)$ will denote the state space of A. Given a compact Hausdorff topological space X, $C(X)$ will denote the space of continuous functions on X, equipped with the supremum norm. When X is just a set, $(X, \mathcal{B})$ will denote a standard Borel space defined as follows:

Definition 1.3.1. A measurable space $(X, \mathcal{B})$ is said to be a standard Borel space if there is a metric d on X such that the metric space (X, d) is complete and separable and its Borel sets coincides with $\mathcal{B}$.

Moreover, $(X, \mathcal{B}, \mu)$ is said to be a *standard probability space* if $(X, \mathcal{B})$ is a standard Borel space and μ is a probability measure thereon. An easy argument shows that μ is automatically regular. We will usually denote the triple $(X, \mathcal{B}, \mu)$ just by (X, μ). When we speak of a partition of (X, μ) we always mean a finite measurable partition. $L^\infty(X/\mathcal{P})$ will then denote the space of functions $X \to \mathbb{C}$ that are constant on members of $\mathcal{P}$. For another partition $\mathcal{C}$ of (X, μ), we write $\mathcal{C} \geq \mathcal{P}$ if any member of $\mathcal{C}$ is contained in a member of $\mathcal{P}$.

Importantly, a standard probability space is unique up to a measure-preserving Borel isomorphism. Hence the unit interval $[0, 1]$ with the usual Borel structure and Borel measure is the archetypical example of an atomless standard probability space. Another example is the Cantor space $\{0, 1\}^{\mathbb{N}}$ equipped with the infinite product σ-algebra and infinite product probability measure.

There is also a recurring real-valued function we will use,

namely $\eta : [0, 1] \to \mathbb{R}$ defined by $\eta(x) = \begin{cases} -x \log(x), & 0 < x \leq 1 \\ 0, & x = 0 \end{cases}$.

Another recurring and fundamental notion in the thesis is the tensor product of two C*-algebras. If A and B are two unital C*-algebras we denote by $A \odot B$ the algebraic tensor product of A and B when these are considered as vector spaces. By definition, $A \odot B$ is the unique vector space with a bilinear map $i : A \times B \to A \odot B$ with the property that for every bilinear map $f : A \times B \to C$, with C a vector space, there is a unique linear map $g : A \odot B \to C$ such that $f = g \circ i$. The image $i((a, b)) \in A \odot B$ is denoted $a \odot b$ and it is evident that the linear span of such elements is all of $A \odot B$. Indeed, taking $C = A \odot B$ and $f = i$ above we see that there couldn't possibly be a *unique* linear map $g : A \odot B \to A \odot B$ such that $i = g \circ i$, unless $\mathrm{span}(i(A \times B)) = A \odot B$.

Now, it turns out that there is a unique involution on $A \odot B$, denoted *, such that $(a \odot b)^* = a^* \odot b^*$ for $a \in A$ and $b \in B$. Similarly there is a unique multiplication on $A \odot B$ such that $(a_1 \odot b_1)(a_2 \odot b_2) = (a_1 a_2) \odot (b_1 b_2)$. These turn $A \odot B$ into a *-algebra. Consider then a sub-multiplicative norm $\| \cdot \|$ satisfying the *-identity and $\|a \odot 1\| = \|a\|_A$ and $\|1 \odot b\| = \|b\|_B$ for $a \in A$ and

$b \in B$. It turns out these always exist and we call them *-norms. We can then complete $(A \odot B, \|\cdot\|)$ and define multiplication and involution by continuity. This is a C*-algebra and it can depend wildly on the norm $\|\cdot\|$ we started with. When A or B is finite dimensional $A \odot B$ equipped with a *-norm already is complete and the uniqueness of the *-norm is immediate from the fact that it is a C*-norm on a C*-algebra. More generally, if A or B is *nuclear*, there is a unique *-norm ([BO08][pp. 104, Theorem 3.8.7]) and thus a unique such C*-algebra completion. In this case we denote it by $A \otimes B$ and, in this space, we denote the elements $a \odot b$ by $a \otimes b$. Nuclear C*-algebras are abound, for example are abelian C*-algebras nuclear.

Since $A \odot B$ lies inside $A \otimes B$ as a dense subspace it enjoys a universal property similar to $A \odot B$. More precisely, the map $i : A \oplus B \to A \odot B \subset A \otimes B$ has the property that for any normed space C and *continuous* linear map $f : A \oplus B \to C$ there is a unique linear map $g : A \otimes B \to C$ such that $f = g \circ i$. In the case where $P : A \to C$ and $Q : B \to C$ are linear continuous maps and $f = P \oplus Q$ we denote the map $g : A \otimes B \to C$ we get by $P \otimes Q$.

Given Hilbert spaces H and K, $H \otimes K$ similarly denotes the Hilbert space completion of $H \odot K$ equipped with the inner product

$$\langle a_1 \odot b_1, a_2 \odot b_2 \rangle := \langle a_1, a_2 \rangle_H \langle b_1, b_2 \rangle_K.$$

Then the map $\Phi : B(H) \otimes B(K) \to B(H \otimes K)$ given by $\Phi(S \otimes T)(h \otimes k) = S(h) \otimes T(k)$, $S \in B(H)$, $T \in B(K)$ and $h \in H$, $k \in K$ gives an isomorphism of C*-algebras.

If a is a normal element of a C*-algebra A then there is a canonical isomorphism of C*-algebras $C(\mathrm{spec}(a)) \to C^*(a)$, where $\mathrm{spec}(a)$ denotes the spectrum of a and $C^*(a)$ denotes the C*-subalgebra generated by a. Similarly, if a is a normal element of a von Neumann algebra M, there is a canonical homomorphism $B(\mathrm{spec}(a)) \to W^*(a)$ where $B(\mathrm{spec}(a))$ denotes the space of bounded Borel functions on $\mathrm{spec}(a)$ equipped with the supremum norm, and $W^*(a)$ the sub-von Neumann algebra generated by a. For the both of this maps, the images of a function $f : \mathrm{spec}(a) \to \mathbb{C}$ will be referred to as $f(a)$.

CHAPTER 2

Groups and Classical Entropy

sec:second

We begin this chapter by defining and proving various results about amenable and sofic groups. We then consider group actions $G \curvearrowright (X, \mu)$ where (X, μ) is a standard probability space, first considering the case where G is amenable, then the case where G is sofic. Under either of these assumptions we can develop a rich theory of entropy to which we will devote the remainder of the chapter.

2.1 Amenable groups

Definition 2.1.1. A group G is said to be amenable if there exists a state, σ, on $\ell^\infty(G)$, with the property that $\sigma(sf) = \sigma(f)$ for all $s \in G$ and $f \in \ell^\infty(G)$. Here, sf denotes the function on G defined by $(sf)(s') = f(s^{-1}s')$, $s' \in G$. We then call σ a left-invariant mean on G.

This definition is convenient for proving things, but a bit opaque. To get a better grasp on it we start with some basic observations about amenable group. We remark that (except from (i)), these do not require that G be countable.

prop:P1

Proposition 2.1.2 ([KL16][pp. 74–75, Proposition 4.2])**.**

(i) *Finite groups are amenable.*

(ii) *Quotients and subgroups of amenable groups are amenable. If $N \subseteq G$ is a normal amenable subgroup and the quotient group G/N is also amenable, then G is amenable.*

(iii) *If G and H are amenable, so is $G \times H$.*

(iv) *If $\{F_i\}_{i \in I}$ is a net of amenable subgroups of G, ordered under inclusion, with $G = \bigcup_{i \in I} F_i$, then G is amenable.*

(v) *If every finitely generated subgroup of G is amenable, then G is.*

Proof.

(i) $\sigma(f) = \frac{1}{|G|} \sum_{s \in G} sf$ defines a left invariant state on $\ell^\infty(G)$.

(ii) See [KL16][pp. 75, Proposition 4.2].

(iii) See [KL16][pp. 75, Proposition 4.2].

(iv) For each $i \in I$, let θ_i be an F_i-invariant state on $\ell^\infty(F_i)$ and define states σ_i on $\ell^\infty(G)$ by letting $\sigma_i = \theta_i(f_{|F_i})$. These form a net $\{\sigma_i\}_{i\in I}$ in the unit ball of $\ell^\infty(G)^*$ which, by compactness, has a weak*-cluster point, say σ. For some subnet $\{i_j\}_{j\in J}$, we have $\sigma(1) = \lim_j \sigma_{i_j}(1) = 1$ and for $f \geq 0$, $\sigma(f) = \lim_j \sigma_{i_j}(f) \geq 0$ so σ is a state. Fixing $s \in G$ we have $s \in F_{i_0}$ for some $i_0 \in I$ by assumption. For $i_j \geq i_0$ we then have $s \in F_{i_j}$, hence $\sigma_{i_j}(sf) = \theta_{i_j}((sf)_{|F_{i_j}}) = \theta_{i_j}(sf_{|F_{i_j}}) = \theta_{i_j}(f_{|F_{i_j}}) = \sigma_{i_j}(f)$, taking limits along J we see that $\sigma(sf) = \sigma(f)$, showing left invariance.

(v) Let I be the net of finite subsets of G ordered under inclusion, put $F_i = \langle i \rangle$ and apply (iv).

■

To show that amenable groups indeed are common, and exhibit concrete examples of them, we need to find other other descriptions that characterize amenability. Consider the following definition.

Definition 2.1.3. We say that a sequence of finite subsets $\{F_n\}_n$ of G is a *left Følner* sequence if for each $s \in G$ we have $\frac{|sF_n \triangle F_n|}{|F_n|} \to 1$. Here $A \triangle B$ denotes the symmetric difference of sets A and B and $|A|$ denotes the number of elements in A.

def:paradoxical **Definition 2.1.4** ([KL16][pp. 75, Definition 4.3])**.** We say that two sets $C, D \subseteq G$ are *equidecomposable*, and write $C \sim D$ if there exists subsets $C_1, \dots, C_n \subseteq C$ and $s_1, \dots, s_n \in G$ such that $C = C_1 \sqcup \dots \sqcup C_n$ and $D = s_1 C_1 \sqcup \dots \sqcup s_n C_n$. Clearly $\sim$ is an equivalence relation. We now say that G is *paradoxical* if there exists disjoint sets C and D such that $C \sim D \sim G$.

We now show that these notions capture the notion of amenability.

prop:P2 **Proposition 2.1.5** ([KL16][pp. 75–76, Theorem 4.4])**.** *For a group G, the following are equivalent:*

(i) *G is amenable,*

(ii) *G is not paradoxical,*

(iii) *G has a left Følner sequence.*

Proof. The implication (i)⇒ (ii) is the easiest: if G were paradoxical take subsets C and D of G as in the definition of paradoxicality and $C_1, \dots, C_n \subseteq C$ and $s_1, \dots, s_n$ such that $C = C_1 \sqcup \dots \sqcup C_n$ and $G = s_1 C_1 \sqcup \dots \sqcup s_n C_n$. For a left-invariant state σ on $\ell^\infty(G)$ we have

$$\sigma(\mathbb{1}_C) = \sigma(\mathbb{1}_{C_1 \sqcup \dots \sqcup C_n}) = \sum_{i=1}^n \sigma(\mathbb{1}_{C_i}) = \sum_{i=1}^n \sigma(\mathbb{1}_{s_i C_i}) = \sigma(\mathbb{1}_{s_1 C_1 \sqcup \dots \sqcup s_n C_n}) = \sigma(\mathbb{1}_G) = 1.$$

Similarly $\sigma(\mathbb{1}_D) = 1$. This is a contradiction since then $2 = \sigma(\mathbb{1}_C + \mathbb{1}_D) \leq \sigma(\mathbb{1}_G) = 1$.

To show (iii)$\Rightarrow$ (i) equip G with the discrete topology and consider its Stone-Cech compactification βG. By definition, the left translation action $G \curvearrowright G$ extends to a continuous action $G \curvearrowright \beta G$. Picking an arbitrary point $x \in \beta G$ and letting $\{F_n\}_n$ be a left Følner sequence for G we consider, for each $n \in \mathbb{N}$, the states $f \mapsto \frac{1}{|F_n|} \sum_{s \in F_n} f(sx)$ on $C(\beta G)$.

By Banach-Alaoglu the set of states on $C(\beta G)$ is weak*-compact, hence this sequence has a weak*-cluster point, say σ. As can be checked σ is then left G-invariant. Since $C(\beta G)$ is G- equivariantly isomorphic to $\ell^\infty(G)$ as a C*-algebra this is what we need.

The final implication, (ii) $\rightarrow$ (i) takes longer to prove so we omit the proof here. ∎

One might wonder why we have phrased the notions of amenability, Følner sequences and paradoxicality in terms of left multiplication and not right multiplication. Luckliy, it turns out this does not matter: if a group has left invariant mean it also has right invariant one, indeed, it even has a mean which is simultaneously *both* left and right invariant (this is not to say that any given left invariant mean will also be right invariant). Similarly, an amenable group has a right Følner sequence and even a sequence that is both simultaneously left and right Følner.

A simple consequence of Proposition 2.1.5 is that $\mathbb{Z}$ is amenable; it is easily checked that the sequence $\{[-n,n]\}_{n=1}^\infty$ is a Følner sequence. From this it follows that *any* abelian group is amenable; by Proposition 2.1.2(v) it suffices to check that any finitely generated subgroup is amenable, but by the well known classification of these ([Fra03][pp. 108–109, Theorem 11.12]), these are of the form $\mathbb{Z}^{r_0} \times (\mathbb{Z}/r_1\mathbb{Z}) \times \ldots \times (\mathbb{Z}/r_n\mathbb{Z})$ for some $r_0, \ldots r_n \in \mathbb{N}$. Since these are direct products of $\mathbb{Z}$ and finite groups, they are amenable.

There are plenty of groups that aren't amenable, however. An important example is the free group F_2 on 2 generators. To see this, let $\{a,b\}$ be the standard generating set and for $s \in \{a,b,a^{-1},b^{-1}\}$ let V_s denote the set of reduced word beginning with s. Then we see that $F_2 = V_{a^{-1}} \sqcup (F_2 \backslash V_{a^{-1}}) \sim V_{a^{-1}} \sqcup a(F_2 \backslash V_{a^{-1}}) = V_{a^{-1}} \sqcup V_a$, where $\sim$ is as in Definition 2.1.4. Similarly $F_2 \sim V_b \sqcup V_{b^{-1}}$. Now, the sets $V_a \sqcup V_{a^{-1}}$ and $V_b \sqcup V_{b^{-1}}$ are disjoint and F_2 is $\sim$-equivalent to both. This shows that F_2 is paradoxical, and thus not amenable. More generally the free group on n generators is not amenable for any $n \geq 2$. If we want to study group actions of such non-amenable groups we need a weaker notion than amenability. This is why we later introduce the notion of *sofic* groups. At first we give an important about Følner sequences that we will later use.

thm: subadditivity

Theorem 2.1.6 ([KL16][pp. 95, Theorem 4.38]). *Suppose ϕ is a non-negative real valued function on the set of finite subsets of G satisfying :*

(i) $\phi(As) = \phi(A)$ *(right invariance).*

(ii) $\phi(A \cup B) \leq \phi(A) + \phi(B)$ *(subadditivity).*

for all finite subsets $A, B \subseteq G$ and $s \in G$. Then $\frac{\phi(F_n)}{|F_n|}$ converges for any Følner sequence $\{F_n\}_n$ of G and to the same value independent of the choice of Følner sequence.

The proof of this is quite technical, but we will include it since this is such a fundamental result for the theory of entropy. To begin we introduce the following.

def:invariant

Definition 2.1.7 ([KL16][pp. 92, Definition 4.32]). If $F \subseteq G$ is finite and $\epsilon > 0$ we say that a finite set $A \subset G$ is (F, ϵ)–invariant if $|\{s \in A : Fs \subseteq A\}| \geq (1 - \epsilon)|A|$.

We see that this definition captures the notion of a Følner sequence. Indeed, if $\{F_n\}_n$ is Følner, fix $\epsilon > 0$ and a finite set $F \subset G$. We see that

$$\{s \in F_n : Fs \in F_n\} \supset \bigcap_{t \in F} t^{-1}F_n \cap F_n \supset \bigcap_{t \in F} F_n \backslash (t^{-1}F_n \vartriangle F_n).$$

Applying the Følner condition to the elements t^{-1} for $t \in F$ we can choose sufficiently large n such that each of the sets in the last intersection have at least $(1 - \frac{\epsilon}{|F|})|F_n|$ elements. Then the intersection itself has more than $(1 - \epsilon)|F_n|$ elements and in particular $|\{s \in F_n : Fs \in F_n\}| \geq (1 - \epsilon)|F_n|$ showing that F_n eventually becomes (F, ϵ)–invariant.

Conversely, if $\{F_n\}_n$ is a sequence of finite sets in G that for any finite set $F \subset G$ and $\epsilon > 0$ eventually become (F, ϵ)–invariant, then $\{F_n\}_n$ is Følner. Indeed, fixing $t \in G$ and letting $F = \{t\} \cup \{t^{-1}\}$ we obtain that F_n eventually becomes $(F, \epsilon/2)$–invariant.

We have

$$tF_n \cap F_n^C \subset \{f \in F_n : Fs \notin F_n\}$$

and similarly

$$t^{-1}F_n \cap F_n^C \subset \{f \in F_n : Fs \notin F_n\}.$$

By $(F, \epsilon/2)$–invariance we get

$$|tF_n \cap F_n^C|, |t^{-1}F_n \cap F_n^C| < \frac{\epsilon}{2}|F_n|$$

. But

$$|t^{-1}F_n \cap F_n^C| = |t(t^{-1}F_n \cap F_n^C)| = |F_n \cap tF_n^C| = |F_n \cap (tF_n)^C| < \frac{\epsilon}{2}|F_n|$$

also, so $|tF_n \vartriangle F_n| = |tF_n \cap F_n^C| + |F_n \cap (tF_n)^C| < \epsilon|F_n|$, showing that $\{F_n\}_n$ is Følner.

Another way of describing approximate invariance is the following.

Definition 2.1.8 ([KL16][pp. 93, Definition 4.34])**.** For finite sets $F, A \subseteq G$ we define the F *boundary* of A, denoted $\partial_F A$, to be the set $\{s \in A : Fs \cap A \neq \varnothing \text{ and } Fs \cap A^C \neq \varnothing\}$. This coincides with the set $\bigcap\limits_{s \in F} s^{-1}A \setminus \bigcup\limits_{s \in F} s^{-1}A$.

The above definition also captures Følnerness in the sense that a finite sequence of sets $\{F_n\}_n$ is Følner if and only if $\frac{|\partial_F F_n|}{|F_n|} \to 0$ for every finite set $F \subset G$. We also need the following notions:

Definition 2.1.9 ([KL16][pp. 91, Definition 4.29])**.** Let $F \subset G$ be finite and $\lambda, \epsilon \geq 0$. We say that a collection $\{F_i\}_{i \in I}$ of subsets of G

(i) $\lambda-$*covers* F if $|\bigcup\limits_{i \in I} F_i| \geq \lambda |F|$,

(ii) is a $\lambda-$*even covering* of F if there exists a positive integer M such that $\sum\limits_{i \in I} \mathbb{1}_{F_i} \leq M \mathbb{1}_F$ and $\sum\limits_{i \in I} |F_i| \geq \lambda M |F|$. M is called the *multiplicity* of the covering.

(iii) is $\epsilon-$*disjoint* if there exists pairwise disjoint sets $\hat{F}_i \subset F_i$ such that $|\hat{F}_i| \geq (1 - \epsilon)|F|$ for all $i \in I$.

It is intuitive that for a sufficiently even covering most of the sets are forced to be $\epsilon-$disjoint for small ϵ. More precisely, we have the following:

lem:L1

Lemma 2.1.10 ([KL16][pp. 92, Lemma 4.31])**.** *For $0 \leq \epsilon \leq 1/2$ and $0 < \lambda \leq 1$ any $\lambda-$even covering $\{F_i\}_{i \in I}$ of F admits an $\epsilon-$disjoint subcovering $\{F_i\}_{i \in J}$ that $\epsilon\lambda-$covers F.*

There is also a connection between approximate invariance and coverings:

lem:L2

Lemma 2.1.11 ([KL16][pp. 93, Lemma 4.33])**.** *Let $\epsilon > 0$ and let $F, A \subset G$ be nonempty finite sets such that A is $(F, \epsilon)-$invariant. Then the collection of right translates $\{Fs : s \in A,\ Fs \subset A\}$ is a $(1 - \epsilon)-$even covering of A with multiplicity $|F|$.*

Proof. The set I of all $s \in A$ witnessing $Fs \subset A$ has, by $(F, \epsilon)-$invariance, at least $(1 - \epsilon)|A|$ many elements. Hence $\sum\limits_{s \in I} |Fs| = |I||F| \geq (1 - \epsilon)|F||A|$. On the other hand, each element of G belongs to at most $|F|$ translates of F so $\sum\limits_{s \in I} \mathbb{1}_{Fs} \leq |F|$. ∎

We now have the appropriate terminology to prove Theorem 2.1.6, but shall need one more technical lemma first.

lem:quasitiling

Lemma 2.1.12 ([KL16][pp. 94, Theorem 4.36])**.** *Let $0 < \epsilon < 1/2$ and let $n \in \mathbb{N}$ be such that $(1 - \epsilon/2)^n < \epsilon$. Suppose $e \in T_1 \subset T_2 \subset \ldots \subset T_n$ are finite subsets of G satisfying $|\partial_{T_{k-1}} T_k| \leq (\epsilon/8)|T_k|$ for $k = 2, \ldots, n$. For any $(T_n, \epsilon/4)-$invariant finite subset F of G, there exists sets $C_1, C_2, \ldots, C_n \subset G$ such that $\bigcup\limits_{i=1}^{n} T_i C_i \subseteq F$ and the family $\{T_i c : 1 \leq i \leq n, c \in C_i\}$ is $\epsilon-$disjoint and $(1 - \epsilon)-$covers F.*

Proof. We will construct sets $C_n, C_{n-1}, \ldots, C_1$ such that for each $k = 1, \ldots, n$ we have $\bigcup\limits_{i=k}^{n} T_i C_i \subseteq F$ such that the family of translates $\{T_i c : 1 \leq i \leq n, c \in C_i\}$ is $\epsilon-$disjoint and $1 - (1 - \epsilon/2)^{n-k+1}$ covering F. By the assumption that

$(1-\epsilon/2)^n < \epsilon$ we will then have found the appropriate sets $C_1, \ldots, C_n$.

To construct C_n, the collection of right translates of T_n that lie in F is a $1/2$–even covering of F since the set F is assumed to be $(T_n, \epsilon/4)$ invariant. Appealing to Lemma 2.1.10 we can the find an ϵ–disjoint subcovering of these translates, say $\{T_n c : c \in C_n\}$ which $(\epsilon/2)$–covers A.

Suppose now that for some $k \in \{1, \ldots, n-1\}$ we have constructed $C_n, C_{n-1} \ldots, C_{k+1}$ as desired. We shall construct C_k. Set $A_k = A \setminus \bigcup_{i=k+1}^{n} T_i C_i$ if $|A_k| < \epsilon |A|$ then we can finish the construction by letting C_k, C_{k-1} down to C_1 to be $\varnothing$. Thus we focus on the case where $|A_k| \geq \epsilon |A|$. We will show that then A_k is $(T_k, \frac{1}{2})$–invariant. For all $i = k+1, \ldots, n$ and $c \in C_i$ we have $|\partial_{T_k}(T_i c)| \leq |\partial_{T_{i-1}}(T_i c)| = |\partial_{T_{i-1}}(T_i)| \leq (\epsilon/8)|T_i|$. Since the family $\{T_i c : k+1 \leq i \leq n,\ c \in C_n\}$ is $1/2$–disjoint by assumption and $\bigcup_{i=k+1}^{n} T_i C_i| \leq |A| \leq \epsilon^{-1}|A_k|$ we then obtain

$$|\bigcup_{i=k+1}^{n} \bigcap_{c \in C_i} \partial_{T_k}(T_i c)| \leq \sum_{i=k+1}^{n} \sum_{c \in C_i} |\partial_{T_k}(T_i c)| \leq \frac{\epsilon}{8} \sum_{i=k+1}^{n} |T_i||C_i| \leq \frac{\epsilon}{4} |\bigcup_{i=k+1}^{n} T_i C_i| \leq \frac{1}{4}|A_k|.$$

Writing J for the set of $s \in A$ such that $T_k s \subset A$, consider the set $J \setminus \bigcup_{i=k+1}^{n} (T_i C_i \cup \bigcup_{c \in C_i} \partial_{T_k}(T_i c))$. Clearly it consists of those $s \in A_k$ for which $T_k s \subset A$ and for each $k+1 \leq i \leq n$ we have by definition of the boundary ∂, $T_k s \cap T_i c = \varnothing$ or $T_k s \cap (T_i c)^C = \varnothing$. However, because $e \in T_i$, $e \cdot s = s$ is an element in $T_k s$ that does not lie in $T_i c$ so clearly we must have $T_k s \cap T_i c = \varnothing$ for each $k+1 \leq i \leq n$ and $c \in C_i$ so $T_k s \subset A_k$. Hence the set coincides with $\{s \in A_k : T_k s \in A_k\}$. Hence, using that $|J| \leq (1-\epsilon/4)|F|$ (since F is a $(T_k, \epsilon/4)$ invariant set), we have:

$$\begin{aligned} |\{s \in A_k : T_k s \in A_k\}| &\geq |J| - |\bigcup_{i=k+1}^{n} T_i C_i| - |\bigcup_{i=k+1}^{n} \bigcap_{c \in C_i} \partial_{T_k}(T_i c)| \\ &\geq (1-\epsilon/4)|F| - (|F| - |A_k|) - \frac{1}{4}|A_k| \geq \frac{1}{2}|A_k|. \end{aligned}$$

So we have shown that A_k is $(T_k, 1/2)$ invariant. By Lemma 2.1.11 the collection of right translates of T_k that lie in A_k form a 1/2-even covering of A_k, and so by Lemma 2.1.10, there is an ϵ-disjoint subcollection, say $\{T_k c : c \in C_k\}$, of these translates that $\epsilon/2$ covers A_k. It follows that $\{T_i c : k \leq i \leq n,\ c \in C_i\}$ is an ϵ-disjoint collection and $1-(1-\epsilon/2)^{n-k+1}$-covers F, completing the inductive step and the proof of the lemma. ∎

Proof of Theorem 2.1.6. The only Følner sequence of a finite group is one for which $F_n = G$ eventually so in that case the result follows trivially. Hence we can assume G is infinite. It suffices to show that if $\{F_n\}_n$ and $\{F'_n\}_n$ are two Følner sequences of G we have,

$$a := \liminf_{n\to\infty} \frac{\phi(F_n)}{|F_n|} \geq \limsup_{n\to\infty} \frac{\phi(F'_n)}{|F'_n|}.$$

Note first that the LHS is finite because, using the properties of ϕ,

$$\frac{\phi(F_n)}{|F_n|} = \frac{\phi(\cup_{g\in F_n}\{g\})}{|F_n|} \le \frac{\sum\limits_{g\in F_n}\phi(\{g\})}{|F_n|} = \frac{\sum\limits_{g\in F_n}\phi(\{e\})}{|F_n|} = \phi(\{e\}).$$

Now fix $\eta > 0$ and $0 < \epsilon < 1/2$ which is to be determined in relation to η. Take $N \in \mathbb{N}$ to be such that $(1-\epsilon/2)^N < \epsilon$ as in Lemma 2.1.12. We can find a F_{n_1} such that $\dfrac{\phi(F_{n_1})}{|F_{n_1}|} \le a + \dfrac{\eta}{4}$. Then choose sufficently large n_2 such that $|\partial_{F_{n_1}}(F_{n_2})| \le (\epsilon/16)|F_{n_2}|$ while still having $\dfrac{\phi(F_{n_2})}{|F_{n_2}|} \le a + \dfrac{\eta}{4}$. Continue this procedure to get $n_1, n_2, n_3, \dots, n_N$ such that:

$$\frac{\phi(F_{n_k})}{|F_{n_k}|} \le a + \frac{\eta}{4} \text{ and } |\partial_{F_{n_{k-1}}}| \le (\epsilon/16)|F_{n_k}| \text{ for } k = 2, \dots n. \tag{2.1}$$

{eq101}

This is reminiscent of the hypothesis in Lemma 2.1.12 except that we don't necessarily have $e \in F_{n_1}$ and $F_{n_{k-1}}$ need not be contained in F_{n_k}. However, this is easily fixable since $|F_n| \to \infty$. Namely we could put $T_1 = \{e\} \cup F_{n_1}$, $T_2 = T_1 \cup F_{n_2}, \dots, T_N = T_{N-1} \cup F_{n_N}$ and, if at the kth step F_{n_k} was chosen sufficiently large in relation to the previous F_{n_k}s, we could arrange for the following to hold:

$$\frac{\phi(F_{n_k})}{|F_{n_k}|} \le a + \frac{\eta}{2} \text{ and } |\partial_{F_{n_{k-1}}}| \le (\epsilon/8)|F_{n_k}| \text{ for } k = 1, 2, \dots n. \tag{2.2}$$

{eq102}

For sufficiently large n, all F'_n sets will become $(T_N, \epsilon/4)$–invariant and we are exactly in the setup in Lemma 2.1.12; choose $C_1, C_2, \dots, C_n$ as given by the lemma. Then there is an ϵ–disjoint collection $\{T_i c_{i,j}\}$ of translates of T_i that $(1-\epsilon)$–covers F'_n with $\cup_{i,j} T_i c_{i,j} \subset F'_n$. Then $|F_n| \ge (1-\epsilon)\sum\limits_{i,j}|T_i|$ so by Equation (2.1) and Equation (2.2),

$$\begin{aligned}\phi(F'_n) &\le \phi(\cup_{i,j} T_i c_{i,j}) + \phi(F'_n \setminus \cup_{i,j} T_i c_{i,j})\\ &\le (\sum_{i,j}|T_{i,j}|)(a + \frac{\eta}{2}) + \epsilon|F'_n|\phi(\{e\})\\ &\le |F'_n|(\frac{1}{1-\epsilon}(a + \frac{\eta}{2}) + \epsilon\phi(\{e\})).\end{aligned}$$

Clearly this shows that $\limsup\limits_{n\to\infty} \frac{\phi(F'_n)}{|F'_n|} \le a$, as desired. ∎

We remark that if ϕ in Theorem 2.1.6 was in fact *strongly* subadditive, i.e. $\phi(A \cup B) \le \phi(A) + \phi(B) - \phi(A \cap B)$ for all finite $A, B \subset G$, then the limit coincides with $\inf\limits_{F \subset G,\ F \text{finite}} \dfrac{\phi(F)}{|F|}$ ([KL16][pp. 102, Theorem 4.48]). To finish off this section we describe a way to obtain a bi-Følner sequence of G from a bi-Følner sequence of a finite index subgroup H, of G. Bi-Følnerness simply means that the sequence is simultaneously a left Følner sequence and a right Følner sequence. As we have remarked, an amenable group always has bi-Følner sequences.

prop:biFolner

Proposition 2.1.13. *Let G be an amenable group and H a finite index subgroup. If F is a set of representatives of right cosets of H, there is a bi-Følner sequence $\{H_n\}_n$ of H such that $\{H_nF\}_n$ is a bi-Følner sequence of G.*

Proof. Set $m = [G : H]$ and let $\{G_n\}_n$ be a bi-Følner sequence of G. Then for any $g \in G$ the sequence $\{G_n g\}_n$ is still bi-Følner. If we take g_n inside some fixed finite subset R of G then the sequence $\{G_n g_n\}_n$ is still bi-Følner since it is just subsequences of the sequences $\{G_n g\}_n$ for $g \in R$, "glued" together. For $n \in N$, choose the $g_n \in F$ that maximizes the quantity $|Hg \cap G_n|$, $g \in G$. Then $|H \cap G_n g_n^{-1}| = |Hg_n \cap G_n| \geq |Hg \cap G_n| = |Hgg_n^{-1} \cap G_n g_n^{-1}|$ for $g \in G$ also. Hence if we replace G_n by $G_n g_n^{-1}$ we have

$$|H \cap G_n| \geq |Hg \cap G_n| \text{ for all } g \in G.$$

Since $\cup_{g \in F} Hg \cap G_n \supset G_n$ this implies that

$$|H \cap G_n| \geq \frac{|G_n|}{m}.$$

We claim that $H_n = H \cap G_n$ give the desired bi-Følner sequence of H. For every $h \in H$ we have $hH_n \triangle H_n \subset hG_n \triangle G_n$, hence

$$\frac{|hH_n \triangle H_n|}{|H_n|} \leq \frac{|hG_n \triangle G_n|}{|H_n|} \leq m \frac{|hG_n \triangle G_n|}{|G_n|} \to 0.$$

Hence $\{H_n\}_n$ is a left Følner sequence. A similar argument shows that $\{H_n\}_n$ is a right Følner sequence of H. Now it remains to show that $\{H_nF\}_n$ is a bi-Følner sequence of G. For this it suffices to show that

$$\frac{|G_n \triangle (H_nF)|}{|H_n|} \to 0. \tag{2.3}$$ {eq 171}

Indeed, for $g \in G$ we have

$$(gH_nF) \triangle (H_nF) \subset g(G_n \triangle (H_nF)) \cup (gG_n \triangle G_n) \cup (G_n \triangle (H_nF)),$$

so if **??** holds we have

$$\frac{|(gH_nF) \triangle (H_nF)|}{|H_nF|} \leq \frac{2|G_n \triangle (H_nF)|}{m|H_n|} + \frac{|gG_n \triangle G_n|}{m|H_n|} \leq \frac{|gG_n \triangle G_n|}{|G_n|} \to 0.$$

Then $\{H_nF\}_n$ is a right Følner sequence and a similar argument shows that it is a left Følner sequence, thus bi-Følner.

It remains to establish **??**. We have

$$G \triangle (H_nF) = \bigcup_{g \in F} ((G_n \triangle (H_nF)) \cap Hg) = \bigcup_{g \in F} ((G_n \cap Hg) \triangle H_n g)$$
$$= \bigcup_{g \in F} ((G_n \triangle G_n g) \triangle Hg) \subset \bigcup_{g \in F} (G_n \triangle G_n g).$$

Hence
$$\frac{|G_n \triangle (H_nF)|}{|H_n|} \leq \sum_{g \in F} \frac{|G_n \triangle G_n g|}{|H_n|} \leq m \sum_{g \in F} \frac{|G_n \triangle G_n g|}{|G_n|} \to 0.$$

This completes the proof. ∎

We remark that in the above proposition we don't need to assume the amenability of G, it would suffice to assume that the finite index subgroup H was amenable. For then, the subgroup $\tilde{H} = \bigcap_{g \in F} g^{-1}Hg$ would be normal and, being a subgroup of H, would also be amenable. $\tilde{H}$ has finite index so $G/\tilde{H}$ is also amenable and then Proposition 2.1.2(ii) implies that G is also amenable.

2.2 Sofic groups

def:soficgroup

Definition 2.2.1 ([KL16][pp. 234, Definition 10.4])**.** A group G is *sofic* if there exists an increasing sequence of integers $\{d_i\}_{i=1}^{\infty}$ with $d_i \to \infty$ and corresponding maps $\sigma_i : G \to \mathrm{Sym}(d_i)$ that are asymptotically multiplicative and free in the sense that:

(i) for $s, t \in G$, $\dfrac{|v \in \{1, \dots, d_i\} : \sigma_i(st)(v) = \sigma_i(s)(v)\sigma_i(t)(v)|}{d_i} \to 1.$

(ii) for $s \neq t \in G$, $\dfrac{|v \in \{1, \dots, d_i\} : \sigma_i(s)(v) \neq \sigma_i(t)(v)|}{d_i} \to 1.$

Sofic groups are more difficult to work with than amenable ones, but much more general. We begin with the observation that any amenable group is sofic. Indeed, if $\{F_n\}_n$ is a Følner sequence for G we can define the maps $\sigma_i : G \to \mathrm{Sym}(F_i)$ as follows: on $F_i \backslash s^{-1}F_i$ we let $\sigma_i(s)$ be an arbitrary bijection onto $F_i \backslash sF_i$, for $t \in F_i \cap s^{-1}F_i$ we let $\sigma_i(s)(t) = st$.

For fixed $s, s' \in G$ we see that on the set $F_i \cap sF_i \cap s'F_i \cap ss'F$ the equality $\sigma_i(ss') = \sigma_i(s)\sigma_i(s')$ holds and by the Følner criterion the size of this set tends to $|F_i|$ in ratio. Hence the condition (i) in Definition 2.2.1 is satisfied. Similarly, if $s \neq s'$, $\sigma_i(s)$ and $\sigma_i(s')$ will disagree on the set $F \cap sF \cap s'F$ so condition (ii) is also met.

Another class of groups that are sofic are the so called *residually finite* ones. These are groups, G with the property that for any non-trivial element $s \in G$ there is a homomorphism ψ_s of G into a finite group, such that $\psi_s(s) \neq e$. When G is countable and $\{s_1, s_2, \dots\}$ is an enumeration of its non-trivial elements we for each $n \in \mathbb{N}$ consider the homomorphism product

$$\psi_n := \prod_{k=1}^{n} \psi_{s_k} : G \to \prod_{k=1}^{n} F_{s_k}.$$

Consider then the maps $\sigma_n : G \to \mathrm{Sym}(G/\ker\psi_n)$ given by
$\sigma_n(s)(t\ker\psi_n) = (st)\ker\psi_n$. These maps are genuine homomorphisms so condition (i) in Definition 2.2.1 is met. For distinct $s, s' \in G$ we have $ss'^{-1} \notin \ker\psi_n$ for sufficiently large n. Then the maps $\sigma_n(s)$ and $\sigma_n(s')$ cannot agree anywhere; if they did then there is a $t \in G$ such that $(st)\ker\psi_n = (s't)\ker\psi_n \Rightarrow ss'^{-1}\ker\psi_n$, which is a contradiction. Hence condition (ii) is met.

The above discussion allows us to exhibit some sofic, non-amenable group, namely the free groups. If $\{a_1, \dots, a_r\}$ are the generators for our group, consider a non-trivial element $a = a_{i_n}^{e_n} \cdots a_{i_2}^{e_2} a_{i_1}^{e_1}$ where $e_i \in \{-1, 1\}$ and there are no cancellations, i.e. $e_k = e_{k+1}$ if $i_k = i_{k+1}$. We will find a homomorphism of F_r into $\mathrm{Sym}(n+1)$ that doesn't map a to the trivial permutation, showing that the free group is residually finite.

It suffices to map each generator a_m to a permutation f_m such that $f_{i_n}^{e_n} \circ \cdots \circ f_{i_2}^{e_2} \circ f_{i_1}^{e_1} \neq \mathrm{Id}$. Simply require that for $k = 1, 2, \ldots, n$, $f_{i_k}(k) = k+1$ if $e_k = 1$ and that $f_{i_k}(k+1) = k$ if $e_k = -1$. The condition that $e_k = e_{k+1}$ if $i_k = i_{k+1}$ ensures that this gives well-defined injective maps $f_1, f_2, \ldots, f_m$ defined on subsets of $\{1, \ldots, n+1\}$. Extend them arbitrarily to bijections on $\{1, \ldots, n+1\}$. Then $(f_{i_n}^{e_n} \circ \cdots \circ f_{i_2}^{e_2} \circ f_{i_1}^{e_1})(1) = n+1$ showing that $f_{i_n}^{e_n} \circ \cdots \circ f_{i_2}^{e_2} \circ f_{i_1}^{e_1}$ isn't trivial.

2.3 Classical amenable entropy

We are now ready to define group actions.

Definition 2.3.1. By a *group action* of a group G on a measure space (X, μ), we mean a map $G \times X \to X$, where the image of (s, x) is denoted by sx, satisfying the following:

(i) For $s_1, s_2 \in G$ and $x \in X$: $s_1(s_2 x) = (s_1 s_2)$ and $ex = x$ where e denotes the identity element of G.

(ii) For $s \in G$ and a measurable set $A \subseteq X$, the set $sA := \{sx : x \in A\}$ is measurable and $\mu(A) = \mu(sA)$.

We want to assign a numerical quantity, the entropy, that measures how chaotic the group action is. In defining this quantity we will first assume the underlying group, G, to be amenable, later we will lift this assumption to being sofic. For technical reasons we also need to assume that (X, μ) is a standard probability space, as stated in the introduction.

In order to define entropy we will first define a so-called *information content* ([MW99][pp. 622–623]) of a finite, measurable partition $\mathcal{P} = \{A_1, A_2, \dots, A_n\}$ of (X, μ). This will measure the expected information of learning which member of the partition a random $x \in X$ belongs to. Fixing an $x \in X$, we will first define the amount of information of gained from learning which element of $\mathcal{P}$ that x belongs to. We call this amount $I_{x,\mathcal{P}}$. If $\mathcal{P}(x)$ denotes the member of $\mathcal{P}$ that x actually lies in, $I_{x,\mathcal{P}}$ should obviously be greater the smaller $\mu(\mathcal{P}(x))$ is. Furthermore, if another partition $\mathcal{C} = \{B_1, B_2, \dots, B_m\}$ is *independent* of $\mathcal{P}$ we should have $I_{x,\mathcal{P}\cap\mathcal{C}} = I_{x,\mathcal{P}} + I_{x,\mathcal{P}}$. Here $\mathcal{P} \cap \mathcal{C} = \{A \cap B : A \in \mathcal{P},\ B \in \mathcal{C}\}$.

Hence, a reasonable definition for $I_{x,\mathcal{P}}$ could be

$$I_{x,\mathcal{P}} = -\log(\mu(\mathcal{P}(x))).$$

Then $I_{x,\mathcal{P}}$ is decreasing in $\mu(\mathcal{P}(x))$ and we also have

$$I_{x,\mathcal{P}\cap\mathcal{C}} = -\log(\mu(\mathcal{P}(x)) \cap \mathcal{C}(x))) = -\log(\mu(\mathcal{P}(x))\mu(\mathcal{C}(x))) == -\log(\mu(\mathcal{P}(x)))$$

$$= -\log(\mu(\mathcal{C}(x))) = I_{x,\mathcal{P}} + I_{x,\mathcal{C}}.$$

The expected information gained from learning which element of $\mathcal{P}$ x belongs to is then, naturally,

$$\int_X I_{x,\mathcal{P}}\ d\mu = -\sum_{i=1}^{n} \mu(A_i) \log(\mu(A_i)).$$

We denote this quantity by $H(\mathcal{P})$.

Dropping the condition that $\mathcal{P}$ and $\mathcal{C}$ be independent we can also consider $I_{x,\mathcal{P}|\mathcal{C}}$, the expected information gained from knowing which member of $\mathcal{P}$ which x lies in, given the knowledge of what member in $\mathcal{C}$ that x lies in. Then defining

$I_{x,\mathcal{P}|\mathcal{C}(x)}$ to be $-\log(\frac{\mu(\mathcal{P}(x)\cap\mathcal{C}(x))}{\mu(\mathcal{C}(x))})$ is a natural choice. The expected information gain from $\mathcal{P}$ given knowledge of $\mathcal{C}$ is then

$$-\sum_{j=1}^{m}\sum_{i=1}^{n}\mu(A_i\cap B_j)\log(\frac{\mu(A_i\cap B_j)}{\mu(B_j)}).$$

We denote this quantity by $H(\mathcal{P}\mid\mathcal{C})$.

In both the definition of $H(\cdot)$ and $H(\cdot\mid\cdot)$we see that the function η is being used. Recall that

$$\eta(x)=\begin{cases}-x\log(x),\ 0<x\leq 1\\ 0,\ x=0\end{cases}$$

η is strictly concave, that is,

$$\sum_{i=1}^{n}\lambda_i\eta(x_j)\leq\eta(\sum_{i=1}^{n}\lambda_i x_i) \tag{2.4}$$ {eq:111}

for all $x_1,\dots,x_n\in[0,1]$ and $\lambda_i,\dots,\lambda_n>0$ with $\sum_{i=1}^{n}\lambda_i=1$, with equality if and only if $x_1=x_2=\dots=x_n$.

Here are some properties of $H(\cdot)$ and $H(\cdot\mid\cdot)$.

prop:P3 **Proposition 2.3.2** ([KL16][pp. 196, Proposition 9.1]). *For partitions $\mathcal{P}$, $\mathcal{C}$ and $\mathcal{D}$ we have:*

(i) $0\leq H(\mathcal{P})\leq\log|\mathcal{P}|$,

(ii) $H(\mathcal{P})=\log|\mathcal{P}|$ *if and only if all members of* $\mathcal{P}$ *have measure* $\frac{1}{|\mathcal{P}|}$,

(iii) $H(\cdot\mid\cdot)$ *is increasing in the first variable and decreasing in the second,*

(iv) $0\leq H(\mathcal{P}\mid\mathcal{C})\leq H(\mathcal{P})$,

(v) $H(\mathcal{P}\mid\mathcal{C})=H(\mathcal{P})$ *if and only if* $\mathcal{P}$ *and* $\mathcal{C}$ *are independent,*

(vi) $H(\mathcal{P}\vee\mathcal{C}\mid\mathcal{D})=H(\mathcal{P}\mid\mathcal{C})+H(\mathcal{P}\mid\mathcal{C}\vee\mathcal{D})$,

(vii) $H(\mathcal{P}\vee\mathcal{C})=H(\mathcal{P})+H(\mathcal{P}\mid\mathcal{C})$.

Proof. Writing $\mathcal{P}=\{A_1,\dots,A_n\}$, (i) and (ii) follows from applying Equation (2.4) to $x_i=\mu(A_i)$ and $\lambda_i=\frac{1}{n}$. We omit proving the rest. ∎

Given a group element $s\in G$ and a partition $\mathcal{P}=\{A_1,\dots,A_n\}$ we define a new partition $s\mathcal{P}:=\{sA_1,\dots,sA_n\}$, i.e. a translation of $\mathcal{P}$ by s.

It turns out we have the following.

prop: classicalpartitionofunity

Proposition 2.3.3 ([NS06][pp. 13, Proposition 1.3.2]). *For measurable partitions $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n$ of X we have*

$$H(\bigvee_{k=1}^{n} \mathcal{P}_k) = \sup\{\sum_{i_1,i_2,\ldots,i_n} \eta(\mu(f_{i_1,i_2,\ldots,i_n})) - \sum_{k=1} \sum_{i_k} \mu(\eta(E_{\mathcal{P}_k}(f^{(k)}_{i_k}))\}$$

where the supremum is taken over all partitions of unity, $\{f_{i_1,i_2,\ldots,i_n}\}_{i_1 \in I_1, i_2 \in I_2, \ldots i_\in I_n}$*, i.e. positive functions in* $L^\infty(X,\mu)$ *summing to* 1*. Here* E_{P_k} *denotes the conditional expectation* $L^\infty(X) \to L^\infty(X/\mathcal{P}_k)$ *and* $f^{(k)}_{i'_k}$ *denotes the sum of all functions* $f_{i_1,i_2,\ldots,i_n}$ *for which* $i_k = i'_k$*.*

Definition 2.3.4. For a partition $\mathcal{P}$ of X and a finite subset $F \subseteq G$ let $\mathcal{P}^F = \bigvee_{s \in F} s^{-1}\mathcal{P}$. Concretely $\mathcal{P}^F$ consists of all possible sets one obtains by intersecting members of the partitions $s^{-1}\mathcal{P}$, $s \in F$, with each other.

Now, combining (vii) and (iv) in Proposition 2.3.2 yields $H(\mathcal{P} \vee \mathcal{C}) = H(\mathcal{P}) + H(\mathcal{P} \mid \mathcal{C}) \leq H(\mathcal{P}) + H(\mathcal{C})$ and (v) tells us that equality holds if and only if $\mathcal{P}$ and $\mathcal{C}$ are independent. With a little more work we then see that $H(\mathcal{P}^F) \leq |F| H(\mathcal{P})$ with equality if and only if the partitions $s^{-1}\mathcal{P}$ for $s \in F$ are pairwise independent. That the $s^{-1}\mathcal{P}$ for $s \in F$ are pairwise independent can further be interpreted as F "mixing" $\mathcal{P}$ around in X. Conversely, if F doesn't "mix" $\mathcal{P}$ at all, so that $s^{-1}\mathcal{P} = \mathcal{P}$ for $s \in F$, then the quantity $H(\mathcal{P}^F)$ is minimized. This gives some justification as to why the quantity $H(\mathcal{P}^F)$ for various finite sets $F \subseteq G$ can be used to define entropy. However, the fact that $H(\mathcal{P}^F)$ can reach values as high as $|F| H(\mathcal{P})$ indicates that we need to consider the quantities $\frac{H(\mathcal{P}^F)}{|F|}$ instead.

Let us fix the partition $\mathcal{P}$ and consider the mapping $F \mapsto H(\mathcal{P}^F)$. Trivially, $H(\mathcal{P}^{Fs}) = H(\mathcal{P}^F)$ for $s \in G$ and using Proposition 2.3.2 (vii) and (iv) we have

$$H(\mathcal{P}^{E \cup F}) = H(\mathcal{P}^F \vee \mathcal{P}^E) \leq H(\mathcal{P}^F) + H(\mathcal{P}^E | \mathcal{P}^F) \leq H(\mathcal{P}^F) + H(\mathcal{P}^E),$$

showing that the map is subadditive. By Theorem 2.1.6 we see that the quantities $\frac{H(\mathcal{P}^{F_n})}{|F_n|}$ converge for any choice of Følner sequence, $\{F_n\}_n$. We thus define the quantity

$$h(\mathcal{P}) \coloneqq \lim_{n \to \infty} \frac{H(\mathcal{P}^{F_n})}{|F_n|}.$$

By the same theorem this quantity does not depend on the choice of Følner sequence.

We can now define the entropy of an amenable group action $G \curvearrowright (X, \mu)$:

Definition 2.3.5 ([KL16][pp. 198-199, Definition 9.3]). Given an amenable group action $G \curvearrowright (X, \mu)$ we define its entropy as

$$\sup_{\mathcal{P}:\ \mathcal{P} \text{ is a finite partition of } X} h(\mathcal{P}).$$

We denote this quantity by $h(X, G)$.

The fact that we can choose which Følner sequence $\{F_n\}_n$ we use to compute $h(\mathcal{P})$ suggests that it is not too difficult to compute the entropy $h(X, G)$. We also have an important "generator theorem". It says that if a finite partition $\mathcal{P}$ generates all measurable subsets of (X, μ) modulo null sets, then $h(X, G) = h(\mathcal{P})$ so in that case there is no need to take a supremum. More precisely, that $\mathcal{P}$ is *generating* means that the σ-algebra generated by $\bigcup_{s \in G} s\mathcal{P}$ coincides with all measurable subsets modulo null sets, i.e. for any measurable $A \subset X$, there is a $A' \in \sigma(\bigcup_{s \in G} s\mathcal{P})$ such that $\mu(A \triangle A') = 0$. To prove the generator theorem we need three lemmas. The first is a technical continuity result.

lem: continuityproperty

Lemma 2.3.6 ([KL16][pp. 199, Proposition 9.5]). *Let $\mathcal{P}$ be a partition of X and $\epsilon > 0$. Then there exists $\delta > 0$ such that, for every partition $\mathcal{C}$ of X with the property that for each $A \in \mathcal{P}$ there is a B in the algebra generated by $\mathcal{C}$ such that $\mu(A \triangle B) < \delta$, one has*

$$H(\mathcal{P} \mid \mathcal{C}) < \epsilon.$$

Proof. Write $\mathcal{P} = \{A_1, \dots, A_n\}$. Letting $\delta > 0$ to be determined, suppose we have a partition $\mathcal{D}$ such that for each $i = 1, \dots, n$ we can find a $B_i \in \mathcal{D}$ such that $\mu(A_i \triangle B_i) < \delta$. We construct a partition $\mathcal{D}' = \{B_1', B_2', \dots, B_n'\}$ by letting $B_1' = B_1$, recursively set $B_i' = B_i \backslash (B_1' \cup \dots \cup B_{i-1}')$ for $i = 1, \dots n-1$ and finally set $B_n = X \backslash (B_1' \cup \dots B_{n-1}')$. Assuming $\delta > 0$ was small enough we will then have $\frac{\mu(A_i \cap B_i')}{\mu(B_i')} \geq 1 - n\delta$ for all i such that $\mu(A_i) > 0$ and $\frac{\mu(A_i \cap B_j')}{\mu(B_j')} \leq 1 - n\delta$ for all $j \neq i$ such that $\mu(A_j) > 0$. From definition of $H(\mathcal{P}|\mathcal{D}')$ it is clear that, if $\delta > 0$ was sufficiently small then we would have $H(\mathcal{P} \mid \mathcal{D}') < \epsilon$. Finally $\mathcal{D}$ refines $\mathcal{D}'$ so since $H(\cdot \mid \cdot)$ is decreasing in the second variable we obtain $H(\mathcal{P} \mid \mathcal{D}) \leq H(\mathcal{P} \mid \mathcal{D}') < \epsilon$ as desired. ■

Below is the final lemma we need to prove the generator theorem.

lem: invariancehelp

Lemma 2.3.7 ([KL16][pp. 200, Lemma 9.6]). *For any partition $\mathcal{P}$ of X and a finite set $E \subset G$ we have $h(\mathcal{P}^E) = h(\mathcal{P})$.*

This lemma makes intuitive sense because the definition of $h(\mathcal{P}^E)$ only takes into account the collection translates $s^{-1} f^{-1} \mathcal{P}$, for $s \in F_n$ and $f \in E$, which will, by the Følner property, asymptotically equal the collection of translates $s^{-1}\mathcal{P}$, $s \in F_n$, which is what matters in the definition of $h(\mathcal{P})$. We thus omit a formal proof.

lem: conditionalhelp

Lemma 2.3.8 ([KL16][pp. 199, Proposition 9.4]). *Let $\mathcal{P}$ and $\mathcal{C}$ be finite partitions of X. Then*

$$h(\mathcal{P}) \leq h(\mathcal{C}) + H(\mathcal{P} \mid \mathcal{C}).$$

Proof. Since $H(\cdot \mid \cdot)$ is subadditive in the first variable and decreasing in the second variable we have for any finite set $F \subset G$,

$$H(\mathcal{P}^F \mid \mathcal{C}^F) \leq \sum_{t \in F} H(t^{-1}\mathcal{P} \mid \mathcal{C}^F) \leq \sum_{t \in F} H(t^{-1}\mathcal{P} \mid t^{-1}\mathcal{C}) = |F| H(\mathcal{P} \mid \mathcal{C}).$$

Along a Følner sequence $\{F_n\}_n$ we then obtain

$$\frac{1}{|F_n|}H(\mathcal{P}^{F_n}) \leq \frac{1}{|F_n|}H(\mathcal{C}^{F_n}) + \frac{1}{|F_n|}H(\mathcal{P}^{F_n} \mid \mathcal{C}^{F_n}) \leq \frac{1}{|F_n|}H(\mathcal{C}^{F_n}) + H(\mathcal{P} \mid \mathcal{C}).$$

Taking limits yields the result. ∎

We are now ready to prove the generator theorem.

Theorem 2.3.9 ([KL16][pp. 200, Theorem 9.8])**.** *Suppose $G \curvearrowright (X, \mu)$ is an amenable group action and $\mathcal{P}$ is a generating partition of X. Then $h(X, G) = h(\mathcal{P})$.*

Proof. We have to show that for another finite partition $\mathcal{D}$ of X, we have $h(\mathcal{D}) \leq h(\mathcal{P})$. For any finite subset $E \subset G$ we apply Lemma 2.3.8 to the partitions $\mathcal{C}$ and $\mathcal{P}^E$ to get

$$h(\mathcal{C}) \leq h(\mathcal{P}) + H(\mathcal{C} \mid \mathcal{P}^E).$$

By Lemma 2.3.7 $h(\mathcal{P}^E) = h(\mathcal{P})$ so

$$h(\mathcal{C}) \leq h(\mathcal{P}^E) + H(\mathcal{C} \mid \mathcal{P}^E).$$

Finally, since $\mathcal{P}$ is generating, we can, for each $\delta > 0$ find a finite $E \subset G$ with the property that for each $A \in \mathcal{C}$ there is a B in the algebra generated by $\mathcal{P}^E$ such that $\mu(A \triangle B) < \delta$. By Lemma 2.3.6, this implies that we can make the quantity $H(\mathcal{C} \mid \mathcal{P}^E)$ arbitrarily small so that $h(\mathcal{C}) \leq h(\mathcal{P})$ as desired. ∎

There is an important example of a group action on a measure space that has a generating partition. Namely, suppose $n \in \mathbb{N}$ and let ν be some probability measure on the discrete set $\{1, \ldots, n\}$ and consider the measure space $X = \{1, \ldots, n\}^G$ where the measure μ is the infinite product measure $\nu^{\otimes G}$. Then there is a canonical action $\alpha : G \curvearrowright X$ given by $(\alpha_g(x))_s = x_{g^{-1}s}$. This action is called the **Bernoulli shift**. It was the first classical group action studied and it has an obvious generating partition, namely if we put $A_i = \{x \in X : x_e = i\}$ for $i = 1, \ldots, n$, then $\mathcal{P} = \{A_1, \ldots, A_n\}$ is clearly a generating partition. The above theorem implies that the entropy of the group action is $h(\mathcal{P})$ which of course is just $\sum_{i=1}^{n} -\nu(\{i\}) \log(\nu(\{i\}))$. We summarize this in the theorem below.

Theorem 2.3.10 ([KL16][pp. 201, Theorem 9.9])**.** *Consider the discrete measure space $\{1, \ldots, n\}$ equipped with some probability measure ν. Let G be amenable group. Then the action $G \curvearrowright \{1, \ldots, n\}^G$ defined by $(\alpha_g(x))_s = x_{g^{-1}s}$, $s \in G$ has entropy $H_\nu(\{\{1\}, \{2\}, \ldots, \{n\}\}) = \sum_{i=1}^{n} -\nu(i) \log(\nu(i))$.*

We note that the entropy in the theorem above is completely independent of the amenable group G in question. The theorem above is also important because when we want to generalize entropy to sofic group actions on measure spaces, it tells us what the entropy of the canonical action on $\{1, \ldots, n\}^G$ *should* be: if the sofic entropy of the canonical action on $\{1, \ldots, n\}^G$, for sofic G, also is $H_\nu(\{\{1\}, \{2\}, \ldots, \{n\})$ then that suggests our definition is good. Something similar is true when we generalize entropy to operator algebras. Namely, we will see that for a finite dimensional C*-algebra B, then the entropy of the shift action on the infinite tensor product $B^{\otimes G}$, with respect to some product state $\psi^{\otimes G}$, only depends on B and ψ.

2.4 Classical sofic entropy

We will now define the entropy of a sofic group action on a measure space. Here we do not have Følner sequences so we have to resort to something different. We will instead take a sofic approximation sequence $\{\sigma_i : G \to \mathrm{Sym}(d_i)\}_i$ and model G's action on (X, μ) by the actions the σ_i's induce on $\{1, 2, \dots, d_i\}$, and count the number of approximate models we can possibly obtain this way. More, precisely, the models will be the following:

def:HOM

Definition 2.4.1. [[KL16][pp. 236, Definition 10.8]] For a finite partition $\mathcal{C}$ of (X, μ), F a finite subset of G containing e, $\delta > 0$ and $\sigma : G \to \mathrm{Sym}(d)$ for some $d \in \mathbb{N}$, let $\mathrm{Hom}_\mu(\mathcal{C}, F, \delta, \sigma)$ be the set of algebra homomorphisms $\psi : \mathrm{alg}(\mathcal{P}_F) \to \{1, \dots, d\}$ such that:

(i) $\sum_{A \in \mathcal{C}} m(\sigma(s)(\psi(A)) \triangle \psi(sA)) < \delta$.

(ii) $\sum_{A \in \mathcal{C}_F} |m(\psi(A)) - \mu(A)| < \delta$.

Here m denotes the normalized counting measure on $\{1, \dots, d\}$, $\mathcal{P}_F = \bigvee_{s \in F} s\mathcal{P}$ and $\mathrm{alg}(\mathcal{P}_F)$ is simply all possible unions of sets in $\mathcal{P}_F$.

Also, given a coarser partition $\mathcal{C} \geq \mathcal{P}$ we define $|\mathrm{Hom}_\mu(\mathcal{C}, F, \delta, \sigma)|_\mathcal{P}$ to be the number of maps $\mathcal{P} \to \{1, \dots d\}$ we can get by restricting maps in $\mathrm{Hom}_\mu(\mathcal{C}, F, \delta, \sigma)$ to $\mathcal{P}$.

The above definitions are a bit complicated, but for intuition one should think of it as requiring that for $A \in \mathcal{C}$ and $B \in \mathcal{C}_F$ we are requiring $(\sigma(s) \circ \psi)(A) \approx \psi(sA)$ and $m(\psi(A)) \approx \mu(A)$ "up to δ" accuracy.

Note that in the above definition $\mathcal{P}_F$ denotes the partition $\bigvee_{s \in F} s\mathcal{P}$ and *not* the partition $\bigvee_{s \in F} s^{-1}\mathcal{P}$ which we worked with in the amenable context.

Now, for a sofic group G we now fix a sofic approximation sequence $\{\sigma_i : G \to \mathrm{Sym}(d_i)\}_{i=1}^\infty$ and denote by Σ, and define the quantity

$$h_\Sigma(\mathcal{P}) = \inf_{\mathcal{C} \geq \mathcal{P}, F, \delta > 0} \limsup_{i \to \infty} \log(|\mathrm{Hom}_\mu(\mathcal{C}, F, \delta, \sigma_i)|_\mathcal{P}).$$

For the expression above we define $\log(0)$ to be $-\infty$. We can now define sofic entropy.

Definition 2.4.2. Let G be a sofic group with a sofic approximation sequence Σ. For a group action $G \curvearrowright (X, \mu)$ we define the sofic entropy of this action, relative to Σ as,

$$h_\Sigma(X, G) \coloneqq \sup_\mathcal{P} h_\Sigma(\mathcal{P}).$$

We will also denote the quantity $\limsup_{i \to \infty} \log(|\mathrm{Hom}_\mu(\mathcal{C}, F, \delta, \sigma_i)|_\mathcal{P})$ by $h_\Sigma(\mathcal{P}, \mathcal{C}, F, \delta)$ and $\inf_{F, \delta > 0} \limsup_{i \to \infty} \log(|\mathrm{Hom}_\mu(\mathcal{C}, F, \delta, \sigma_i)|_\mathcal{P})$ by $h_\Sigma(\mathcal{P}, \mathcal{C})$. Note that the quantity $|\mathrm{Hom}_\mu(\mathcal{C}, F, \delta, \sigma)|_\mathcal{P}$ is decreasing in $\mathcal{C}$. Indeed suppose that

$\mathcal{C} \leq \mathcal{C}_2$.
If $\psi \in \mathrm{Hom}(\mathcal{C}_2, F, \delta, \sigma)$ then by the triangle inequality,

$$\sum_{A \in \mathcal{C}_{1\,F}} |m(\psi(A)) - \mu(A)| \leq \sum_{A \in \mathcal{C}_{2\,F}} |m(\psi(A)) - \mu(A)| < \delta.$$

Furthermore, for sets $A_1, \ldots, A_n$ we have

$$(A_1 \cup \ldots \cup A_n) \vartriangle (sA_1 \cup \ldots \cup sA_n) \subset (A_1 \vartriangle sA_1) \cup \ldots \cup (A_n \vartriangle sA_n)$$

so we see that

$$\sum_{A \in \mathcal{C}_2} m(\sigma(s)(\psi(A)) \vartriangle \psi(sA)) \leq \sum_{A \in \mathcal{C}_1} m(\sigma(s)(\psi(A)) \vartriangle \psi(sA)) < \delta$$

as well. Hence we have a map $\mathrm{Hom}(\mathcal{C}_2, F, \delta, \sigma) \to \mathrm{Hom}(\mathcal{C}_1, F, \delta, \sigma)$, $\psi \mapsto \psi_{|\,\mathrm{alg}(\mathcal{C}_1)}$ which clearly becomes an injection when we identify algebra homomorphisms that agree on $\mathcal{P}$. Hence $|\,\mathrm{Hom}_\mu(\mathcal{C}_2, F, \delta, \sigma)|_{\mathcal{P}} \leq |\,\mathrm{Hom}_\mu(\mathcal{C}_1, F, \delta, \sigma)|_{\mathcal{P}}$. Similarly, $|\,\mathrm{Hom}_\mu(\mathcal{C}, F, \delta, \sigma)|_{\mathcal{P}}$ is also decreasing in F and δ and increasing in $\mathcal{P}$ so in the definitions of $h_\Sigma(\mathcal{P})$ and $h_\Sigma(X, G)$ we could equally well have taken limits instead of suprema and infima.

It is important to note that the definition of sofic entropy depends on the choice of sofic approximation sequence Σ, and in choosing a different approximation sequence we could well change the entropy.

To establish properties of sofic entropy we need couple of results, some of which we prove.

lem:Sterling

Lemma 2.4.3 ([KL16][pp. 233, Lemma 10.1]). *For $n \in \mathbb{N}$ one has*

$$e\left(\frac{n}{e}\right)^n \leq n! \leq en\left(\frac{n}{e}\right)^n.$$

Proposition 2.4.4 ([KL16][pp. 233, Proposition 10.2]). *Let $\epsilon > 0$ and $\mathcal{P} = \{A_1, A_2, \ldots, A_n\}$ be an ordered partition of a probability space (X, μ). Then there is a $\delta > 0$, such that for all sufficiently large $d \in \mathbb{N}$, the set of all ordered partitions with n members $\{V_1, \ldots, V_n\}$ satisfying $\sum_{i=1}^{n} |\frac{|V_i|}{d} - \mu(A_i)| < \delta$, has cardinality between $e^{(H(\mathcal{P})-\epsilon)d}$ and $e^{(H(\mathcal{P})+\epsilon)d}$.*

Proof. By the continuity properties of $H(\cdot)$ we can find a $\delta > 0$ such that for all ordered partitions $\mathcal{C} = \{B_1, \ldots, B_n\}$ of a probability space (Y, ν) satisfying $\sum_{i=1}^{n} |\nu(B_i) - \mu(A_i)| < \delta$ we have $|H(\mathcal{C}) - H(\mathcal{P})| < \epsilon/4$.

To get the lower bound in the proposition statement, note that for sufficiently large $d \in \mathbb{N}$ there exists $\lambda = (\lambda_1, \ldots, \lambda_n) \in \{0, 1/d, 2/d, \ldots, 1\}^n$ satisfying $\sum_{i=1}^{n} \lambda_i = 1$ and $\sum_{i=1}^{n} |\lambda_i - \mu(A_i)|$. Write W_λ for the set of all ordered n–partitions $\{V_1, \ldots, V_n\}$ of $\{1, \ldots, d\}$ such that $|V_i|/d = \lambda_i$. Applying Lemma 2.4.3 we have, for sufficiently large d,

$$|W_\lambda| = \frac{d!}{(\lambda_1 d)!(\lambda_2 d)! \ldots (\lambda_n d)!} \geq e^{1-n} d^{-n} \prod_{i=1}^{n} \lambda_i^{-\lambda_i d} \geq e^{-\epsilon d/4} e^{\log(\prod_{i=1}^{n} \lambda_i^{-\lambda_i d})}$$

$$= e^{-\epsilon d/4} e^{\sum_{i=1}^{n} -\lambda_i d \log \lambda_i} = e^{-\epsilon d/4} e^{dH(\mathcal{P})} \geq e^{-\epsilon d/4} e^{H(\mathcal{P})-\epsilon d/4} = e^{H(\mathcal{P})-\epsilon d/2}$$

This establishes the lower bound. To get the upper bound, denote by Λ the set of *all* tuples $\lambda \in \{1, \ldots n\}^n$ satisfying $\sum_{i=1}^{n} \lambda_i = 1$ and $\sum_{i=1}^{n} |\lambda_i - \mu(A_i)| < \delta$. Taking just the last criteria into account we see that $|\Lambda| \leq (1+2\delta d)^n$. If we, for $\lambda \in \Lambda$ define W_λ as before, a similar argument to the one above shows that $|W_\lambda| \leq e^{H(\mathcal{P}+\epsilon/2d)}$. Hence the set $\bigcup_{\lambda \in \Lambda} W_\lambda$ has cardinality at most $(1+2\delta d)^n e^{H(\mathcal{P}+\epsilon/2d)}$. For sufficiently large d this quantity is bounded by $e^{H(\mathcal{P})+\epsilon d}$. Since the set of all $n-$partitions $\{V_1, \ldots, V_n\}$ of $\{1, \ldots, d\}$ satisfying $\sum_{i=1}^{n} |\frac{|V_i|}{d} - \mu(A_i)| < \delta$ is just $\bigcup_{\lambda \in \Lambda} W_\lambda$ this establishes the upper bound in the proposition. ∎

lem:genhelp

Lemma 2.4.5 ([KL16][pp. 240, Lemma 10.13]). *Let $\mathcal{P}$ be a finite partition, F a finite subset of G containing the identity e, and let $\delta > 0$. Suppose $\mathcal{D}$ is an algebra of measurable subsets of X, such that the σ-algebra generated by $\mathcal{D}$ coincides with all measurable subsets, modulo null sets. Suppose also $\mathcal{P}$ is a partition in $\mathcal{D}$ and $\mathcal{C}$ is any finite partition refining $\mathcal{P}$. Then there is a partition $\mathcal{P}'$ in $\mathcal{D}$ that refines $\mathcal{P}$ and satisfies:*

$$h_\Sigma(\mathcal{P}, \mathcal{P}', F, \delta/4) \leq h_\Sigma(\mathcal{P}, \mathcal{C}, F, \delta).$$

In the theorem below we make repeated use of the basic facts that for sets A_i and B_i we have $(\bigcup_{i=1}^{n} A_i) \triangle (\bigcup_{i=1}^{n} B_i) \subset \bigcup_{i=1}^{n} (A_i \triangle B_i)$ and $(\bigcap_{i=1}^{n} A_i) \triangle (\bigcap_{i=1}^{n} B_i) \subset \bigcup_{i=1}^{n} (A_i \triangle B_i)$.

Theorem 2.4.6 ([KL16][pp. 241, Theorem 10.14]). *Let $\mathcal{P}$ be a generating partition of X. Then*

$$h_\Sigma(X, G) = h_\Sigma(\mathcal{P}, \mathcal{C})$$

for any partition $\mathcal{C} \geq \mathcal{P}$. In particular $h_\Sigma(X, G) = h_\Sigma(\mathcal{P})$.

Proof. Fix a partition $\mathcal{C}$ that refines $\mathcal{P}$. We will first show that $H(X, G) \leq h_\Sigma(\mathcal{P}, \mathcal{C})$.

For that, let $\mathcal{D}$ be an arbitrary partition, fix $\kappa > 0$ and let $\epsilon > 0$ to be determined in relation to κ. Now, the translates $s\mathcal{P}$, $s \in G$, generate a $\sigma-$algebra that agrees with the measurable subsets of X modulo null sets, hence any measurable subset can be approximated arbitrarily well by sets in the *algebra* generated by $\{s\mathcal{P} : s \in G\}$. Hence there is a finite set $K \subset G$ such that for every $A \in \mathcal{D}$ there is an $A' \in \mathrm{alg}(\mathcal{P}_K)$ satisfying

$$\mu(A \triangle A') < \epsilon/8. \tag{2.5}$$ {eq125}

Write $A' = \bigcup_{Y \in \Omega_A} \bigcap_{s \in K} sY_s$ for the appropriate collection Ω_A of maps $K \to \mathcal{P}$. By definition of $h_\Sigma(\mathcal{P}, \mathcal{C})$ we can find a finite set $F \subset G$ containing $K \cup \{e\}$ and a $\delta > 0$ such that

$$h_\Sigma(\mathcal{P}, \mathcal{C}, F, \delta) \leq h_\Sigma(\mathcal{P}, \mathcal{C}) + \kappa \tag{2.6}$$ {eq1}

We claim that for arbitrary $\sigma : G \to \mathrm{Sym}(d)$ with d large enough, and small enough $\delta > 0$, we will have

$$|\operatorname{Hom}(\mathcal{C} \vee \mathcal{D}, F, \delta, \sigma)|_{\mathcal{D}} \leq e^{\kappa d} |\operatorname{Hom}(\mathcal{C}, F, \delta, \sigma)|_{\mathcal{P}}. \tag{2.7}$$

{eq2}

Letting σ_i range over a sofic approximation, taking logarithms, then dividing by d_i and taking limit suprema on both sides of Equation (2.7) we will have $h_\Sigma(\mathcal{D}, \mathcal{C} \vee \mathcal{D}, F, \delta) \leq h_\Sigma(\mathcal{P}, \mathcal{C}, F, \delta) + \kappa$. Combining this with Equation (2.6) we then have $h_\Sigma(\mathcal{D}, \mathcal{C} \vee \mathcal{D}, F, \delta) \leq h_\Sigma(\mathcal{P}, \mathcal{C}) + 2\kappa$ yielding $H(X, G) \leq h_\Sigma(\mathcal{P}, \mathcal{C})$ since $\mathcal{D}$ was arbitrary.

It remains to establish Equation (2.7). For this consider the obvious map $\operatorname{Hom}(\mathcal{C} \vee \mathcal{D}, F, \delta, \sigma) \to \operatorname{Hom}(\mathcal{C}, F, \delta, \sigma)$, $\psi \mapsto \psi_{|\operatorname{alg}(\mathcal{C}_F)}$. It need not be an injection when we mod out by the relations of agreement on $\mathcal{D}$ and $\mathcal{P}$, respectively. However we will show that, modding out, it is sufficiently injective in a certain sense.

Note that if $\delta \leq \epsilon/8$ we will have

$$m(\psi(A \triangle A')) < \mu(A \triangle A') + \epsilon/8 < \epsilon/4.$$

for $\psi \in \operatorname{Hom}(\mathcal{C} \vee \mathcal{D}, F, \delta, \sigma)$. Here we have used that $A \triangle A' \in \operatorname{alg}((\mathcal{C} \vee \mathcal{D})_F)$ and **??**. Now, on $\operatorname{Hom}(\mathcal{C} \vee \mathcal{D}, F, \delta, \sigma)$ define the pseudometric
$p(\psi, \phi) := \max_{A \in \mathcal{D}} \frac{1}{d} |\psi(A) \triangle \phi(A)|$. Now, if $\psi, \phi \in \operatorname{Hom}(\mathcal{C} \vee \mathcal{D}, F, \delta, \sigma)$ agree on $\mathcal{P}$, and $\delta \leq \epsilon/8$, we have for $A \in \mathcal{D}$:

$$\frac{1}{d}|\psi(A) \triangle \phi(A)| \leq \frac{1}{d}(|\psi(A \triangle A')| + |\psi(A') \triangle \phi(A')| + |\psi(A' \triangle A)|)$$

$$\leq \epsilon/4 + \sum_{Y \in \Omega_A} \sum_{s \in K} \frac{1}{d} |\psi(sY_s) \triangle \phi(sY_s)| + \epsilon/4$$

$$\leq \epsilon/2 + \sum_{Y \in \Omega_A} \sum_{s \in K} \frac{1}{d}(|\psi(sY_s) \triangle \sigma(s)(\psi(Y_s))| + |\sigma(s)(\phi(Y_s)) \triangle \phi(sY_s)|) + \epsilon/8 + \epsilon/8$$

$$\leq \epsilon/2 + 2|\mathcal{P}_K||K|\delta < \epsilon$$

Hence, $p(\psi, \phi) < \epsilon$. It follows that the image of any ϵ separated set in $\operatorname{Hom}(\mathcal{C} \vee \mathcal{D}, F, \delta, \sigma)$ under the restriction map $\operatorname{Hom}(\mathcal{C} \vee \mathcal{D}, F, \delta, \sigma) \to \operatorname{Hom}(\mathcal{C}, F, \delta, \sigma)$ consists of homomorphisms that are all distinct when restricted to $\mathcal{P}$. Now for a set $V \subset \{1, \ldots, d\}$ there are at most $(1 + \epsilon d) \binom{d}{\lfloor \epsilon d \rfloor}$ sets W satisfying $m(V \triangle W)$. For large enough d, $(1 + \epsilon d) \binom{d}{\lfloor \epsilon d \rfloor}$ is less than $e^{\eta d}$ for some number η with $\eta \to 0$ as $\epsilon \to 0$. So for sufficiently small ϵ then every ϵ-ball in the p-pseudometric contains at most $e^{\kappa d}$ distinct restrictions to $\mathcal{P}$. From this Equation (2.7) is evident.

Now we show the reverse inequality, that $h_\Sigma(\mathcal{P}, \mathcal{C}) \leq H(X, G)$. Let $\kappa > 0$ and find a partition $\mathcal{D}$ refining $\mathcal{P}$, a finite set $F \subset G$ containing e and a $\delta > 0$ such that $h_\Sigma(\mathcal{P}, \mathcal{D}, F, \delta) \leq h_\Sigma(\mathcal{P}) + \kappa$. Since $\mathcal{P}$ is generating, we may, by replacing δ with $\delta/4$ and appealing to Lemma 2.4.5, assume that there is a finite set $E \subset G$

such that $\mathcal{D} \leq \mathcal{P}_E$.

Let $\sigma : G \to \mathrm{Sym}(d)$ and $0 < \delta' \leq \delta$. If we are given a $\psi \in \mathrm{Hom}(\mathcal{C}, FE, \delta', \sigma)$ we can restrict it to $\mathrm{alg}(\mathcal{D}_F)$ since $\mathcal{D}_F \leq (\mathcal{P}_E)_F = \mathcal{P}_{FE} \leq \mathcal{C}_{FE}$. We now argue that this restriction lies in $\mathrm{Hom}(\mathcal{D}, F, \delta, \sigma)$. For every $A \in \mathcal{D}$ we can write $A = \bigcup\limits_{Y \in \Omega_A} \bigcap_{s \in E} sY_s$ for some collection Ω_A of maps $E \to \mathcal{P}$. Then, for every $t \in F$,

$$\begin{aligned}
\sum_{A \in \mathcal{D}} \frac{1}{d} |\psi(tA) \vartriangle \sigma_t \phi(A)| &\leq \sum_{A \in \mathcal{D}} \sum_{Y \in \Omega_A} \sum_{s \in E} \frac{1}{d} |\psi(tsY_s) \vartriangle \sigma_t \psi(sY_s)| \\
&\leq \sum_{A \in \mathcal{D}} \sum_{Y \in \Omega_A} \sum_{s \in E} \frac{1}{d} (|\psi(tsY_s) \vartriangle \sigma_{ts} \psi(Y_s)| \\
&\quad + |\sigma_{ts}\psi(Y_s) \vartriangle \sigma_t \sigma_s \psi(Y_s)| + |\sigma_t(\sigma_s \psi(Y_s) \vartriangle \psi(sY_s))|)
\end{aligned}$$

All of the first terms occuring in the triple sum, $\frac{1}{d}|\psi(tsY_s) \vartriangle \sigma_{ts}\psi(Y_s)|$, are of the form $\frac{1}{d}|\psi(rB) \vartriangle \sigma_r \psi(B)|$ with $r \in FE$ and $B \in \mathrm{alg}(\mathcal{C}_{FE})$, so since $\psi \in \mathrm{Hom}(\mathcal{C}, FE, \delta', \sigma)$ their sum is dominated by δ'. For the same reason all of the third terms $|\sigma_t(\sigma_s\psi(Y_s) \vartriangle \psi(sY_s))| = |\sigma_s\psi(Y_s) \vartriangle \psi(sY_s)| < \delta'$ since $s \in E \subset FE$. When σ is a sufficiently good sofic approximation sequence we can also ensure that sum of all the middle terms is less than say $\delta/2$. Then

$$\sum_{A \in \mathcal{D}} \frac{1}{d} |\psi(tA) \vartriangle \psi(A)| < |\mathcal{D}||\mathcal{P}_E||E|\delta' + \delta/2 + |\mathcal{D}||\mathcal{P}_E||E|\delta' < \delta$$

provided we just chose $\delta' > 0$ small enough. Now, the triangle inequality implies that $\sum\limits_{A \in \mathcal{D}_F} |m(\psi(A)) - \mu(A)| < \delta' \leq \delta$.

This is all to say that when σ is a sufficiently good approximation and $\delta' > 0$ is small enough, we have a map $\mathrm{Hom}_\mu(\mathcal{C}, FE, \delta', \sigma) \to \mathrm{Hom}_\mu(\mathcal{D}, F, \delta, \sigma)$ given by restriction to $\mathrm{alg}(\mathcal{D}_F)$. When we identify algebra homomorphisms that agree on $\mathcal{P}$ the map becomes injective. Hence,

$$\begin{aligned}
h_\Sigma(\mathcal{P}, \mathcal{C}) &\leq \limsup_{i \to \infty} \frac{1}{d_i} \log |\mathrm{Hom}_\mu(\mathcal{C}, FE, \delta', \sigma_i)|_{\mathcal{P}} \\
&\leq \limsup_{i \to \infty} \frac{1}{d_i} \log |\mathrm{Hom}_\mu(\mathcal{D}, F, \delta, \sigma_i)|_{\mathcal{P}} \\
&\leq h_\Sigma(\mathcal{P}) + \kappa \leq h_\Sigma(X, G) + \kappa
\end{aligned}$$

Since κ was arbitrary, this proves the other inequality. ■

Just like in the amenable case, this generator theorem can be used to prove that for a sofic group G, the Bernoulli shift on $(\{1, \ldots, n\}, \nu^{\otimes G})$ has sofic entropy $h_\nu(\{\{1\}, \ldots, \{n\}\})$. This time however, the proof is not straightforward and requires some meticulous probabilistic arguments. We summarize the result in the theorem below.

Theorem 2.4.7 ([KL16][pp. 244, Theorem 10.15]). *Consider the discrete measure space $\{1, \ldots, n\}$ equipped with some probability measure ν. Let G be a sofic group. Then the action $G \curvearrowright \{1, \ldots, n\}^G$ defined by $(\alpha_g(x))_s = x_{g^{-1}s}$ has entropy $h_\nu(\{\{1\}, \{2\}, \ldots, \{n\}\}) = \sum\limits_{i=1}^{n} -\nu(\{i\}) \log(\nu(\{i\}))$.*

Note that any group action $\alpha : G \curvearrowright (X,\mu)$ induces a natural action $\alpha' : G \to L^2(X,\mu)$ by $\alpha'_s(f) = sf$. Here the function sf is defined by $(sf)(x) = f(s^{-1}x)$. Since α is assumed to be measure preserving, each operator α'_s on $L^2(X,\mu)$ is an isometry. Recalling that the entropy should measure how much the group "mixes" the measure space around we would expect that if each orbit of α' was precompact then the sofic entropy is 0 or $-\infty$. Indeed, this would give is reassurance that our definition of sofic entropy is good.
To prove this we will first need a combinatorial argument that is typical of sofic group theory:

lem:compactlemma

Lemma 2.4.8 ([KL16][pp. 247, Lemma 10.17])**.** *Let $\epsilon > 0$. Then there are a $\delta > 0$ and an $n \in \mathbb{N}$ such that, for all sufficiently large $d \in \mathbb{N}$, if $\sigma_1, \dots, \sigma_n$ are permutations of $\{1, \dots d\}$ such that*

$$|\{v \in \{1,\dots,d\} : \sigma_j(v) \neq \sigma_k(v) \text{ for } j \neq k\}| \geq (1-\delta)d,$$

then the number of sets $A \subset \{1,\dots,d\}$ with $|\sigma_j(A) \triangle A| < \delta d$ for all $j = 1,\dots,n$ is less than $e^{\epsilon d}$.

Proof. Suppose $\delta > 0$, and $d \in \mathbb{N}$ and that the permutations $\sigma_1, \dots, \sigma_n$ satisfy the hypothesis in the lemma. Let Γ denote the set of functions $\gamma : \{1,\dots,n\} \to \{0,1\}$ such that $|\sigma_j(\gamma^{-1}(1)) \triangle \gamma^{-1}(1)| < \delta d$ for $j = 1,\dots,n$. To prove the lemma it will suffice to find appropriate numbers $\delta > 0$ and $d, n \in N$ for which $|\Gamma|$ can be bounded by $e^{\epsilon d}$. Denote by H the subgroup of $\mathrm{Sym}(d)$ generated by the σ_i's and partition $\{1,\dots,d\}$ into minimal G-invariant sets. Let $V_1, V_2, \dots, V_m$ be among the sets in the partition with cardinality at least n. Then the set $W := \{1,\dots,d\} \setminus \bigcup_{i=1}^{n} V_i$ has no more than δd elements, by assumption on the σ_i's.

Now, given a $\gamma \in \Gamma$, let $B_\gamma = \cup_{j=1}^n (\sigma_j \gamma^{-1}(1) \triangle \gamma^{-1}(1))$. Then B_γ is just the set of $v \in \{1,\dots,n\}$ such that $\gamma(\sigma_j^{-1}(v)) \neq \gamma(v)$ for some $j \in \{1,\dots,n\}$, and clearly $|B_\gamma| \leq n\delta d$. Now let $\mathcal{B} := \{B_\gamma : \gamma \in \Gamma\}$. Since $|B_\gamma| \leq n\delta d$ for each γ it is clear that, if $n\delta < 1/2$, then $|\mathcal{B}| \leq (1 + n\delta d)\binom{d}{n\delta d}$. Now, for a $B \in \mathcal{B}$ note that any $\gamma \in \Gamma$ for which $B_\gamma = B$ will be constant on each V_i interval that does not intersect B. Indeed, B not intersecting V_i means that $\gamma(\sigma_j^{-1}(v)) = \gamma(v)$ for all $v \in V_i$. But fixing $v \in V_i$ we see that since V_i is a *minimal* H-invariant subset,

$$V_i = H(v) = \{\omega(v) : \omega \text{ is a word in the } \sigma_j\text{'s}\} \tag{2.8}$$

{eq:50}

so clearly γ will be constant on all of V_i. Let R_B be a set containing exactly one element from each V_i that does not intersect B.

Now note that any $\gamma \in \Gamma$ such that $B_\gamma = B$ is determined on the set $W \cup B \cup R_B$. Indeed, all of the V_i sets on which γ is constant all contain an element that lies in R_B, so γ's values on these V_i's can be deduced from γ's values on R_B. Finally, take a v belonging to a V_i set on which γ is *not* constant. Then V_i contains elements of $w \in B$ and by Equation (2.8) there is a word ω in the σ_j^{-1}'s such that $w = \omega(v)$. Pick a *shortest* word $\omega = \sigma_{j_m}^{-1}\sigma_{j_{m-1}}^{-1} \dots \sigma_{j_1}^{-1}$ such that $\omega(v) \in B$. Now consider the largest number k, $1 \leq k \leq m$, such that $\gamma((\sigma_{j_k}^{-1} \dots \sigma_{j_1}^{-1})(v)) = \gamma(v)$. If $k < m$ then of course $\gamma((\sigma_{j_{k+1}}^{-1}\sigma_{j_k}^{-1} \dots \sigma_{j_1}^{-1})(v)) \neq \gamma((\sigma_{j_k}^{-1} \dots \sigma_{j_1}^{-1})(v))$ showing that $(\sigma_{j_k}^{-1} \dots \sigma_{j_1}^{-1})(v) \in B$ which contradicts that ω was the shortest word such that $\omega(v) \in B$. Hence

$k = m$, i.e. $\gamma(v) = \gamma(\omega(v))$. In particular $\gamma(v)$ can be deduced from what values γ attains on $W \cup B \cup R_B$.

Having shown that any $\gamma \in \Gamma$ such that $B_\gamma = B$ is determined on the set $W \cup B \cup R_B$, it follows that for all $B \in \mathcal{B}$:

$$|\{\gamma \in \Gamma : B_\gamma = B\}| \leq 2^{|W \cup B \cup R_B|} \leq 2^{\delta d + n\delta d + \frac{d}{n}}.$$

Then,

$$|\Gamma| = |\bigcup_{B \in \mathcal{B}} \{\gamma \in \Gamma : B_\gamma = B\}| \leq |\mathcal{B}| 2^{\delta d + n\delta d + \frac{d}{n}} = (1 + n\delta d)\binom{d}{n\delta d} 2^{\delta d + n\delta d + \frac{d}{n}}.$$

By Lemma 2.4.3 we can choose n sufficiently large and then δ sufficiently small to ensure that this RHS is at most $e^{\epsilon d}$ for all large enough d. This completes the proof. ∎

thm:compactsofic

Theorem 2.4.9 ([KL16][pp. 247, Theorem 10.81])**.** *Suppose G is an infinite sofic group. For any compact action $G \curvearrowright (X, \mu)$ we have $h_\Sigma(X, G) = 0$ or $-\infty$.*

Proof. Fix a partition $\mathcal{P} = \{A_1, \ldots, A_m\}$ of (X, μ). Let $\epsilon > 0$ and choose a corresponding $\delta > 0$ and $n \in \mathbb{N}$ so that the statement in Lemma 2.4.8 holds. Set $f = \sum_{k=1}^{m} k \mathbb{1}_{A_k}$. Then the orbit Gf is totally bounded in $L^2(X, \mu)$ so there exists an *infinite* subset of G, say I, such that $\|sf - tf\|_2^2 < \frac{\delta}{3}$ for all $s, t \in I$. Picking a $t \in I$ and replacing I by $t^{-1}I$ we may assume that $e \in I$ so that $\|sf - f\|_2^2 < \frac{\delta}{3}$ for all $s \in I$. This implies that

$$\max_{k=1,\ldots,m} \mu(A_k \triangle sA_k) \leq \|sf - f\|_2^2 < \frac{\delta}{3} \text{ for all } s \in I.$$

Letting n be as above, choose a subset $F \subset I$ with cardinality n. Letting $\{\sigma_i : G \to \mathrm{Sym}(d_i)\}_i$ be a sofic approximation sequence for G we will, since such a sequence is asymptotically free, have

$$|\{v \in \{1, \ldots, d_i\} : \sigma_i(s)(v) \neq \sigma_i(t)(v) \text{ for } s, t \in F\}| \geq (1 - \delta)d.$$

for sufficiently large i. Note that, again by asymptotic freeness, since G is infinite we must have $d_i \to \infty$. In particular, for sufficiently large i, we have d_i large enough so that the hypothesis of Lemma 2.4.8 is satisfied. Then there can be at most $e^{\epsilon d_i}$ subsets with $|\sigma_s(B) \triangle B| < \delta d_i$. But given a $\phi \in \mathrm{Hom}(\mathcal{P}, F, \delta/3, \sigma_i)$, $s \in F$ and $A \in \mathcal{P}$ we have

$$\begin{aligned}\frac{1}{d}|\sigma_s\phi(A) \triangle \phi(A)| &\leq \frac{1}{d}|\sigma_s\phi(A) \triangle \phi(sA)| + \frac{1}{d}|\phi(sA \triangle A)| \\ &< \delta/3 + \mu(A \triangle sA) + \delta/3 < \delta.\end{aligned}$$

Thus $\mathrm{Hom}(\mathcal{P}, F, \delta/3, \sigma_i)$ has at most $e^{\epsilon d_i}$ restrictions to $\mathcal{P}$, i.e. $|\mathrm{Hom}(\mathcal{P}, F, \delta/3, \sigma_i)|_{\mathcal{P}} \leq e^{\epsilon d_i}$ for all sufficiently large i. In particular $h(\mathcal{P}, \mathcal{P}) \leq \epsilon$. Since ϵ and $\mathcal{P}$ were arbitrary, clearly $h(X, G) = 0$ or $-\infty$ as desired. ∎

In this chapter we have defined amenable- and sofic group entropy. A natural question is whether the sofic entropy of a group action $G \curvearrowright (X, \mu)$ coincides with the amenable entropy when G is amenable. After all, these entropies are defined in completely different ways so a priori their is no reason for them to agree. It turns out that the answer is yes, however, the entropies are the same, but this isn't so simple to prove. A proof can be found in [KL10][pp. 34, Theorem 6.7]. In particular then above theorem holds for amenable Bernoulli shifts as well. It also suggests the definition of sofic entropy, which appears as if from nowhere, is good.

CHAPTER 3

Dynamical Entropy

thir:third

3.1 Completely positive maps and conditional expectations

When defining entropy in the non-commutative setting, we need to find an analogue of partitions. A natural idea would be finite dimensional C*-subalgebras of a C*-algebra. It turns out there are two few of these to base the theory on them and instead we will work with images of finite dimensional C^* algebras under completely positive maps. We begin by defining them and then proving a few properties.

Definition 3.1.1. A map $\theta : A \to B$ of C*-algebras is said to be *completely positive* if, for all $n \in \mathbb{N}$, the tensor map $\theta \otimes \mathrm{id}_{M_n(\mathbb{C})} : A \otimes M_n(\mathbb{C}) \to B \otimes M_n(\mathbb{C})$ is positive.

Completely positive maps, being positive, are bounded. Indeed, if $P : A \to B$ is positive then the collection of positive linear functionals $\{\phi \circ P : \phi \in S(B)\}$ is a pointwise bounded family of bounded linear operators, so by the Uniform Boundedness principle, it is uniformly bounded in norm, say by $K \geq 0$. For $a \in A^+$ with $||a|| \leq 1$ we have,

$$||P(a)|| = \sup_{\phi \in S(B)} ||(\phi \circ P)(a)|| \leq K.$$

Hence P is bounded with $||P|| \leq 4K$. Furthermore, if A is unital $||P|| = |P(1)|$. We now list some immediate properties of completely positive maps.

prop:cp

Proposition 3.1.2.

(i) *The restriction of a c.p. map to a C*-subalgebra is also c.p.*

(ii) *If $\theta_1 : A_1 \to B_1$ and $\theta_2 : A_2 \to B_2$ are c.p. maps then,*

$$\theta_1 \oplus \theta_2 : A_1 \oplus A_2 \to B_1 \oplus B_2 \text{ is c.p.}$$

$$\theta_1 \otimes \theta_2 : A_1 \otimes A_2 \to B \otimes B_2 \text{ is c.p.}$$

(iii) *Suppose $\{A_i\}_{i \in I}$ is an increasing net of C*-subalgebras of A and $\theta_i : A_i \to B$ is a completely positive map for each $i \in I$ such that $\theta_{j|A_i} = \theta_i$ whenever $i \leq j$. Then there is a completely positive map $\theta : A \to B$ extending all the θ_i's.*

All elements in $A \otimes M_n(\mathbb{C})$ can be written in the form $\sum_{i,j=1}^{n} a_{ij} \otimes e_{ij}$ so a generic positive element looks like

$$(\sum_{i,j=1}^{n} a_{ij} \otimes e_{ij})^*(\sum_{k,l=1}^{n} a_{kl} \otimes e_{kl}) = (\sum_{i,j=1}^{n} a_{ji}^* \otimes e_{ij})(\sum_{k,l=1}^{n} a_{kl} \otimes e_{kl})$$

$$= \sum_{j=1}^{n} \sum_{i,l=1}^{n} a_{ji}^* a_{jl} \otimes e_{il}.$$

The take away from this is that an element in $A \otimes M_n(\mathbb{C})$ is positive if and only if it can be written as the sum of n elements of the form $\sum_{i,j=1}^{n} a_j^* a_i \otimes e_{ij}$ for some $a_1, a_2, \dots, a_n \in A$.

We now claim that the element $\sum_{i,j=1}^{n} b_{ij} \otimes e_{ij} \in B \otimes M_n(\mathbb{C})$ is positive if and only if $\sum_{i,j=1}^{n} b_i^* b_{ij} b_j \in B$ is positive for any $b_1, b_2, \dots, b_n \in B$. Now, an element a of a C*-algebra is positive if and only if $\phi(b) \geq 0$ for each pure state ϕ. Recalling the connection between pure states and irreducible Gelfand-Naimark-Segal (GNS) representations we then see that the positivity of $\sum_{i,j=1}^{n} b_{ij} \otimes e_{ij}$ is equivalent to having $\rho(\sum_{i,j=1}^{n} b_{ij} \otimes e_{ij}) \geq 0$ for each irreducible representation ρ on $B \otimes M_n(\mathbb{C})$. But these are all of the form $\pi \otimes \mathrm{id}_{M_n(\mathbb{C})}$ for an irreducible representation $\pi : B \to B(H)$. Such π's are cyclic so this is equivalent to

$$\langle (\pi \otimes \mathrm{id}_{M_n(\mathbb{C})})(\sum_{i,j=1}^{n} b_{ij} \otimes e_{ij})(\sum_{k=1}^{n} \pi(b_k)h \otimes e_k), \sum_{l=1}^{n} \pi(b_l)h \otimes e_l \rangle \geq 0$$

$$\text{for any } b_1, \dots, b_n \in B \text{ and } h \in H$$

$$\Leftrightarrow \langle \pi(\sum_{i,j=1}^{n} b_i^* b_{ij} b_j)h, h \rangle \geq 0 \text{ for any } b_1, \dots, b_n \in B \text{ and } h \in H.$$

This is equivalent to $\pi(\sum_{i,j=1}^{n} b_i^* b_{ij} b_j) \geq 0$ for any $b_1, \dots, b_n \in B$ and irreducible π, i.e. $\sum_{i,j=1}^{n} b_i^* b_{ij} b_j \geq 0$ for any $b_1, \dots, b_n \in B$.

Combining these two characterizations of positivity gives that θ is a completely positive exactly when

$$\sum_{i,j=1}^{n} b_i^* \theta(a_i^* a_j) b_j \geq 0 \text{ for any } a_1, \dots, a_n \in A \text{ and } b_1, \dots, b_n \in B.$$

There is another useful result for completely positive maps.

Theorem 3.1.3 ([NS06][pp. 265, A.2]). *If $\theta : A \to B$ is completely positive and $B \subset B(H)$, then there exists a Hilbert space K, a representation $\pi : A \to B(K)$ and a bounded operator $V : H \to K$ such that $\theta = V^*\pi(\cdot)V$ and $\overline{\pi(A)VH} = K$.*

The triple (K, π, V) is called the *Stinespring dilation* of θ. Conversely, any map which is of the form $a \mapsto V^*\pi(a)V$ for a bounded operator V is easily checked to be c.p. In particular, using the GNS-representation, all positive linear functional are of this form so they are completely positive. Upon further inspection we see that the above theorem actually generalizes the GNS-construction.

So far we have only proven characterizations of completely positive maps, nothing which indicates their existence. The below result tells us that they are common.

thm:abeliancp

Theorem 3.1.4 ([NS06][pp. 266, A.3]). *If $\theta : A \to B$ is a positive linear map, and either A or B is abelian, then θ is completely positive.*

Proof. Suppose B is abelian. By the Gelfand correspondence there is a compact Hausdorff space X such that $B \simeq C(X)$ as C*-algebras. For $n \in \mathbb{N}$ note that $B \otimes M_n(\mathbb{C}) \simeq C(X) \otimes M_n(\mathbb{C}) \simeq C(X; M_n(\mathbb{C}))$. Here $C(X; M_n(\mathbb{C}))$ denotes the C*-algebra of continuous functions $X \to M_n(\mathbb{C})$ where $M_n(\mathbb{C})$ is equipped with the norm topology. The last isomorphism is obtained by the map $f \otimes A \mapsto (x \otimes A \mapsto f(x)A)$. Under this identification of $B \otimes M_n(\mathbb{C})$ we have

$$(\theta \otimes \mathrm{id})(\sum_{i=1}^{k} a_i \otimes A_i)(x) = \sum_{i=1}^{k} \theta(a_i)(x)A_i = \sum_{i=1}^{k}((\chi_x \circ \theta) \otimes \mathrm{id})(a_i \otimes A_i)$$

$$= ((\chi_x \circ \theta) \otimes \mathrm{id})(\sum_{i=1}^{k} a_i \otimes A_i).$$

Here χ_x denotes evaluation at x. The above equation shows that to verify that θ is completely positive, it suffices to check that $\chi_x \circ \theta$ is completely positive for $x \in X$. These are all positive linear functionals so by the above remark they are c.p.

Now suppose A is abelian. Any positive linear map $\mathbb{C} \to B$ is a positive functional hence completely positive. By Proposition 3.1.2(ii) we then see that any positive, linear map $\mathbb{C}^n \to B$ is c.p. Appealing to Proposition 3.1.2(iii) we get that any abelian AF-algebra is completely positive. In the general case, consider the second transpose map $\theta^{**} : A^{**} \to B^{**}$. Viewing A^{**} as $\pi(A)''$ where π denotes the universal representations of A, we see that A is a weakly dense subalgebra of A^{**}, hence θ^{**} is also positive.

Since abelianness passes to the weak closure, we see that A^{**} is an abelian von Neumann algebra. Then it is AF: we can order the collection of finite sets of mutually orthogonal projections, denoted I, by declaring that for $\mathcal{P}, \mathcal{C} \in I$, $\mathcal{P} \leq \mathcal{C}$ if for all projections $q \in \mathcal{C}$ there is a $p \in \mathcal{P}$ with $q \leq p$. For $\mathcal{P}$ we let $A_{\mathcal{P}}$ denote the C*-subalgebra generated by $\mathcal{P}$. Then $\{A_{\mathcal{P}}\}_{\mathcal{P} \in I}$ is a net of finite dimensional subalgebras ordered under inclusion. By Borel functional calculus, we have $A^{**} = \overline{\bigcup_{\mathcal{P} \in I} A_{\mathcal{P}}}$. Thus A^{**} is an abelian AF-algebra so by the previous case θ^{**} is completely positive. By Proposition 3.1.2 (i) the restricted map $\theta^{**}_{|A} = \theta$ is c.p.. ■

When either the domain or the target space of a completely positive map is a full matrix algebra, $M_n(\mathbb{C})$, completely positive maps are related to states on tensor algebras of the domain and target space. More precisely, we have the following.

prop:p7

Proposition 3.1.5 ([NS06][pp. 267, A.6]). *Let $\theta : A \to M_n\mathbb{C}$ be a linear map and write $\theta(a) = \sum_{i,j=1}^{n} \theta_{ij}(a)e_{ij}$. Then the linear functional ϕ_θ on $A \otimes M_n(\mathbb{C})$ defined by $\phi_\theta(\sum_{i,j=1}^{n} a_{ij} \otimes e_{ij}) = \sum_{i,j=1}^{n} \theta_{ij}(a_{ij})$ is positive if and only if θ is completely positive. Thus we have a one to one correspondence between positive functionals on $A \otimes M_n(\mathbb{C})$ and c.p. maps $A \to M_n(\mathbb{C})$.*

Proof. Assume ϕ_θ is positive. Any non-degenerate representation of $M_n(\mathbb{C})$ is a direct sum of the identity representation, hence the GNS-representation of ϕ_θ is of the form $x \otimes a \mapsto \pi(x) \otimes a$ for some representation $\pi : A \to B(H)$. Suppose $h = \sum_{i=1}^{n} h_i \otimes e_i$ is the cyclic vector in the GNS-rep. of ϕ_θ. Then we obtain

$$\theta_{ij}(a) = \phi_\theta(a \otimes e_{ij}) = \langle(\pi(a) \otimes e_{ij})(\sum_{k=1}^{n} h_k \otimes e_k), \sum_{l=1}^{n} h_l \otimes e_l\rangle = \langle\pi(a)h_j, h_i\rangle.$$

Define $V : \ell^2(n) \to H$ by $Ve_i = h_i$. Then $V^*c = \sum_{i=1}^{n} \langle c, h_i\rangle e_i$. Hence,

$$V^*\pi(a)Ve_j = \sum_{i=1}\langle\pi(a)h_j, h_i\rangle e_i = \theta(a)e_j.$$

Thus $\theta(\cdot) = V^*\pi(\cdot)V$ so θ is completely positive.

Conversely, if θ is completely positive, it has a Stinespring dilation (H, π, V) where $V : \ell^2(n) \to H$. If we put $h = \sum_{i=1}^{n} Ve_i \otimes e_i$ we can check that $\phi_\theta(a \otimes x) = \langle(\pi(a) \otimes x)h, h\rangle$ showing that ϕ_θ is positive. ■

A similar result, but this time about c.p. maps $M_n(\mathbb{C}) \to A$, is the following.

prop:p8

Proposition 3.1.6 ([NS06][pp. 268, A.9]). *A linear map $\gamma : M_n(\mathbb{C}) \to A$ is completely positive if and only the element $T_\gamma = \sum_{i,j=1}^{n} \gamma(e_{ij}) \otimes e_{ij} \in A \otimes M_n(\mathbb{C})$ is positive.*

The above two propositions have some interesting consequences. For example, suppose A is a C*-subalgebra of a unital C*-algebra B and that A contains the unit of B. Then any completely positive map $\theta : A \to B(H)$ can be extended to a completely positive map $\bar{\theta} : B \to B(H)$. Indeed, for finite dimensional H, Proposition 3.1.5 gives us a state $\phi_\theta \in A \otimes B(H)$ corresponding to θ. Extend it, using Hahn-Banach, to a state $\phi \in B \otimes B(H)$ and use Proposition 3.1.5 backwards to identify this with a completely positive map $\bar{\theta} : B \to B(H)$. Since ϕ extends ϕ_θ, $\bar{\theta}$ will extend θ.

For general H, choose a net of finite rank projections, $\{p_i\}_{i\in I}$ such that $p_i \uparrow 1$ strongly. Consider the c.p. maps $A \to p_i B(H) p_i$, $a \mapsto p_i \theta(a) p_i$. Use the finite dimensional case to find completely positive extensions $B \to p_i B(H) p_i$. Since A is unital the norm of these extensions are bounded by $\|\theta\|$. Now the unit ball in $B(H)$, B_1, is compact when equipped with the weak operator topology. By Tychonoff's theorem, the set of functions from $B \to B_1$ is compact in the pointwise weak operator topology. This means that the net of extensions has a pointwise weak operator cluster point. This cluster point is the desired extension.

When A is a von Neumann algebra, say M, Proposition 3.1.6 allows us to approximate completely positive maps by completely positive, *normal* maps. We recall that these are ultraweakly continuous maps where the ultraweak topology on M is induced by the norm closure of $\mathrm{Span}\{\langle \cdot h, k\rangle : h, k \in H\} \subset M^*$.

Then the map $\theta : M \to M_n(\mathbb{C})$ in Proposition 3.1.6 is normal if and only if $\psi \circ \theta$ for each normal functional $\psi \in M_n(\mathbb{C})^*$. This is the same as the map $\langle \theta(\cdot) e_j, e_i\rangle$ being normal for each i and j, which is to say that θ_{ij} (as defined in Proposition 3.1.5) is normal. Similarly, a state $\phi \in (M \otimes M_n(\mathbb{C}))^*$ is normal if and only if $\phi(\cdot \otimes e_{ij})$ is normal for each i and j. We conclude that, in Proposition 3.1.5, θ is a normal map if and only if ϕ_θ is a normal state.

Now, when M^* is considered a real vector space, the set of normal states is convex. We claim its weak*-closure in M^* contains all states. If not, Hahn-Banach gives us a $\lambda \in \mathbb{R}$ and an element $a \in M$ such that

$$\mathrm{Re}(\psi(a)) \leq \lambda < \mathrm{Re}(\phi(a))$$

for all normal states ψ (see [Ped12][pp. 65, 2.4.7]). Letting $b = \frac{a+a^*}{2}$ yields

$$\psi(b) \leq \lambda < \phi(b) \leq \|b\|.$$

But if we just let ψ range over states of the form $\langle \cdot h, h\rangle$, $\psi(b)$ can approximate $\|b\|$ so this is absurd. Again using the correspondence in Proposition 3.1.5 this means that c.p. maps may be approximated by normal c.p. maps pointwise.

Here are two consequences of Proposition 3.1.6.

prop:cpmapapprox

Proposition 3.1.7 ([NS06][pp. 268, A.10])**.** *If B is a finite dimensional C*-algebra, $\{A_i\}_{i\in I}$ is an increasing net of unital C*-subalgebras with $\cup_i A_i$ norm dense in A, then any completely positive map $B \to A$ can be approximated in norm by c.p. maps $B \to \cup_i A_i$. If A is a strongly operator dense subalgebra of a von Neumann algebra M, any c.p. map $B \to M$ can be approximated in the pointwise strong operator topology by c.p. maps $B \to \cup_i A_i$ for $i \in I$. If γ is unital the approximations can be chosen to be unital.*

cor: liftingchannel

Corollary 3.1.8 ([NS06][pp. 269, A.11])**.** *If A is a C*-algebra, I a closed ideal in A, and $\gamma : M_n(\mathbb{C}) \to A/I$ is a c.p. map, then there exists a lifting of γ to a c.p. $\bar{\gamma} : M_n(\mathbb{C}) \to A$. If A is unital and γ is unital, the lifting $\bar{\gamma}$ can be chosen unital.*

Proof. We only prove the first part. Let $p : A \to A/I$ denotes the quotient map and write the positive operator $T_\gamma \in A/I \otimes M_n(\mathbb{C})$, given by Proposition 3.1.6, in the form $\sum_k \sum_{i,j} p(a_{k,j}^*)p(a_{i,j}) \otimes e_{ij}$. Setting $T = \sum_k \sum_{i,j} a_{k,j}^* a_{i,j} \otimes e_{ij} \in A \otimes M_n(\mathbb{C})$ we have $T_\gamma = (p \otimes \mathrm{id}) \circ T$. Hence, if we choose the c.p. map $\bar{\gamma} : M_n(C) \to A$ corresponding to T, $\bar{\gamma}$ will be a lifting of γ. ∎

We finish this section with a particularly nice class of completely positive maps, namely conditional expectations. These are much used in von Neumann algebra theory.

Definition 3.1.9. Let B a unital C*-subalgebra of a unital C*-algebra A. A unital, linear positive map $E : A \to B$ is a *conditional expectation* if $E(bab') = bE(a)b'$ for all $a \in A$ and $b, b' \in B$.

Evidently one has $\sum_{i,j} b_i^* E(a_i^* a_j) b_j = E((\sum_i a_i b_i)^*(\sum_i a_i b_i)) \geq 0$ for any $a_i \in A$ and $b_j \in B$ so conditional expectations are completely positive. This operator algebraic definition of conditional expectations that we have given here generalizes the classical conditional expectation $E(\cdot \mid \mathcal{C}) : L^\infty(X, \mathcal{A}, \mu) \to L^\infty(X, \mathcal{C}, \mu)$ where $\mathcal{C} \subset \mathcal{A}$ is a sub-σ-algebra of $\mathcal{A}$. For L^∞ algebras, we also have the following:

prop: abelianvonneumannalgebras

Proposition 3.1.10. *Let N be a sub-von Neumann algebra of $L^\infty(X, \mathcal{A}, \mu)$. Then there exists a sub-σ-algebra of $\mathcal{A}$, say $\mathcal{C}$, such that $N = L^\infty(X, \mathcal{C}, \mu_{|\mathcal{C}})$ when the RHS is viewed as subspace of $L^\infty(X, \mathcal{A}, \mu)$.*

Proof. Put $\mathcal{C} := \{C \in \mathcal{A} : [\mathbb{1}_C] \in N\}$. Since $1 \in N$, $\mathcal{C}$ is closed under complements and if $\{A_n\}_n \subset \mathcal{C}$, then by the Dominated Convergence Theorem $[\mathbb{1}_{\cup_{k=1}^n A_k}] \to [\mathbb{1}_{\cup_{k=1}^\infty A_k}]$ weakly. Hence $[\mathbb{1}_{\cup_{k=1}^\infty A_k}] \in N$ showing that $\mathcal{C}$ is a σ-algebra.

By construction the span of the projections in N coincides with the simple functions in $L^\infty(X, \mathcal{C}, \mu_{|C})$. But by Borel function calculus the former is norm-dense in N and by measure theory the latter is norm dense in $L^\infty(X, \mathcal{C}, \mu_{|C})$. Hence $N = L^\infty(X, \mathcal{C}, \mu_{|C})$.

∎

The above proposition can be used to produce conditional expectations from abelian von Neumann algebras M onto a sub von Neumann-algebra N: identify M with some $L^\infty(X, \mathcal{A}, \mu)$ space and N with $L^\infty(X, \mathcal{C}, \mu)$ for some $\mathcal{C} \subset \mathcal{A}$ and use $E(\cdot \mid \mathcal{C}) : M \to N$.

If $B(K) \subset B(H)$ is a finite dimensional matrix algebra we may think of H as $K \otimes K'$ for some Hilbert space K'. Under this identification $B(K) \simeq B(K) \otimes 1_{K'}$ where $1_{K'}$ is the identity on K'. Picking a state ϕ on $B(K')$ the map $\mathrm{id} \otimes \phi : B(K) \otimes B(K') \simeq B(H) \to B(K)$ gives a conditional expectation. We conclude that for any finite dimensional C*-algebra B of a C*-algebra A there will be a conditional expectation $A \to B$.

Von Neumann algebras are closed in the weak operator topology and are therefore closed under least upper bounds. That is, if $M \subset B(H)$ is a von Neumann algebra and $\{a_{i \in I}\} \subset M$ an increasing sequence of self adjoints with a least upper bound, denoted l.u.b., then the l.u.b. lies in M. Thus the class of maps that preserve l.u.b's are of interest. It turns out these are exactly the *normal* maps. In other words, a linear map $\theta : M \to N$ von Neumann algebras is normal if and only if for any monotone net $\{a_i\}_{i\in I}$ of self adjoint elements in M with a least upper bound, a, the net $\{\theta(a_i)\}_i$ has a least upper bound $\theta(a)$.

It turns out that whenever two (concrete) von Neumann algebras, $M \subset B(H)$ and $N \subset B(K)$ are isomorphic as C*-algebras, the isomorphism will be normal. To finish of the section, we give a discussion of normal conditional expectations $E : M \to N$ between von Neumann algebras. We claim that then there is a smallest projection $p \in M$ with the property that $E(p) = 1$. Indeed, consider the set $\mathcal{Q} = \{q \in M :\ q \text{ is a projection} : E(q) = 0\}$. Since E is normal and M is closed in the strong operator topology, any chain $\mathcal{C} \subset \mathcal{Q}$ has an upper bound. By Zorn's lemma it follows that $\mathcal{Q}$ has a maximal element $q \in M$. $p := 1 - q$ will then have the desired property. p is called the *support* of θ. In a similar vein we define the support of a positive element $a \in M$, denoted $s(a)$, to be the smallest projection $s(a) \in M$ for which $a \leq \|a\| s(a)$.

Note that $E(a) = E(pa)$ for any $a \in M$ because, for any state ϕ on N the CS-inequality applied to the sesquilinear form $(a, b) \mapsto \phi(b^* a)$ yields,

$$\begin{aligned}|(\phi \circ E)((1-p)a)| &\leq (\phi \circ E)((1-p)^*(1-p))^{1/2}(\phi \circ E)(a^* a)^{1/2} \\ &= \phi(E(1-p))^{1/2}(\phi \circ E)(a^* a)^{1/2} = 0.\end{aligned}$$

Similarly $E(a) = E(ap)$. Suppose now $a \in M$ is such that $s(a) \leq p$. We claim that $s(a) \leq s(E(a))$ and use the argument in [NS06][pp. 269]

We start by showing that p commutes with N. By Borel-functional calculus the span of unitaries is norm dense in N, so it suffices to show that a unitary $u \in N$ commutes with p. We have $E(upu^*) = uE(p)u^* = 1$ so by the defining property of p, $p \leq upu^*$. Replacing u by u^* we have also $p \leq u^* pu$. Combining the two yields

$$p \leq upu^* \leq uu^* puu^* \leq p.$$

Hence $p = upu^*$, as desired.

Now put $q = 1 - s(E(a))$. Then $s(a) \leq s(E(a))$ will follow if we can show that $qaq = 0$. Since q commutes with p, and $s(a) \leq p$, we have $qaq = qpaqp$, so that $s(qaq) \leq p$. By p's property, E is faithful on pMp so it suffices to check that $E(qaq) = 0$. This is obvious since $E(qaq) = qE(a)q = 0$.

3.2 State entropy

Let ϕ and ψ be two positive linear functionals on a finite dimensional C^*-algebra A. To begin the study of entropy on operator algebras we want to define the *mutual entropy* with respect to these states, denoted $S(\phi, \psi)$. In order to do this we need a one to one correspondence between positive linear functionals on A and positive elements of A. Namely, we claim that any positive linear functional ϕ on A is of the form $\mathrm{Tr}(\cdot\ Q_\phi) = \mathrm{Tr}(Q_\phi\ \cdot)$ for a unique positive element $Q_\phi \in A$ (Here Tr denotes the canonical trace on A which is the direct sum of the usual matrix traces when A is identified with a direct sum of matrix algebras). By linearity it suffices to verify this in the case $A \simeq \mathrm{Mat}_n \mathbb{C}$.

If the equation $\phi(\cdot) = \mathrm{Tr}(\cdot\ Q_\phi)$ (1) is to hold we can recover Q_ϕ by applying both sides to matrix units $e_{i'j'}$. Writing $Q_\phi = \sum_{i,j=1}^{n} a_{ij} e_{ij}$ for simplicity, we then have

$$\phi(e_{i'j'}) = \mathrm{Tr}(e_{i'j'}(\sum_{i,j=1}^{n} a_{ij} e_{ij})) = \mathrm{Tr}(\sum_j a_{a_{j'j}} e_{i'j}) = a_{j'i'}.$$

Hence $Q_\phi = (\phi(e_{ji})_{ij})$ and (1) then holds. Since ϕ is *-preserving we have $a_{ij} = \phi(e_{ji}) = \overline{\phi(e_{ij})} = \overline{a_{ji}}$ showing that Q_ϕ is symmetric. Let U^*DU be its diagonalization. Applying (1) to elements of the form $U^*e_{ii}U$ yields, $0 \leq \phi(U^*e_{ii}U) = \mathrm{Tr}(U^*e_{ii}UU^*DU) = \mathrm{Tr}(Ue_{ii}DU^*) = \mathrm{Tr}(e_{ii}D)$, which shows that D has positive entries on its diagonal, showing that $Q_\phi = U^*DU$ is positive.

In summary, the map $(A^*)^+ \mapsto A^+$, $\phi \mapsto Q_\phi$ is an injective, order-preserving isomorphism. Any map of the form $\mathrm{Tr}(\cdot Q)$ with $Q \geq 0$ is positive since $\mathrm{Tr}(B^*BQ^{1/2}Q^{1/2}) = \mathrm{Tr}(Q^{1/2}B^*BQ^{1/2}) = \mathrm{Tr}((BQ^{1/2})^*BQ^{1/2}) \geq 0$ for $B \in M_n(\mathbb{C})$. Hence the map is surjective.

We remark that the centralizer of ϕ, i.e. the set $Z_\phi := \{a \in A : \phi(ab) = \phi(ba)\ \forall b \in A\}$ consists precisely of those elements commuting with Q_ϕ. Indeed,

$$\begin{aligned} a \in Z_\phi &\Leftrightarrow \mathrm{Tr}(abQ_\phi - baQ_\phi) = 0 \text{ for all } b \in A \\ &\Leftrightarrow \mathrm{Tr}(bQ_\phi a - baQ_\phi) = 0 \text{ for all } b \in A \\ &\Leftrightarrow \mathrm{Tr}((Q_\phi a - aQ_\phi)^*(Q_\phi a - aQ_\phi)) = 0 \\ &\Leftrightarrow Q_\phi a - aQ_\phi = 0 \\ &\Leftrightarrow a \in Q'_\phi. \end{aligned}$$

Here we have used the Cauchy-Schwarz inequality applied to the sesquilinear form $(x, y) \mapsto \mathrm{Tr}(y^*x)$ to get the the third equivalence, and the faithfulness of $\mathrm{Tr}(\cdot)$ to get the fourth.

We furthermore note that a state ϕ is pure if and only if Q_ϕ is a one rank projection. Indeed, all pure states on a $K(H)$ algebra are of the form $\langle \cdot h, h\rangle$ for a unit vector $h \in H$. If p is the orthogonal projection onto $\mathrm{Span}\{h\}$, then $\langle \cdot h, h\rangle = \mathrm{Tr}(\cdot p)$ since we can compute the trace along an orthonormal basis containing h. Hence $Q_{\langle \cdot h, h\rangle} = p$.

We are now ready to define relative entropy of two states.

def: relativeentropy

Definition 3.2.1 ([NS06][pp. 15, Definition 2.11]). For positive linear functionals ϕ and ψ, the relative entropy of ϕ and ψ is

$$S(\phi, \psi) = \begin{cases} \mathrm{Tr}(Q_\phi(\log Q_\phi - \log Q_\psi)), & \text{if } \phi \leq \lambda\psi \text{ for some } \lambda > 0 \\ +\infty \text{ otherwise} \end{cases}$$

The condition that $\phi \leq \lambda\psi$ for some $\lambda > 0$ is equivalent to supp $\phi \leq$ supp ψ. Here supp ϕ denotes the smallest projection, p for which $\phi(p) = \phi(1)$.

The quantity $\mathrm{Tr}(Q_\phi(\log Q_\phi - \log Q_\psi))$ looks scary, but there are cases where it is easier to deal with. For example, suppose the two density matrices are simultaneously diagonalizable, i.e. there exists a unitary $U \in A$ and diagonal matrices D and D' such that $Q_\phi = U^*DU$ and $Q_\psi = U^*D'U$. For any $f \in C(\mathrm{spec}(Q_\phi))$ and automorphism α on A we have $f(\alpha(a)) = \alpha(f(a))$ (this is true when f is a polynomial and then by continuity for all f). Hence $\log(U^*DU) = U^*(\log D)U$ so,

$$S(\phi, \psi) = \mathrm{Tr}(Q_\phi(\log Q_\phi - \log Q_\psi))$$

$$= \mathrm{Tr}(U^*DU(\log(U^*DU) - \log(U^*DU))) = \mathrm{Tr}(U^*DU(U^*\log DU - U^*\log D'U))$$

$$= \mathrm{Tr}(D(\log D - \log D')) = \sum_i \lambda_i(\log \lambda_i - \log \mu_i),$$

where the λ_i's and μ_i's are the eigenvalues of Q_ϕ and Q_ψ, respectively.

We now list some basic properties of relative entropy. Recall that a linear map $\alpha : B \to A$ of C*-algebras is a Schwarz map if $\alpha(b^*b) \geq \alpha(b)^*\alpha(b)$ for $b \in B$.

thm:stateentropy **Theorem 3.2.2** ([NS06][pp. 16, Theorem 2.1.2]). *We have:*

(i) *$S(\phi, \psi) \geq 0$ for states ϕ and ψ with equality if and only if $\phi = \psi$.*

(ii) *$S(\phi, \psi)$ is decreasing in ψ.*

(iii) *$(\phi, \psi) \mapsto S(\phi, \psi)$ is lower semicontinuous and continuous on the closed sets of the form $\{(\phi, \psi) : \phi \leq \lambda\psi\}$ for $\lambda > 0$.*

(iv) *$(\phi, \psi) \mapsto S(\phi, \psi)$ is a convex function.*

(v) *if $\alpha : B \to A$ is a unital Schwarz map, then*

$$S(\phi \circ \alpha, \psi \circ \alpha) \leq S(\phi, \psi).$$

(vi) *if ϕ and ψ are states on A, B is a C*-subalgebra of A, and $E : A \to B$ is a a ψ-preserving conditional expectation, i.e. $\psi = \psi \circ E$, then*

$$S(\phi, \psi) = S(\phi_{|B}, \psi_{|B}) + S(\phi, \phi \circ E).$$

(vii) *For any decomposition $\phi = \sum_{i=1}^{n} \phi_i$ we have $\sum_i S(\phi_i, \psi) = s(\phi, \psi) + \sum_i S(\phi_i, \phi)$.*

Proof.

(i) Elementary calculus tells us that $\log t \leq t - 1$ with equality if and only if $t = 1$. Replacing t by $\frac{x}{y}$ we obtain that, for $x, y \geq 0$,

$$x(\log x - \log y) \geq x - y \tag{3.1}$$ {eq4}

with equality if and only if $x = y$. Now, Q_ϕ and Q_ψ are diagonalizable so we can decompose them as $Q_\phi = \sum_i \lambda_i p_j$ and $Q_\psi = \sum_i \mu_j q_j$ for mutually orthogonal projections $p_1, \ldots, p_n$, and mutually orthogonal projections $q_1, \ldots, q_n$. Using Equation (3.1), and assuming that $S(\phi, \psi) \neq \infty$, we obtain,

$$S(\phi, \psi) = \sum_{i,j} \lambda_i (\log \lambda_i - \log \mu_j) \operatorname{Tr}(p_i q_j) \geq \sum_{i,j} (\lambda_i - \mu_j) \operatorname{Tr}(p_i q_j) = 0$$

Here, equality occurs if and only if $\lambda_i = \mu_j$ for each i and j. In this case,

$$Q_\phi = \sum_j (\sum_i \lambda_i p_i) q_j = \sum_i (\sum_i \mu_j q_j) p_i = Q_\psi,$$

i.e. $\phi = \psi$.

(ii) log is an operator monotone function, so for $\psi_1 \leq \psi_2$ we have $Q_{\psi_1} \leq Q_{\psi_2}$ and then $\log Q_\phi \leq Q_\psi$. Hence $S(\phi, \psi_1) \geq S(\phi, \psi_2)$. Since η isn't operator monotone $S(\cdot, \cdot)$ is not increasing in the first argument.

(iii) We only prove semicontinuity so assume $\phi_n \to \phi$ and $\psi_n \to \psi$. Whenever ψ is faithful, we have, by the remark preceding Definition 3.2.1,

$$S(\phi, \psi) = \operatorname{Tr}(Q_\psi (\log Q_\phi - \log Q_\psi))) = -\operatorname{Tr}(\eta(Q_\phi)) - \phi(\log Q_\psi).$$

Similarly ψ_n will be faithful for sufficiently large n and

$$S(\phi_n, \psi_n) = -\operatorname{Tr}(\eta(Q_{\phi_n})) - \phi_n(\log Q_{\psi_n})$$

so we obtain $S(\phi_n, \psi_n) \to S(\phi, \psi)$.
Even if we drop the assumption that ψ is faithful, $S(\cdot, \cdot)$ is decreasing in the second argument so for $\epsilon > 0$,

$$S(\phi_n, \psi_n) \geq S(\phi_n, \psi_n + \epsilon \operatorname{Tr})$$

Taking lim inf on both sides and using what we proved in the faithful case we get,

$$\liminf_{n \to \infty} S(\phi_n, \psi_n) \geq S(\phi, \psi + \epsilon \operatorname{Tr}) = -\operatorname{Tr}(\eta(Q_\phi)) - \phi(\log(Q_\psi + \epsilon 1)) \quad (3.2)$$

{eq100}

We claim that $-\operatorname{Tr}(\eta(Q_\phi)) - \phi(\log(Q_\psi + \epsilon 1))$ converges to $S(\phi, \psi)$ as $\epsilon \to 0$. If $\operatorname{supp} \phi \leq \operatorname{supp} \psi$ this is obvious. If $\operatorname{supp} \phi \nleq \operatorname{supp} \psi$ note that the spectral projection $\mathcal{X}_{\{\epsilon\}}(Q_\psi + \epsilon 1)$ is the projection onto $\ker Q_\phi$. Hence $\mathcal{X}_{\{\epsilon\}}(Q_\psi + \epsilon 1) = 1 - \operatorname{supp} \psi$. We then have

$$\begin{aligned} -\operatorname{Tr}(\eta(Q_\phi)) - \phi(\log(Q_\psi + \epsilon 1)) &\geq -\phi(\log \epsilon \; \mathcal{X}_{\{\epsilon\}}(Q_\psi + \epsilon 1)) \\ &= -\log \epsilon \; \phi(1 - \operatorname{supp} \psi) \to \infty. \end{aligned}$$

This shows that also in this case we have

$$-\operatorname{Tr}(\eta(Q_\phi)) - \phi(\log(Q_\psi + \epsilon 1)) \to S(\phi, \psi).$$

Combining this with Equation (3.2) we get $\liminf_{n \to \infty} S(\phi_n, \psi_n) \geq S(\phi, \psi)$, establishing that the relative entropy is lower semicontinuous.

(iv) We omit the proof.

(v) We omit the proof.

(vi) We omit the proof.

(vii) We omit the proof.

■

One might wonder why we have restricted ourselves to finite dimensional C*-algebras. After all there are infinite dimensional C*-algebras having a unique trace, for instance the hyperfinite type Π_1 factor, which we denote by N. Denote this trace on N by tr. It is ultraweakly continuous. Note that, unlike the standard trace of $B(H)$, this is just a functional and does not attain the value ∞. In fact, if we have states ϕ and ψ on N which are of the form,

$$\phi = \mathrm{tr}(\cdot\ Q_\phi) \text{ and } \psi(a) = \mathrm{tr}(\cdot\ Q_\psi)$$

for operators $Q_\phi, Q_\psi \in N$, we could proceed as before and define the the relative entropy of ϕ and ψ to be the quantity $\mathrm{tr}(Q_\phi(\log Q_\phi - \log Q_\psi))$. Evidently states that are of this form are ultraweakly continuous, i.e. they are in the predual of N, denoted N_*. A problem is that not all states on N lie in N_* so defining entropy in this way cannot possibly be done for any pair of states.

Indeed, since N is a von Neumann alegebra it contains a non-trivial projection p_2. Since N is infinite dimensional either $p_2 N p_2$ or $(1-p_2)N(1-p_2)$ has dimension greater than 1. Suppose without loss of generality that $\dim((1-p_2)N(1-p_2)) \geq 2$. Then it contains a non-trivial projection p_3 satisfying $p_3 \leq 1 - p_2$. Containing this procedure we get mutually orthogonal projections p_n, $n \geq 2$. Setting $p_1 = 1 - \sum_n p_n$ we then have a unital embedding

$$\ell^\infty(\mathbb{N}) \ni (x_n)_n \mapsto \sum_n x_n p_n \in N.$$

But there are non-normal states on ℓ^∞, for example $(a_n + ib_n)_n \mapsto \limsup_n a_n + i \limsup_n b_n$, and under the embedding above these extend to non-normal states on N also. This argument of course applies to any infinite dimensional von Neumann algebra.

Definition 3.2.3 ([NS06][pp. 21, Definition 2.2.1])**.** The *von Neumann entropy* of a positive linear functional ϕ on a finite dimensional C*-algebra is

$$S(\phi) := \mathrm{Tr}(\eta(Q_\phi)) = -S(\phi, \mathrm{Tr})$$

Note that $S(\phi)$ can be computed in terms of Q_ϕ's eigenvalues; letting $Q_\phi = U^* D U$ be its diagonalization we have,

$$S(\phi) = \mathrm{Tr}(\eta(U^* D U)) = \mathrm{Tr}(U^* \eta(D) U) = \mathrm{Tr}(\eta(D)) = \sum_{i=1}^{n} \eta(\lambda_i) \tag{3.3}$$

{eq5}

Here the λ_i's are the eigenvalues of Q_ϕ counted with multiplicity. Von Neumann entropy is then a natural generalization of classical entropy; in the case where $A \simeq \mathbb{C}^n \simeq C(\{1,\dots,n\})$ and μ is a measure on $\{1,\dots,n\}$ the classical entropy of the partition consisting of singletons coincides with $S(\mu)$ in the definition above. If a finite dimensional C*-algebra is isomorphic to $\bigoplus_{i=1}^{k} M_{n_i}(\mathbb{C})$ we define rank $A := \sum_{i=1}^{k} n_i$. One can show that rank A is the dimension of any maximal abelian subalgebra of A. We have the following basic properties of von Neumann entropy.

thm:vnentropy

Theorem 3.2.4 ([NS06][pp. 21, Theorem 2.2.2])**.**

(i) $0 \le S(\phi) \le \log(\operatorname{rank} A)$ *for any state* ψ *and* $S(\phi) = 0$ *if and only if* ϕ *is pure, whereas* $S(\phi) = \log(\operatorname{rank} A)$ *if and only if* ϕ *is the normalized trace.*

(ii) $\phi \mapsto S(\phi)$ *is continuous, concave and we have*

$$S(\phi + \psi) \le S(\phi) + S(\psi).$$

(iii) *For any convex combination* $\phi = \sum_i \lambda_i \phi_i$ *of states, with* $\lambda_i \neq 0$ *for all* i*, we have,*

$$S(\phi) \ge \sum_i \lambda_i S(\phi_i, \phi),$$

with equality if and only if all the ϕ_i*'s are pure.*

(iv) *For states* ϕ *on* A*, and* ψ *on* B*, we have,*

$$S(\phi \otimes \psi) = S(\phi) + S(\psi).$$

(v) *if* ϕ *is a positive linear functional on* $A \otimes B \otimes C$*, then*

$$S(\phi) + S(\phi_{|B}) \le S(\phi_{|A\otimes B}) + S(\phi_{|B\otimes C}).$$

(vi) *if* ϕ *is a state on* $A \otimes B$*, then*

$$|S(\phi) - S(\phi_{|B})| \le S(\phi_{|A}).$$

(vii) *if* B *is a maximal abelian subalgebra of* A*, then* $S(\phi_{|B}) \ge S(\phi)$*, with equality if and only if* B *is in the centralizer,* Z_ϕ*, of* ϕ*.*

To prove this theorem we require the following lemma. Recall that for integers n and m we have $M_n(\mathbb{C}) \otimes M_m(\mathbb{C}) \simeq M_{nm}(\mathbb{C})$ where the isomorphism is given by

$$M_n(\mathbb{C}) \otimes M_m(\mathbb{C}) \ni A \otimes B \mapsto \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{n1}B & a_{n2}B & \dots & a_{nn}B \end{pmatrix} \in M_{nm}(\mathbb{C}).$$

We note that under this isomorphism one has $\operatorname{Tr}_{|M_n(\mathbb{C})} \otimes \operatorname{Tr}_{|M_m(\mathbb{C})} = \operatorname{Tr}_{|M_{nm}(\mathbb{C})}$.

lem:pure

Lemma 3.2.5 ([NS06][pp. 22, Lemma 2.2.3])**.**

(i) *if ϕ is a pure state on the algebra $M_n(\mathbb{C}) \otimes M_m(\mathbb{C}) \simeq M_{nm}(\mathbb{C})$, then $S(\phi_{|M_n(\mathbb{C})}) = S(\phi_{|M_m(\mathbb{C})})$.*

(ii) *Every state on $M_n(\mathbb{C}) \otimes 1 \subset M_n(\mathbb{C}) \otimes M_n(\mathbb{C})$ extends to a pure state on $M_n(\mathbb{C}) \otimes M_n(\mathbb{C})$.*

Proof. We will denote $\mathrm{Tr}_{|M_{nm}(\mathbb{C})}$ as Tr_{nm} and $\mathrm{Tr}_{|M_n(\mathbb{C})}$ and $\mathrm{Tr}_{|M_m(\mathbb{C})}$ as Tr_n and Tr_m respectively. Since ϕ is pure it is of the form $\langle \cdot h, h \rangle$ for some unit vector $h \in \mathbb{C}^n \otimes \mathbb{C}^m$. Write $h = \sum_{i,j} \lambda_{i,j} e_i \otimes e_j$. Then it is easy to verify that,

$$Q_\phi = \sum_{i,j,k,l} \lambda_{ij} \overline{\lambda_{kl}} e_{ik} \otimes e_{jl}.$$

We now claim that the density matrices of $\phi_{|M_n(\mathbb{C})}$ and $\phi_{|M_m(\mathbb{C})}$ are $(\mathrm{id} \otimes \mathrm{Tr}_m)(Q_\phi)$ and $(\mathrm{Tr}_n \otimes \mathrm{id})(Q_\phi)$. To verify the former, write $Q_\phi = \sum_r A_r \otimes B_r$ and note that for $A \in M_n(\mathbb{C})$ we have

$$\mathrm{Tr}_n((\mathrm{id} \otimes \mathrm{Tr}_m)(Q_\phi)A) = \mathrm{Tr}_n((\mathrm{id} \otimes \mathrm{Tr}_m)(\sum_r A_r \otimes B_r)A)$$

$$= \mathrm{Tr}_n((\sum_r \mathrm{Tr}_m(B_i)A_r)A) = \sum_r \mathrm{Tr}_n(A_r A)\,\mathrm{Tr}_m(B_r) = \sum_r (\mathrm{Tr}_n \otimes \mathrm{Tr}_m)((AA_i) \otimes B)$$

$$= \mathrm{Tr}_{nm}(Q_\phi(A \otimes 1)) = \phi(A \otimes 1).$$

If we define the $n \times m$ matrix $T = \sum_{i,j} \lambda_{ij} e_{ij}$ we can check that

$$(\mathrm{id} \otimes \mathrm{Tr}_m)(Q_\phi) = TT^*$$

and similarly $(\mathrm{Tr}_n \otimes \mathrm{id})(Q_\phi) = (T^*T)'$ where $(T^*T)'$ is the transpose of T^*T. Now, the nonzero eigenvalues of TT^* and $(T^*T)'$ are the same so by Equation (3.3) we have $S(\phi_{|M_n(\mathbb{C})}) = S(\phi_{|M_m(\mathbb{C})})$. Hence (i) is proved.

We first show (ii) in the case where $Q_\phi \in M_n(\mathbb{C})$ is a diagonal matrix. The trace of Q_ϕ is 1 since it is a rank one projection so if we write $Q_\phi = \sum \lambda_i e_{ii}$, then $\sum_i \lambda_i = 1$. Consider then the operator

$$P = \sum_{i,j} \lambda_i^{1/2} \lambda_j^{1/2} e_{ij} \otimes e_{ij} \in M_n(\mathbb{C}) \otimes M_n(\mathbb{C}).$$

We have $P^*P = P$ and,

$$\mathrm{Tr}_{nm}(P) = \sum_{i,j} \lambda^{1/2} \lambda_j^{1/2} \,\mathrm{Tr}_n(e_{ij})\,\mathrm{Tr}_n(e_{ij}) = \sum_i \lambda_i = 1.$$

P is therefore a rank one projection and the state associated to P, say $\overline{\phi} \in (M_n(\mathbb{C}) \otimes M_n(\mathbb{C}))^*$, is pure. Furthermore $Q_\phi = (\mathrm{id} \otimes \mathrm{Tr}_n)(P)$ so writing $P = \sum_r A_r \otimes B_r$ with $A_r \in M_n(\mathbb{C})$ and $B_r \in M_m(\mathbb{C})$ reveals that for $A \in M_n(\mathbb{C})$,

$$\overline{\phi}(A \otimes 1) = (\mathrm{Tr}_n \otimes \mathrm{Tr}_n)(P(A \otimes 1)) = (\mathrm{Tr}_n \otimes \mathrm{Tr}_n)((\sum_r A_r \otimes B_r)(A \otimes 1))$$

$$= \sum_r (\mathrm{Tr}_n \otimes \mathrm{Tr}_n)(A_r A \otimes B_r) = \sum_r \mathrm{Tr}_n(A_r A)\,\mathrm{Tr}_n(B) = \mathrm{Tr}_n(\sum_r \mathrm{Tr}_n(B_r) A_r A)$$

$$= \mathrm{Tr}_n((\mathrm{id} \otimes \mathrm{Tr}_n)(\sum_r A_r \otimes B_r)A) = \mathrm{Tr}_n((\mathrm{id} \otimes \mathrm{Tr}_n)(P) \otimes A)) = \phi(A)$$

Hence $\overline{\phi}$ is a pure state extending ϕ.

In the general case, let $Q_\phi = U^* DU$ be Q_ϕ's diagonalization. Let ψ be the state corresponding to D. By the above ψ has an extension $\overline{\psi}$ which is pure. Define the state $\overline{\phi} := (U^* \otimes U^*)\overline{\psi}(\cdot)(U \otimes U)$. It is pure since pure states are stable under unitary conjugation and it extends ϕ.

■

Proof of Theorem 3.2.4.

(i) By Equation (3.1) and Proposition 2.3.2(i) we see that $0 \le S(\phi) \le \log(\operatorname{rank} A)$ with $S(\phi) = 0$ if and only if $\eta(\lambda_i) = 0$ for each eigenvalue λ_i of Q_ϕ. This is to say that for each eigenvalue λ_i of Q_ϕ we have either $\lambda_i = 0$ or $\lambda_i = 1$. ϕ is a state so on the other hand $\sum_i \lambda_i = \mathrm{Tr}(Q_\phi) = \phi(1) = 1$. So we are forced to conclude that $\lambda_i = 1$ for one particular i and that the other eigenvalues are 0. Then $\phi = \langle \cdot h, h\rangle$ for a unit vector h, i.e. ϕ is pure.

By 3.1 and Proposition 2.3.2(ii) we have $S(\phi) = \log(\operatorname{rank} A)$ if and only if $\lambda_i = 1/(\operatorname{rank} A)$ for each i. The diagonalizing, we have $Q_\phi = U^* \frac{1}{\operatorname{rank} A} U$ for some unitary U. Then $Q_\phi = \frac{1}{\operatorname{rank} A} I$, i.e. ϕ is the canonical normalized trace.

(ii) Since $S(\phi) = -S(\phi, \mathrm{Tr})$ von Neumann entropy is concave by Theorem 3.2.2(iv). By Theorem 3.2.2(iii), $-S(\phi, \mathrm{Tr})$ is also known to be continuous on the sets $\{\phi : \phi \le \lambda \mathrm{Tr}\}$, but the union of these sets over all $\lambda > 0$ is the set of all positive linear functionals so $S(\phi)$ is continuous everywhere. The inequality in (ii) follows from the operator monotonicity of log:

$$S(\phi + \psi) \le -\phi(\log(Q_\phi + Q_\psi)) - \psi(\log(Q_\phi + Q_\psi))$$
$$\le -\phi(\log Q_\phi) - \psi(\log Q_\psi) = S(\phi) + S(\psi).$$

(iii) Note that,

$$\sum_i \lambda_i S(\phi_i, \phi) = \sum_i \lambda_i \phi_i(\log Q_{\phi_i} - \log Q_\phi)$$
$$= -\sum_i \lambda_i S(\phi_i) - (\sum_i \lambda_i \phi_i)(\log Q_\phi) = -\sum_i \lambda_i S(\phi_i) - \phi(\log Q_\phi)$$
$$= S(\phi) - \sum_i \lambda_i S(\phi_i) \le S(\phi)$$

This is an equality if and only if $S(\phi_i) = 0$ for all i, which, by (i), is to say that each ϕ_i is pure.

(iv) For $a \in A$ and $b \in B$ we have

$$\mathrm{Tr}((Q_\phi \otimes Q_\psi)(a \otimes b)) = \mathrm{Tr}_A(Q_\phi a)\,\mathrm{Tr}_B(Q_\psi b) = (\phi \otimes \psi)(a \otimes b).$$

This simple computation reveals that $Q_{\phi\otimes\psi} = Q_\phi \otimes Q_\psi$. Let $Q_\phi = U^* D_1 U$ and $Q_\psi = V^* D_2 V$ be diagonalizations. For diagonal matrices functional calculus reduces to applying functions to the diagonal entries. Hence, for diagonal D and D' we have $\log(DD') = \log(D) + \log(D')$. Therefore,

$$S(\phi\otimes\psi) = \mathrm{Tr}((D_1\otimes D_2)\log(D_1\otimes D_2)) = \mathrm{Tr}((D_1\otimes D_2)\log((D_1\otimes 1)(1\otimes D_2)))$$
$$= \mathrm{Tr}((D_1 \otimes D_2)(\log(D_1 \otimes 1) + \log(1 \otimes D_2)))$$
$$= (\mathrm{Tr}_A \otimes \mathrm{Tr}_B)((D_1 \otimes D_2)(\log(D_1) \otimes 1 + 1 \otimes \log(D_2))) = S(\phi) + S(\psi).$$

(v) To prove the inequality, note that

$$S(\phi_{|A\otimes B}) - S(\phi_{|B}) = -\phi_{|A\otimes B}(\log Q_{\phi_{|A\otimes B}}) + \phi_{|B}(\log Q_{\phi_{|B}})$$
$$= -\phi_{|A\otimes B}(\log Q_{\phi_{|A\otimes B}}) + \phi_{|A\otimes B}(\log(1_A \otimes Q_{\phi_{|B}}))$$
$$= -\phi_{|A\otimes B}(\log Q_{\phi_{|A\otimes B}} - \log(1_A \otimes Q_{\phi_{|B}})) = -S(\phi_{|A\otimes B}, \mathrm{Tr}_A \otimes\phi_{|B}).$$

Similarly,

$$S(\phi) - S(\phi_{|B\otimes C}) = -S(\phi, \mathrm{Tr}_A \otimes\phi_{|B\otimes C}).$$

Now $S(\phi, \mathrm{Tr}_A \otimes\phi_{B\otimes C}) \geq S(\phi_{|A\otimes B}, \mathrm{Tr}_A \otimes\phi_{|B})$ so applying Theorem 3.2.2(v) to the inclusion mapping $A \otimes B \to A \otimes B \otimes C$ gives (v).

(vi) Viewing ϕ as state on $A \otimes \mathbb{C} \otimes B$ (v) yields $S(\phi) + S(\phi_{|\mathbb{C}}) \leq S(\phi_{A\otimes\mathbb{C}}) + S(\phi_{\mathbb{C}\otimes B})$, i.e.

$$S(\phi) \leq S(\phi_{|A}) + S(\phi_{|B}).$$

To establish (iv), we have then to show

$$S(\phi_{|B}) \leq S(\phi) + S(\phi_{|A}) \tag{3.4}$$

{eq6}

We first prove Equation (3.4) this in the case $A \simeq M_n(\mathbb{C})$ and $B \simeq M_m(\mathbb{C})$. Let $C = A \otimes B$. By Lemma 3.2.5(ii) we may extend ϕ to a pure state, say $\bar{\phi}$ on $A \otimes B \otimes C$. Lemma 3.2.5(i) then implies that $S(\overline{\phi}_{|B\otimes C} = S(\overline{\phi}_{|A}))$ so if we apply (v) to $\overline{\phi}$ we get:

$$S(\overline{\phi}) + S(\overline{\phi}_{|B}) \leq S(\overline{\phi}_{|A\otimes B}) + S(\overline{\phi}_{|A})$$

But $\overline{\phi}$ is pure so $S(\overline{\phi}) = 0$ and the other terms can be rewritten in terms of ϕ. We get,

$$S(\phi_{|B}) \leq S(\phi) + S(\phi_{|A}).$$

So Equation (3.4) holds. To verify Equation (3.4) for general A and B, consider the decomposition of A into matrix algebras, say $\bigoplus_{i=1}^{k} M_{n_i}(\mathbb{C})$. Put $n = \operatorname{rank} A$ and $m = \operatorname{rank} B$. Then the embedding,

$$A \ni A_1 \oplus A_2 \oplus \ldots \oplus A_k \mapsto \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \ldots & 0 & A_k \end{pmatrix} \in M_n(\mathbb{C}) \tag{3.5}$$

is a trace preserving embedding Similarly we have trace preserving embeddings $B \to M_m(\mathbb{C})$ and $A \otimes B \to M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$. Extend ϕ

to a state $\tilde{\phi}$ on $M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$. If we write $Q_{\phi_{|A}} = Q_1 \oplus Q_2 \oplus \ldots \oplus Q_k$, then $Q_{\tilde{\phi}_{|M_n(\mathbb{C})}}$ will be of the form

$$\begin{pmatrix} Q_1 & ? & \ldots & ? \\ ? & Q_2 & \ldots & \vdots \\ \vdots & \vdots & \ddots & ? \\ ? & \ldots & ? & Q_k \end{pmatrix}$$

It follows that $S(\tilde{\phi}_{|M_n(\mathbb{C})}) = S(\phi_{|A})$. Similarly $S(\widetilde{\phi_{|M_m(\mathbb{C})}}) = S(\phi_{|B})$ and $S(\tilde{\phi}) = S(\phi)$ so Equation (3.4) follows from the full matrix algebra case.

(vii) Since B is maximal abelian we have $\mathrm{Tr}_{A|B} = \mathrm{Tr}_B$. Since Tr_A is faithful, tracial state, a general result that holds even for von Neumann algebras tells us there exists a Tr_A preserving conditional expectation $E_B : A \to B$. We have $Q_{\phi \circ E_B} = Q_{\phi_{|B}}$, because for $a \in A$:

$$\mathrm{Tr}_A(aQ_{\phi_{|B}}) = \mathrm{Tr}_A(E_B(aQ_{\phi_{|B}}))$$

$$= \mathrm{Tr}_A(E_B(a)Q_{\phi_{|B}}) = \mathrm{Tr}_B(E_B(a)Q_{\phi_{|B}})$$

$$= \phi_{|B}(E_B(a)) = (\phi \circ E_B)(a).$$

Hence,

$$S(\phi_{|B}) - S(\phi) = \phi(\log Q_\phi - \log Q_{\phi_{|B}}) = S(\phi, \phi \circ E_B) \geq 0. \tag{3.6}$$ {eq7}

The last inequality holds by Theorem 3.2.2(i) since ϕ and $\phi \circ E_B$ are both states, since E_B is unital. Hence $S(\phi_{|B}) \geq S(\phi)$. Suppose now $B \subset Z_\phi$. Since $Z_\phi = Q'_\phi$ we see then that all elements in B commute with Q_ϕ, but B is maximal abelian so then $Q_\phi \in B$. It then follows that $Q_{\phi_{|B}} = Q_\phi$, so patently $S(Q_{\phi_{|B}}) = S(\phi)$. Conversely, if $S(Q_{\phi_{|B}}) = S(\phi)$ Equation (3.6) reveals that $S(\phi, \phi \circ E_B) = 0$. Since E_B is unital both ϕ and $\phi \circ E_B$ are states so Theorem 3.2.2(i) we get $\phi = \phi \circ E_B$ so $Q_\phi = Q_{\phi \circ E_B} = Q_{\phi_{|B}} \in B$. Since B is abelian we then have $B \subset Q'_\phi = Z_\phi$.

■

3.3 Mutual entropy of channels

We will now make use of relative entropy $S(\cdot, \cdot)$ to define entropy in a non-commutative setting. More precisely given a C*-algebra, A, we want to define the entropy of a group action $G \to \mathrm{Aut}(A)$ with respect to a state, $\phi \in A^*$.

Note that this subsumes the classical case:

Given a classical group action $\alpha : G \to X$ we can define a corresponding action on $\beta : G \to \mathrm{Aut}(C(X))$ by $\beta_g(f) = f \circ \alpha^{-1}$. Conversely, given a group action $\beta : G \to C(X)$ each automorphism β_g is necessarily of the form $f \mapsto f \circ \alpha_g^{-1}$ for some homomorphism α_g on X. Hence there is a one to one correspondence between group actions on X and $C(X)$. We remark that every automorphism on $L^\infty(X, \mu)$ too is induced by a group action on (X, μ), but this is not so easy to show.

Before defining entropy in a non-commutative we want to define a quantity that generalizes the quantity $H_\mu(\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n)$ where $\mathcal{P}_1, \ldots, \mathcal{P}_n$ are partitions. It is reasonable that the state ϕ will replace the role of μ. One might conjecture that finite dimensional C*-subalgebras should play the role of partitions in the new setting. After all, any partition $\mathcal{P}$ of a probability space (X, μ) is in natural correspondence to the subalgebra of $L^\infty(X, \mu)$ of functions that are constant on members of $\mathcal{P}$. Since any von Neumann subalgebra of $L^\infty(X, \mu)$ is generated by its projections, any finite dimensional subalgebra will be of this form for some partition $\mathcal{P}$. However, it turns out that in a non-commutative C*-algebras we can have too few finite dimensional subalgebras so the theory will have to be based on something else. We will instead use unital completely positive maps γ from finite dimensional C*-algebras into A. We give them the following name:

Definition 3.3.1 ([NS06][pp. 34])**.** A *channel* of a C*-algebra A is a completely positive unital map $B \to A$ where B is some finite dimensional C*-algebra.

We are now ready to give an analogue of $H_\mu(\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n)$:

Definition 3.3.2 ([NS06][pp. 34, Definition 3.1.1])**.** Let A be a unital C*-algebra, ϕ a state on A, and $\gamma_k : A_k \to A$, $1 \leq k \leq n$, a collection of channels. Given finite index sets $I_1, I_2, \ldots, I_n$ and a decomposition $\phi = \sum_{(i_1, i_2, \ldots, i_n) \in I_1 \times \ldots \times I_n} \phi_{(i_1, i_2, \ldots, i_n)}$ we define the quantity

$$H(\gamma_1, \ldots, \gamma_n; \{\phi_{i_1, \ldots, i_n}\}) := \sum_{i_1, \ldots, i_n} \eta(\phi_{i_1, \ldots i_n}(1)) + \sum_k \sum_{i_k \in I_k} S(\phi_{i_k}^{(k)} \circ \gamma_k, \phi \circ \gamma_k). \tag{3.7}$$

{eq8}

Here $\phi_{i_k}^{(k)}$ is the sum of all the $\phi_{i_1, \ldots, i_n}$ where the k'th index equals i_k. The supremum of the quantities $H(\gamma_1, \ldots, \gamma_n; \{\phi_{i_1, \ldots, i_n}\})$ of all state decompositions $i_1, \ldots, i_n$ corresponding to so some finite index sets $I_1, \ldots, I_n$, $n \in N$, is the **mutual entropy** of the channels $\gamma_1, \ldots, \gamma_k$ with respect to ϕ. We denote it by $H_\phi(\gamma_1, \ldots, \gamma_k)$.

Note that in the above definition we work with *ordered* decompositions of the form $\phi = \sum_{(i_1, i_2, \ldots i_n) \in I_1 \times \ldots \times I_n} \phi_{(i_1, i_2, \ldots i_n)}$. In Equation (3.7) we see that the k'th index set, I_k, is associated to the k'th channel, γ_k. Using that

$$S(\lambda\psi, \phi) = \lambda S(\psi, \phi) - \psi(1)\eta(\lambda)$$

we can rewrite Equation (3.7) as

$$H(\gamma_1, \ldots, \gamma_n; \{\phi_{i_1, \ldots, i_n}\}) = \tag{3.8}$$

$$\sum_{i_1, \ldots, i_n} \eta(\phi_{i_1, \ldots, i_n}(1)) - \sum_k \sum_{i_k \in I_k} \eta(\phi_{i_k}^{(k)}(1)) + \sum_k \sum_{i_k \in I_k} \phi_{i_k}^{(k)}(1) S(\hat{\phi}_{i_k}^{(k)} \circ \gamma_k, \phi \circ \gamma_k) \tag{3.9}$$

{eq16}

Here $\hat{\psi}$ denotes $\psi(1)^{-1}\psi$. We can also rewrite the double sum in Equation (3.9) to get

$$H(\gamma_1,\dots,\gamma_n;\{\phi_{i_1,\dots,i_n}\}) = \tag{3.10}$$

$$\sum_{i_1,\dots,i_n}\eta(\phi_{i_1,\dots,i_n}(1)) - \sum_k\sum_{i_k\in I_k}\eta(\phi^{(k)}_{i_k}(1)) + \sum_k\sum_{i_k\in I_k}H_\phi(\gamma_k;\{\phi^{(k)}_{i_k}\}) \tag{3.11}$$

{eq56}

In the case where channels $\gamma_k : A_k \to A$, $1 \le k \le n$ are just inclusions of finite dimensional subalgebras of A we will denote $H_\phi(\gamma_1,\dots,\gamma_n)$ by $H_\phi(A_1,\dots,A_n)$.

Observe that the first sum in Equation (3.7) bears resemblance to the classical case. We will thus call this the classical term. Consider now the case of just one channel $\gamma : B \to A$. Then $H_\phi(\gamma)$ is just the supremum of the quantities

$$\sum_i \eta(\phi_i(1)) + \sum_i S(\phi_i\circ\gamma, \phi\circ\gamma) \tag{3.12}$$

{eq9}

for all possible state decompositions $\phi = \sum_i \phi_i$. One can show that the second sum in Equation (3.12) is 0 if and only if the $Q_{\gamma\circ\phi_i}$ matrices are orthogonal. Thus we can think of the second sum as compensating for a lack of orthogonality.

There is a natural way in which the decomposition of a state ϕ arises. Namely, suppose C is a finite dimensional C*-algebra and $C_1,\dots C_n \subset C$ are unital subalgebras. If μ is a state on C and P a unital positive map $A \to C$ such that $\phi = \mu\circ P$, we call the quadruple $(C,\{C_k\}_{k=1}^n,\mu,P)$ an **abelian model** for (A,ϕ).

Now, each C_k, being finite dimensional and abelian, has a unique collection of mutually orthogonal, minimal projections, say $\{p^{(k)}_{i_k} : i_k \in I_k\}$ for some index set I_k. This gives a natural decomposition $\phi = \sum\limits_{i_1,\dots,i_n}\phi_{i_1,i_2\dots,i_n}$ where

$$\phi_{i_1,i_2\dots,i_n} = \mu(P(\cdot)p_{i_1}p_{i_2}\dots p_{i_n}) \tag{3.13}$$

{eq10}

One can check that then

$$H(\gamma_1,\dots,\gamma_n;\{\phi_{i_1,\dots,i_n}\}) = S(\mu_{|\vee_k C_k}) + \sum_k\sum_{i_k\in I_k} S(\mu((P\circ\gamma_k)(\cdot)p^{(k)}_{i_k}),\phi\circ\gamma_k), \tag{3.14}$$

{eq11}

for any channels $\gamma_k : A_k \to A$, $1 \le k \le n$. We denote this quantity by $H(C,\mu,\{C_k\}_{k=1}^n,P)$.

Conversely, note that any decomposition $\phi = \sum\limits_{i_1,\dots i_n}\phi_{i_1,\dots,i_n}$ arises from an abelian model for (A,ϕ). Indeed, let $C = C(I_1\times I_2\times\dots\times I_n)$ and C_k be the algebra of functions only depending on the kth variable. Explicitly,

$$C_k = \{f\in C : f \text{ is constant on the sets } I_1\times\dots I_{k-1}\times\{i_k\}\times I_{k+1}\times\dots I_n \text{ for } i_k\in I_k\}.$$

Products of the form $p_{i_1}p_{i_2}\dots p_{i_n}$ then span C and we define

$$\mu(p_{i_1}p_{i_2}\dots p_{i_n}) = \phi_{i_1,\dots,i_n}(1).$$

Similarly define $P : A \to C$ by

$$P(a)(i_1,\ldots,i_n) = \begin{cases} \dfrac{\phi_{i_1,\ldots,i_n}(a)}{\phi_{i_1,\ldots,i_n}(1)} = \dfrac{\phi_{i_1,\ldots,i_n}(a)}{\mu(p_{i_1}p_{i_2}\ldots p_{i_n})} & \text{if } \phi_{i_1,\ldots,i_n}(1) \neq 0, \\ 0 \text{ otherwise.} \end{cases}$$

Then P is unital and positive and Equation (3.13) is satisfied so the abelian model gives rise to the decomposition we started with.

The following is an elementary result about abelian models.

lem: abelianmodels

Lemma 3.3.3 ([NS06][pp. 36, Lemma 3.1.2]). *Given channels $\gamma_k : A_k \to A$, $1 \le k \le n$ the mutual entropy $H_\phi(\gamma_1, \gamma_2, \ldots, \gamma_n)$ is the supremum of the entropies $H_\mu(P \circ \gamma_1, P \circ \gamma_2, \ldots, P \circ \gamma_n)$ where P ranges over all unital completely positive maps $P : A \to C$ where C is abelian, and over all states μ on C satisfying $\mu \circ P = \phi$.*

Our main reason for introducing abelian models is that they make some proofs significantly easier. For example in the proposition below.

prop: mutualentropy

Proposition 3.3.4 ([NS06][pp. 37, Proposition 3.1.3]). *For a collection of channels $\gamma_k : A_k \to A$, $1 \le k \le n$, we have the following:*

(i) *If B_k are finite dimensional C*-algebras and $\theta_k : B_k \to A_k$, are unital c.p. maps, $1 \le k \le n$, then*

$$H_\phi(\gamma_1 \circ \theta_1, \ldots, \gamma_n) \le H_\phi(\gamma_1, \ldots, \gamma_n),$$

(ii) *if $\theta : A \to B$ is a unital c.p. map and ψ a state on B, then*

$$H_\psi(\theta \circ \gamma_1, \ldots, \theta \circ \gamma_n) \le H_{\psi\circ\theta}(\gamma_1, \ldots, \gamma_n),$$

Equality holds if θ is an isomorphism of C-algebras.*

(iii) *if $k < n$, then $H_\phi(\gamma_1, \ldots, \gamma_k) \le H_\phi(\gamma_1, \ldots, \gamma_n)$*

$$H_\phi(\gamma_1, \ldots, \gamma_n) \le H_\phi(\gamma_1, \ldots, \gamma_k) + H_\phi(\gamma_{k+1}, \ldots, \gamma_n),$$

(iv) *$H_\phi(\gamma_1, \ldots, \gamma_n)$ depends only on the set $\{\gamma_1, \ldots, \gamma_n\}$.*

Proof.

(i) This amounts to showing that fpr a decomposition $\phi = \sum_{i_1,\ldots,i_n} \phi_{i_1,\ldots,i_n}$, we have

$$S(Q_{\phi\circ\gamma_k\circ\theta_k}, Q_{\phi_{i_1,\ldots,i_n}\circ\gamma_k\circ\theta_k}) \le S(Q_\phi \circ \gamma_k, Q_{\phi_{i_1,\ldots,i_n}\circ\gamma_k}).$$

But this follows from applying Theorem 3.2.2(v) to the Schwarz map θ_k.

(ii) Any decomposition $\psi = \sum_{i_1,\ldots,i_n} \psi_{i_1,\ldots,i_n}$ gives rise to the decomposition $\psi \circ \theta = \sum_{i_1,\ldots,i_n} \psi_{i_1,\ldots,i_n} \circ \theta$. By definition

$$H_\psi(\theta \circ \gamma_1, \ldots, \theta \circ \gamma_n; \{\phi_{i_1,\ldots,i_n}\}) = H_{\psi\circ\theta}(\gamma_1, \ldots, \gamma_n; \{\phi_{i_1,\ldots,i_n} \circ \theta\}),$$

so the inequality follows. If θ is an isomorphism then θ^{-1} is unital and completely positive so applying what we just proved yields

$$H_{\psi\circ\theta}(\gamma_1,\ldots,\gamma_n;\{\phi_{i_1,\ldots,i_n}\circ\theta\}) = H_{\psi\circ\theta}(\theta^{-1}\circ\theta\circ\gamma_1,\ldots,\theta^{-1}\circ\theta\circ\gamma_n)$$
$$\leq H_{\psi\circ\theta\circ\theta^{-1}}(\theta\circ\gamma_1,\ldots,\theta\circ\gamma_n)$$

which is what we need.

(iii) The equality $H_\phi(\gamma_1,\ldots,\gamma_k) \leq H_\phi(\gamma_1,\ldots,\gamma_n)$ follows from the fact that any decomposition $\phi = \sum_{i_1,\ldots,i_k} \phi_{i_1,\ldots,i_k}$ can be considered as a decomposition $\phi = \sum_{i_1,\ldots,i_n} \phi_{i_1,\ldots,i_n}$ with $I_{k+1},\ldots,I_n$ singletons. Then $H_\psi(\theta\circ\gamma_1,\ldots,\theta\circ\gamma_n;\{\phi_{i_1,\ldots,i_k}\}) = H_{\psi\circ\theta}(\gamma_1,\ldots,\gamma_n;\{\phi_{i_1,\ldots,i_n}\})$, so the inequality follows.

To show the the other inequality, let $(C,\mu,\{C_j\}_{j=1}^n,P)$ be an abelian model for $(A,\phi,\{\gamma_j\}_{j=1}^n)$.

Then patently $(C,\mu,\{C_k\}_{j=1}^k,P)$ and $(C,\mu,\{C_k\}_{j=k+1}^n,P)$ are abelian models for $(A,\phi,\{\gamma_j\}_{j=1}^k)$ and $(A,\phi,\{\gamma_j\}_{j=k+1}^n)$. Let $\{p_{i_k}^{(k)} : i_k \in I_k\}$ be the atoms of C_k. By Equation (3.14) we have then only to show that $S(\mu_{|\vee_{j=1}^n}) \leq S(\mu_{|\vee_{j=1}^k}) + S(\mu_{|\vee_{j=k+1}^n})$ but this is just Proposition 2.3.2 applied to the probability space $\{1,\ldots,n\}$.

(iv) To show this we need only show that $H_\phi(\gamma_1,\ldots,\gamma_n) = H_\phi(\sigma(\gamma_1),\ldots,\sigma(\gamma_n))$ for any permutation σ on $\{1,\ldots,n\}$, and show that $H_\phi(\gamma_1,\ldots,\gamma_n) = H_\phi(\gamma_1,\ldots,\gamma_n)$. The first part is obvious and we omit the proof of the second part.

■

lem: subalgentropy

Lemma 3.3.5 ([NS06][pp. 39, Lemma 3.1.5]). *If the images of channels $\gamma_k : A_k \to A$, $1 \leq k \leq n$, are contained in a unital C^*-subalgebra $B \subset A$, and there exists a ϕ-preserving conditional expectation $E : A \to B$, then*

$$H_{\phi_{|B}}(\gamma_1,\ldots,\gamma_n) = H_\phi(\gamma_1,\ldots,\gamma_n).$$

Proof. Given a decomposition $\phi = \sum_{i_1,\ldots i_n} \phi_{i_1,\ldots i_n}$ we clearly have $Q_{\phi\circ\gamma_k}, Q_{\phi_{i_1,\ldots i_n}\circ\gamma_k} \in B$ so $H_\phi(\gamma_1,\ldots,\gamma_n;\{\phi_{i_1,\ldots i_n}\}) = H_\phi(\gamma_1,\ldots,\gamma_n;\{\phi_{i_1,\ldots i_n|B}\})$. Given a decomposition $\psi = \sum_{i_1,\ldots i_n} \psi_{i_1,\ldots i_n}$ we can compose both sides with E to get a state decomposition of ϕ. E is the identity on B so restricting this state decomposition to B we get $\sum_{i_1,\ldots i_n} \phi_{i_1,\ldots i_n}$. This completes the proof. ■

lem: commutingsubalgs

Lemma 3.3.6 ([NS06][pp. 39, Proposition 3.1.6]). *Let $A_1,\ldots,A_n \subset B \subset A$. Suppose there exists mutually commuting abelian subalgebras $C_k \subset A_k$, $1 \leq k \leq n$, such that $\vee_k C_k$ is maximal abelian in the centralizer of $\phi_{|B}$. If there also exists a ϕ-preserving conditional expectation $E : A \to B$, then*

$$H_\phi(A_1,\ldots,A_n) = H_\phi(B) = S(\phi_{|B})$$

Proof. There exists a ϕ-preserving conditional expectation $F : B \to C := \vee_k C_k$. Then $(C, \phi_{|C}, \{C_k\}_k, F \circ E)$ is an abelian model for $(A, \phi, \{C_k\})$. But $F \circ E$ is the identity on $C_k \subset C$ so the entropy of this model is the same as the entropy of the model $(C, \phi_{|C}, \{C_k\}_k, \mathrm{id}_{|C})$ for $(C, \phi_{|C}, \{C_k\}_k)$ and this is of course $S(\phi_{|C})$. By Theorem 3.2.4(vii) this again equals $S(\phi_{|B})$. So

$$H_\phi(C_1, \dots, C_n) \geq S(\phi_{|B})$$

On the other hand, Proposition 3.3.4(i) applies to the inclusion maps yields

$$\begin{aligned} H_\phi(C_1, \dots, C_n) &\leq H_\phi(A_1, \dots, A_n) \leq H_\phi(\oplus_k A_k, \dots, \oplus_k A_k) \leq H_\phi(\oplus_k A_k) \\ &\leq H_\phi(B) \leq S(\phi_{|B}). \end{aligned}$$

Then the above inequalities are in fact equalities. ∎

For a state ϕ on a C*-algebra A, let $\pi_\phi : A \to B(H_\phi)$ be its corresponding GNS-representation with cyclic unit vector $h_\phi \in H_\phi$. It is natural to consider the vector state $\langle \cdot\, h_\phi, h_\phi \rangle$ on $\pi_\phi(A)''$. It extends ϕ in the sense that $\phi(a) = \langle \pi_\phi(a) h_\phi, h_\phi \rangle$ for $a \in A$ and we will denote it by $\overline{\phi}$. It is natural to expect that mutual entropy of ϕ and $\overline{\phi}$ to be related. The proposition below tells us that this, and that a plethora of other results hold. They will all be useful when defining the entropy of an amenable group action $\alpha : G \to \mathrm{Aut}(A)$.

prop: channelproperties

Proposition 3.3.7 ([NS06][pp. 40, Proposition 3.1.7])**.**

(i)
$$H_\phi(\gamma_1, \dots, \gamma_n) = H_{\overline{\phi}}(\pi_\phi \circ \gamma_1, \dots, \pi_\phi \circ \gamma_n),$$

(ii) *if ψ is another state on A and $0 \leq \lambda \leq 1$, then*
$$\begin{aligned} &H_{\lambda\phi + (1-\lambda)\psi}(\gamma_1, \dots, \gamma_n) \\ &\geq \lambda H_\phi(\gamma_1, \dots, \gamma_n) + (1-\lambda) H_\psi(\gamma_1, \dots, \gamma_n) - (n-1)(\eta(\lambda) + \eta(1-\lambda)), \end{aligned}$$

(iii) *if ψ is a state on another C*-algebra B, $0 \leq \lambda \leq 1$ and $\theta_k : B_k \to B$, $1 \leq k \leq n$ are channels, then on $A \oplus B$*
$$\begin{aligned} &H_{\lambda\phi \oplus (1-\lambda)\psi}(\gamma_1 \oplus \theta_1, \dots, \gamma_n \oplus \theta_n) \\ &\geq \lambda H_\phi(\gamma_1, \dots, \gamma_n) + (1-\lambda) H_\psi(\gamma_1, \dots, \gamma_n) + \eta(\lambda) + \eta(1-\lambda), \end{aligned}$$

(iv) *under the assumptions of (iii),*
$$H_{\phi\otimes\psi}(\gamma_1 \otimes \theta_1, \dots, \gamma_n \otimes \theta_n) \geq H_\phi(\gamma_1, \dots, \gamma_n) + H_\psi(\theta_1, \dots, \theta_n),$$
and equality holds if B is abelian and the θ_k's are injective homomorphisms.

(v) *For channels $\gamma_k : A_k \otimes B_k \to A$, $1 \leq k \leq n$,*
$$H_\phi(\gamma_1, \dots, \gamma_n) \leq H_\phi(\gamma_{1|A_1}, \dots, \gamma_{n|A_n}) + 2 \sum_k S(\phi \circ \gamma_{k|B_k}).$$

Proof.

(i) For $1 \le k \le n$, let $[\gamma_k]$ denote the channel γ_k composed with the quotient map $A \to A/\ker\phi$. Similarly let $[\phi]$ denotes the state ϕ induces on $A/\ker\phi$. It is obvious that $H_\phi(\gamma_1,\dots,\gamma_n) = H_{[\phi]}([\gamma_1],\dots,[\gamma_n])$. Letting $[\pi_\phi]$ be the map π_ϕ induces on $A/\ker\phi$, we see that $[\pi_\phi] : A/\ker\phi \to B(H_\phi)$ is an isomorphism onto its image. By Proposition 3.3.4 (ii),

$$H_{[\phi]}([\gamma_1],\dots,[\gamma_n]) = H_{\bar\phi_{|\pi_\phi(A)}\circ[\pi_\phi]}([\gamma_1],\dots,[\gamma_n])$$

$$= H_{\overline{\phi}_{|\pi_\phi(A)}}([\pi_\phi]\circ[\gamma_1],\dots,[\pi_\phi]\circ[\gamma_n]) = H_{\overline{\phi}_{|\pi_\phi(A)}}(\pi_\phi\circ\gamma_1,\dots,\pi_\phi\circ\gamma_n)$$

.

Now, from general theory any postive linear functional $\psi \le \overline{\phi}_{|\pi_\phi(A)}$ is actually of the form $\langle\cdot\, ah_\phi,\phi\rangle$ for some $a \in \pi_\phi(A)'$. Thus any state decomposition $\overline{\phi}_{|\pi_\phi(A)} = \sum_{i_1,\dots,i_n} \psi_{i_1,\dots i_n}$ extends to a state decomposition of $\overline{\phi}$. Hence the above is just $H_{\overline{\phi}}(\pi_\phi\circ\gamma_1,\dots,\pi_\phi\circ\gamma_n)$.

(ii) We omit the proof.

(iii) We omit the proof.

(iv) Decompositions $\phi = \sum_{i_1\in I_1,\dots,i_n\in I_n} \phi_{i_1,\dots,i_n}$ and $\psi = \sum_{j_1\in J_1,\dots,j_n\in J_n} \psi_{j_1,\dots,j_n}$ give rives to the decomposition

$$\phi\otimes\psi = \sum_{(i_1,j_1)\in I_1\times J_1,\dots,(i_n,j_n)\in I_n\times J_n} \phi_{i_1,\dots,i_n}\otimes\psi_{j_1,\dots,j_n}$$

Properties of η and techniques we used in proving Theorem 3.2.4 (iv) show that

$$H_{\phi\otimes\psi}(\gamma_1\otimes\theta_1,\dots,\gamma_n\otimes\theta_n;\{\phi_{i_1,\dots,i_n}\otimes\psi_{j_1,\dots,j_n}\}_{(i_1,j_1),\dots,(i_n,j_n)})$$

$$= H_\phi(\gamma_1,\dots,\gamma_n;\{\phi_{i_1,\dots,i_n}\}_{i_1,\dots,i_n}) + H_\psi(\theta_1,\dots,\theta_n;\{\psi_{j_1,\dots,j_n}\}_{j_1,\dots,j_n})$$

which completes the proof of the inequality.

Assume now that B is abelian and the θ_ks are injective homomorphisms. Clearly we can assume $B_k \subset B$ and that the θ_k's are inclusion maps. Finite dimensional C*-algebras are of course v.n. algebras so by Proposition 3.1.10 we have a conditional expectation $B \to \vee_k B_k$.

By Lemma 3.3.5 we may then assume that $\vee_k B_k = B$. Let $\{p_j\}_{j=1}^r$ be the atoms of B, $\{\mathcal{X}_j\}_{j=1}^r$ the corresponding characters. For some $m \in \mathbb{N}$ we have $A\otimes B \simeq A\otimes\mathbb{C}^m$ which can be identified with $\oplus_{j=1}^m A$ under the map $a\otimes(b_1,\dots,b_m) \to (b_1a,\dots,b_ma)$. Under this identification $\phi\otimes\psi = \oplus_j\psi(p_j)\phi$ and

$$\gamma_k(a)\otimes\theta_k(b) = (\mathcal{X}_1(b)\gamma_k(a),\dots,\mathcal{X}_m(b)\gamma_k(a))$$

Thus the channels $\gamma_k\otimes\theta_k$ factorize through the channel

$$\oplus_{j=1}^m\gamma_k : \oplus_j A_k \to \oplus_{j=1}^m A$$

Hence, by part (iii) of this proposition, and Proposition 3.3.4 (i) we have

$$
\begin{aligned}
H_{\phi\otimes\psi}(\gamma_1 \otimes \theta_1, \dots, \gamma_n \otimes \theta_n) &\le H_{\oplus_j \psi(p_j)\phi}(\oplus_j \gamma_1, \dots, \oplus_j \gamma_n) \\
\sum_j \psi(p_j) H_\phi(\gamma_1, \dots, \gamma_n) &+ \sum_j \eta(\psi(p_j)) \\
&= H_\phi(\gamma_1, \dots, \gamma_n) + S(\psi)
\end{aligned}
$$

Now $S(\psi) = H_\psi(B_1, \dots, B_n)$ by Lemma 3.3.6 so we are done.

(v) We omit the proof.

■

Since the definition of $H_\phi(\gamma_1, \dots, \gamma_n)$ involves taking the suprema over all possible state decompositions it is not obvious that it is continuous in the n variables $\gamma_1, \dots, \gamma_n$. We will show that even more is true, namely a form of uniform continuity. Before stating it we give a definition. Given channels $\gamma, \gamma' : B \to A$ we define

$$\|\gamma - \gamma'\|_\phi = \sup_{b \in B,\ \|b\| \le 1} \phi(((\gamma - \gamma')(b))^*((\gamma - \gamma')(b)))^{1/2}.$$

On the space of channels $B \to A$, $\|\cdot\|_\phi$ is a seminorm.

prop: uniformchannels

Proposition 3.3.8 ([NS06][pp. 46, Proposition 3.1.11]). *For every $\epsilon > 0$ and $d \ge 1$ there exists $\delta > 0$ such that for any C^*-algebra A with a state ϕ, $n \in \mathbb{N}$, and channels $\gamma_k, \gamma_k' : A_k \to A$ such that $\dim A_k \le d$ and $\|\gamma_k - \gamma_k'\|_\phi < \delta$, $1 \le k \le n$, we have*

$$|H_\phi(\gamma_1, \dots, \gamma_n) - H(\gamma_1', \dots, \gamma_n')| < n\epsilon.$$

Partly, what makes the above result nice are the universal quantifiers in its conclusion: we get a $\delta > 0$ that works *for all* $n \in \mathbb{N}$ and states ϕ on A. To prove Proposition 3.3.8 we have to take a detour and examine another way in which decompositions of states arise. Namely, suppose that λ is a state on $A \otimes L^\infty(X, \mu)$ with the property that $\lambda_{|A} = \phi$ and $\lambda_{|L^\infty(X,\mu)} = \mu$. We will then say that λ is a **coupling** of (A, ϕ) with (X, μ). Given n partitions of (X, μ), say $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_n$, we then get a decomposition of ϕ

$$\phi = \sum_{Z_1 \in \mathcal{P}_1, \dots Z_n \in \mathcal{P}_n} \lambda(\cdot \otimes \mathbb{1}_{Z_1 \cap \dots \cap Z_n}) \tag{3.15}$$

{eq55}

We shall call this the decomposition induced by the $(\lambda, \mathcal{P})$. Let us compute the entropy of some channels $\gamma_1, \dots \gamma_n$ with respect to this decomposition. We obtain

$$H(\gamma_1, \dots, \gamma_n; \{\phi_{i_1, \dots, i_n}\}_{i_1, \dots, i_n}) \tag{3.16}$$
$$= H_\mu(\vee_k \mathcal{P}_n) + \sum_k \sum_{Z \in \mathcal{P}_k} S(\lambda(\gamma_k(\cdot) \otimes \mathbb{1}_Z), \phi \circ \gamma_k). \tag{3.17}$$

Naturally, we denote this quantity by $H_\lambda(\gamma_1, \dots, \gamma_n; \mathcal{P}_1, \dots, \mathcal{P}_n)$. Using Equation (3.9) we furthermore get

$$H_\lambda(\gamma_1, \dots, \gamma_n; \mathcal{P}_1, \dots, \mathcal{P}_n) = H_\mu(\vee_k \mathcal{P}_k) - \sum_k H_\mu(\mathcal{P}_k) \tag{3.18}$$
$$+ \sum_k \sum_{Z \in \mathcal{P}_k} \mu(Z) S(\mu(Z)^{-1} \lambda(\gamma_k(\cdot) \otimes \mathbb{1}_Z), \phi \circ \gamma_k) \tag{3.19}$$

{eq59}

Putting $\phi_Z = \mu(Z)^{-1} \lambda(\cdot \otimes \mathbb{1}_Z)$ we can rewrite this as

$$H_\mu(\vee_k \mathcal{P}_k) - \sum_k H_\mu(\mathcal{P}_k) + \sum_k \left(S(\phi \circ \gamma_k) - \sum_{Z \in \mathcal{P}_k} \mu(Z) S(\phi_Z \circ \gamma_k) \right) \tag{3.20}$$

{eq17}

Equation (3.11) now becomes

$$H_\lambda(\gamma_1, \dots, \gamma_n; \mathcal{P}_1, \dots, \mathcal{P}_n) = H_\mu(\vee_k \mathcal{P}_k) - \sum_k H_\mu(\mathcal{P}_k) + \sum_k H_\lambda(\gamma_k; \mathcal{P}_k) \quad (3.21)$$

{eq57}

At this point we should remark that any decomposition of ϕ induced by a decomposition of the coupling. Indeed, if $\phi = \sum_{i_1,\dots,i_n} \phi_{i_1,\dots,i_n}$ simply let $X = I_1 \times \dots \times I_n$ and $\mathcal{P}_k = \{I_1 \times \dots I_{k-1} \times \{i_k\} \times I_{k+1} \dots I_n : i_k \in I_k\}$.

There is a special class of couplings that warrant extra attention. These are those where $X = S(A)$, the state space of A, μ is a regular probability measure on X, and $\lambda : A \otimes L^\infty(X, \mu) \to \mathbb{C}$ satisfies

$$\lambda(a \otimes g) = \int_X \psi(a) g(\psi)\ d\mu(\psi)\ , a \in A, g \in L^\infty(X, \mu) \quad (3.22)$$

{eq124}

Note here that the integral is well defined because the for fixed $a \in A$, the mapping $\psi \mapsto \psi(a)$ is continuous and bounded, and $g(\cdot)$ is measurable and essentially bounded on X by definition. If a coupling λ satisfies Equation (3.22) we say that it is *canonical*.

We have already seen that couplings are enough to estimate the quantity $H_\phi(\gamma_1, \dots, \gamma_n)$. It turns out that even canonical couplings are sufficient for this and that this estimation is "nice" in a precise sense:

Given a coupling λ of (A, ϕ) with (Y, μ) and a partition $\mathcal{P}$ of Y, put $X = S(A)$ and consider a map $f : Y \to X$ that maps the set $Z \in \mathcal{P}$ to the state $\phi_Z = \mu(Z)^{-1} \lambda(\cdot, \mathbb{1}_Z)$. Define the pushback measure μ' on X by $\mu'(\cdot) = \mu(f^{-1}(\cdot))$. Note that this is supported on a finite set. Define $\lambda' = \lambda \circ (\mathrm{id}_A \otimes f^*)$ where $f^* : L^\infty(X, \mu') \to L^\infty(Y, \mu)$ is the map induced by f. We shall check that λ' is a canonical coupling of (A, ϕ). That $\lambda'_{|A} = \phi$ and $\lambda'_{|L^\infty(X,\mu')} = \mu'$ is obvious. μ' is a finitely supported measure so it suffices to verify Equation (3.22) on its atoms, i.e. for $g = \mathbb{1}_{\{\phi_Z\}}$ for $Z \in \mathcal{P}$. Then the LHS of Equation (3.22) is

$$\lambda'(a \otimes \phi_Z) = \lambda(a \otimes (\mathbb{1}_{\{\phi_Z\}} \circ f)) = \lambda(a \otimes \mathbb{1}_Z)$$

On the other hand, the RHS is

$$\int_X \psi(a) \mathbb{1}_{\{\phi_Z\}}\ d\mu'(\psi) = \phi_Z(a) \mu'(\phi_Z) = \mu(Z)^{-1} \lambda(a \otimes \mathbb{1}_Z) \mu(Z) = \lambda(a \otimes \mathbb{1}_Z)$$

So indeed λ' is a canonical. Moreover, it is easy to see that for a Borel partition $\mathcal{C}$ of $X = S(A)$, the decompositions of ϕ induced by $(\lambda, f^{-1}(\mathcal{C}))$ and $(\lambda', \mathcal{C})$ are the same. In particular, when $f^{-1}(\mathcal{C}) = \mathcal{P}$, i.e. when each member of $\mathcal{C}$ contains at most one element of the image of f, then

$$H_\lambda(\gamma; \mathcal{P}) = H_{\lambda'}(\gamma; \mathcal{C})$$

.

However, we need not have $f^{-1}(\mathcal{C}) = \mathcal{P}$ in order for $H_{\lambda'}(\gamma; \mathcal{C})$ to be close to $H_\lambda(\gamma; \mathcal{P})$. It turns out that they are close as long as $\mathcal{C}$ is sufficiently fine, even if members of $\mathcal{C}$ contain multiple elements of $f(Y)$:

lem: uniformsinglechannel

Lemma 3.3.9 ([NS06][pp. 44, Lemma 3.1.9]). *Let all notation and variables be as above. For $\delta > 0$ be such that for states ψ_1, ψ_2 on B, $\|\psi_1 - \psi_2\| < \delta \to |S(\psi_1) - S(\psi_2)| < \epsilon$. If $\mathcal{C}$ is a partition of $S(A)$ such that $\|\phi_1 \circ \gamma - \phi_2 \circ \gamma\| < \delta$ whenever ϕ_1 and ϕ_2 are in the same member of $\mathcal{C}$, then $|H_\lambda(\gamma;\mathcal{P}) - H_{\lambda'}(\gamma;\mathcal{C})| < \epsilon$.*

Proof. By construction of f it is clear that $f^{-1}(\mathcal{C}) \leq \mathcal{P}$. Now, for $W \in \mathcal{C}$ set $\psi'_W := \mu'(W)^{-1}\lambda'(\cdot \otimes \mathbb{1}_W)$. We have

$$\psi'_W = \mu'(W)^{-1} \sum_{Z\in\mathcal{P}:Z\subset f^{-1}(W)} \lambda(\cdot \otimes \mathbb{1}_W) = \sum_{Z\in\mathcal{P}:Z\subset f^{-1}(W)} \frac{\mu(Z)}{\mu(f^{-1}(W))}\phi_Z.$$

Now, fixing a $Z' \in \mathcal{P}$ with $Z' \subset f^{-1}(W)$ then for any other $Z \in \mathcal{P}$ with $Z \subset f^{-1}(W)$ the definition of f implies that $\phi_{Z'}, \phi_Z \in W$. Then, by assumption, $\|(\phi_{Z'} - \phi_Z) \circ \gamma\| < \delta$. By the above equation, ψ'_W is a convex combination of such ϕ_Z's so it follows that for any $Z' \in f^{-1}(W)$, $\|(\psi'_W - \phi_{Z'}) \circ \gamma\| < \delta$. Hence

$$|S(\psi'_W \circ \gamma) - S(\phi_{Z'} \circ \gamma)| < \epsilon.$$

Equation (3.20) is simple in the case of one channel and reveals that

$$S(\phi \circ \gamma) - H_\lambda(\gamma;\mathcal{C}) = \sum_{W\in\mathcal{C}} \mu'(W) S(\psi'_W \circ \gamma) = \sum_{W\in\mathcal{C}} \mu(f^{-1}(W)) S(\psi'_W \circ \gamma).$$

Similarly

$$S(\phi \circ \gamma) - H_\lambda(\gamma;\mathcal{P}) = \sum_Z \mu(Z) S(\phi_Z \circ \gamma) = \sum_{W\in\mathcal{C}} \sum_{Z\in\mathcal{P}:Z\subset f^{-1}(W)} \mu(Z) S(\phi_Z \circ \gamma).$$

Convexity then yields $|H_\lambda(\gamma;\mathcal{P}) - H_{\lambda'}(\gamma;\mathcal{C})| < \epsilon$. ∎

This lemma shows that to estimate $H_\lambda(\gamma;\mathcal{C})$ up to ϵ-accuracy we only need to consider decompositions up to a certain size. More precisely, since the unit ball of the state space of B is totally bounded it can be covered say by n_ϵ δ balls. So in Lemma 3.3.9 could have chosen a partition $\mathcal{C}$ of cardinality n_ϵ which of course corresponds to a decomposition of cardinality n_ϵ. So far, we have only dealt with the case of one channel, but we need only modify the techniques slightly to deal with the case of n channels:

Suppose λ again is a coupling of (A, ϕ) with (Y, μ) and $\mathcal{P}_1, \ldots, \mathcal{P}_n$ are partitions of (Y, μ). Set $X = S(A)$ and let $f_k : Y \to X$ map a set $Z \in \mathcal{P}_k$ to the state $\mu(Z)^{-1}\lambda(\cdot \otimes \mathbb{1}_Z) \in X$. Then the map $f = (f_1, f_2, \ldots, f_n) : Y \to X^n$ is measurable and we can define a measure, $\mu'(\cdot)$, on X^n by $\mu'(\cdot) = \mu(f^{-1}(\cdot))$. Then define a coupling λ' of (A, ϕ) with (X^n, μ') by $\lambda' = \lambda(\mathrm{id}_A \otimes f^*)$. Then we have an analogue of Lemma 3.3.9 for n channels $\gamma_k : A_k \to A$. We omit the proof because it is really just a more technical version of the proof of Lemma 3.3.9.

Proposition 3.3.10 ([NS06][pp. 45, Lemma 3.1.10]). *Let the notation and variables be as in the paragraph above and let $\delta > 0$ be such that for states ψ_1, ψ_2 on A_k, $\|\psi_1 - \psi_2\| < \delta \to |S(\psi_1) - S(\psi_2)| < \epsilon$. If $\mathcal{C}_1, \ldots, \mathcal{C}_n$ are partitions of $X = S(A)$ such that $\|\phi_1 \circ \gamma_k - \phi_2 \circ \gamma_k\| < \delta$ whenever ψ_1 and ψ_2 lie in the same member of $\mathcal{C}_k$, for $1 \leq k \leq n$, then*

$$H_\lambda(\gamma_1, \ldots, \gamma_n; \mathcal{C}_1, \ldots \mathcal{C}_n) \leq H_\lambda(\gamma_1, \ldots, \gamma_n; \mathrm{pr}_1^{-1}(\mathcal{C}_1), \ldots, \mathrm{pr}_1^{-1}(\mathcal{C}_n)) + n\epsilon.$$

This enables us to prove Proposition 3.3.8.

We are now ready to define the entropy of an non-commutative dynamical systems.

3.4 Entropy of dynamical systems

Definition 3.4.1. Let G be a group and A a C*-algebra. We call a map $\alpha : G \to \mathrm{Aut}(A)$ a **group action** if $\alpha_{g_1} \circ \alpha_{g_2} = \alpha_{g_1 g_2}$ for all $g_1, g_2 \in G$. Given a state ϕ on A satisfying $\phi \circ \alpha_g = \phi$, $g \in G$, we call the triple (A, ϕ, α) a **C*-dynamical system**. If in addition A is a von Neumann algebra and ϕ is normal, we say that (A, ϕ, α) is a **W*-dynamical system**.

From here on out we will assume that G is an amenable group and $\{F_n\}_{n=1}^{\infty}$ is a Følner sequence of G. We will now combine our work on channels and Theorem 2.1.6 to define the entropy of C*-dynamical systems.

def: dynamicalentropy

Definition 3.4.2 ([NS06][pp. 48, Definition 3.2.1])**.** Given a C*-dynamical system (A, ϕ, α) and a channel $\gamma : B \to A$ we define

$$h_\phi(\gamma; \alpha) = \lim_{n \to \infty} \frac{1}{|F_n|} H_\phi(\{\alpha_g \circ \gamma : g \in F_n\}).$$

The **entropy of** the system (A, ϕ, α) is now defined as the supremum of $h_\phi(\gamma; \alpha)$ for all possible channels γ and we denote it $h_\phi(\alpha)$.

Note that the limit in Definition 3.4.2 exists because the mapping $F \mapsto H_\phi(\alpha_g \circ \gamma : g \in F)$ defined on finite subsets of G is, by Proposition 3.3.4 (ii) and (iii) invariant under left multiplication by elements of G, and subadditive. Theorem 2.1.6 then implies that the above limit exists and is the same regardless of the choice of Følner sequence $\{F_n\}_{n=1}^{\infty}$.

Given a C*-dynamical system (A, ϕ, α) we get an induced W*-dynamical system using the GNS representation $\pi_\phi : A \to B(H_\phi)$ associated to ϕ. Namely, let h_ϕ be the cyclic unit vector in the GNS- representation and for $g \in G$ define

$$U'_g : \pi_\phi(A) h_\phi \to \pi_\phi(A) h_\phi, \; U'_g(\pi_\phi(a) h_\phi) = \pi_\phi(\alpha_g(a)) h_\phi.$$

The computation

$$\begin{aligned} \|\pi_\phi(a) h_\phi\|^2 = \langle \pi_\phi(a) h_\phi, \pi_\phi(a) h_\phi \rangle = \langle \pi_\phi(a^* a) h_\phi, h_\phi = \phi(a^* a) = \phi(\alpha_g(a^* a)) \\ = \|\pi_\phi(\alpha_g(a)) h_\phi\|^2 \rangle \end{aligned}$$

reveals that the U'_g are well-defined surjecive isometries. Extend each U'_g to a unitary U_g on H_ϕ. Put $M = \pi_\phi(A)''$ and let $\overline{\phi}$ be the vector state extension $\langle \cdot \, h_\phi, h_\phi \rangle$ of ϕ to M. $g \mapsto \mathrm{Ad}\, U_g$ gives a group action on M which we will denote by $\overline{\alpha}$. We assert that

$$\pi_\phi \circ \alpha_g = \mathrm{Ad}\, U_g \circ \pi_\phi, \; g \in G \tag{3.23}$$

{eq18}

This follows from the simple computation

$$\langle \pi_\phi(\alpha_g(a)) \pi_\phi(b) h_\phi, \pi_\phi(c) h_\phi \rangle = \phi(c^* \alpha_g(a) b)) = \phi(\alpha_{g^{-1}}(c) a \alpha_{g^{-1}}(b))$$

$$= \langle \pi_\phi(a)\pi_\phi(\alpha_{g^{-1}}(b))h_\phi, \pi_\phi(\alpha_{g^{-1}}(c))h_\phi\rangle = \langle \pi_\phi(a)U_g^*\pi_\phi(b)h_\phi, U_g^*\pi_\phi(c)h_\phi\rangle.$$

Similar computations yield that $\mathrm{Ad}\, U_g$ is $\overline{\phi}$ invariant so $(M, \overline{\phi}, \overline{\alpha})$ is a W*-dynamical system.

If the system (A, ϕ, α) we started with was a W*-dynamical system then the image $\pi_\phi(A)$ is already a von Neumann algebra ([BR12][pp. 79, Theorem 2.4.24]).If ϕ is also faithful π_ϕ is injective and by Equation (3.23) the system is then G-*equivariantly isomorphic* to the W*-dynamical system where the state is a vector state. So if the state is faithful, then in the W*-dynamical case, we can assume the state is a vector state and the group action is given by unitary conjugation. This point of view remains valid in the non-faithful case if we mod out A by $\ker \pi_\phi$.

We now give some basic properties of dynamical entropy. Given an α-invariant subalgebra $B \subset A$ we denote by $\alpha_{|B}$ the restricted group action $G \to \mathrm{Aut}(B), g \mapsto \alpha_{g|B}$.

prop: dymanicalentropyproperties

Proposition 3.4.3 ([NS06][pp. 48, Theorem 3.2.2]). *Let (A, ϕ, α) be a C*-dynamical system. Then*

(i) *If $\beta : A \to B$ is an isomorphism of C*-algebras, $h_{\phi\circ\beta^{-1}}(\beta\circ\alpha\circ\beta^{-1}) = h_\phi(\alpha)$.*

(ii) *With the notation as above, we have $h_{\overline{\phi}}(\overline{\alpha}) = h_\phi(\alpha)$.*

(iii) *for another C*-dynamical system (B, ψ, β) and $0 \le \lambda \le 1$,*

$$h_{\lambda\phi\oplus(1-\lambda)\psi}(\alpha \oplus \beta) = \lambda h_\phi(\alpha) + (1-\lambda)h_\phi(\beta) \text{ and}$$

$$h_{\phi\otimes\psi}(\alpha \otimes \beta) \le h_\phi(\alpha) + h_\psi(\beta).$$

(iv) *If the underlying group is abelian, define the action α^{-1} by $\alpha_g^{-1} = \alpha_{g^{-1}}$, then $h_\phi(\alpha^{-1}) = h_\phi(\alpha)$.*

(v) *if B is an α-invariant subalgebra of A and there exists a ϕ-preserving conditional expectation $A \to B$, then $h_{\psi_{|B}}(\alpha_{|B}) \le h_\phi(\alpha)$.*

(vi) *If $H \subset G$ is a subgroup of G of finite index $[G : H]$, then the group action $\alpha_{|H}$ given by restriction to H, satisfies $h_\phi(|\alpha_H) \le [G : H]h_\phi(\alpha_{|H})$.*

Proof.

(i) This is easy.

(ii) Proposition 3.3.7(i) implies that for a channel $\gamma : B \to A$ we have

$$h_{\overline{\phi}}(\pi_\phi \circ \gamma; \overline{\alpha}) = h_\phi(\gamma; \alpha). \tag{3.24}$$ {eq1001}

so we obtain $h_{\overline{\phi}}(\overline{\alpha}) \ge h_\phi(\alpha)$ immediately. To see why $\le$ holds, take a channel $\gamma : B \to \pi_\phi(A)''$ Corollary 3.1.8 says that it can be approximated by a channel $\gamma' : B \to \pi_\phi(A)$ in the pointwise strong operator topology. But the equality

$$||(\gamma-\gamma')(b)h_\phi||^2 = \langle(\gamma-\gamma')(b)^*(\gamma-\gamma')(b)h_\phi, h_\phi\rangle = \overline{\phi}((\gamma-\gamma'(b))^*(\gamma-\gamma'(b)))$$

then tells us that γ' is close to γ in the $\|\cdot\|_{\overline{\phi}}$-seminorm which by Proposition 3.3.8 means that $h_{\overline{\phi}}(\gamma';\overline{\alpha}) \approx h_{\overline{\phi}}(\gamma;\overline{\alpha})$. By Corollary 3.1.8 γ' can be written as $\pi_\phi \circ \gamma''$ for some channel γ'' into A. Applying Equation (3.24) again finishes the proof.

(iii) This follows from Proposition 3.3.7.

(iv) Let $\{F_n\}_n$ be a bi-Følner sequence of G which we recall is a sequence that is simultaneously left- and right-Følner. Then $\{F_n^{-1}\}_n$ is also bi-Følner: the equality

$$sF_n^{-1} \triangle F_n^{-1} = (F_n s^{-1})^{-1} \triangle F_n^{-1} = (F_n s^{-1} \triangle F_n)^{-1}$$

shows that it is left-Følner and similarly it is right-Følner. Finally, the sequence $\{F_n \cup F_n^{-1}\}_n$ is Følner since

$$\begin{aligned} s(F_n \cup F_n^{-1}) \triangle (F_n \cup F_n^{-1}) &= (sF_n \cup sF_n^{-1}) \triangle (F_n \cup F_n^{-1}) \\ &= (sF_n \triangle F_n) \cup (sF_n^{-1} \triangle F_n^{-1}). \end{aligned}$$

Hence, given a channel $\gamma : B \to A$, we can compute $h_\phi(\alpha;\gamma)$ along the Følner sequence $\{F_n \cup F_n^{-1}\}$. But $h_\phi(\alpha_g \circ \gamma : g \in F_n \cup F_n^{-1}) = h_\phi(\alpha_{g^{-1}} \circ \gamma : g \in F_n \cup F_n^{-1})$ which establishes the equality.

(v) This follows from Lemma 3.3.5.

(vi) Use Proposition 2.1.13 to get a bi-Følner sequence $\{H_n\}_n$ of H such that $\{H_n F\}_n$ is a bi-Følner sequence for G, where F denotes a set of representatives of right cosets of H, that contains the identity e. Then $H_n \subset H_n F$ so for each channel γ into A,

$$\begin{aligned} h_\phi(\alpha_{|H};\gamma) = \frac{1}{|H_n|} H_\phi(\alpha_g \circ \gamma : g \in H_n) &\leq \frac{1}{|H_n|} H_\phi(\alpha_g \circ \gamma : g \in H_n F) \\ &= [G:H] \frac{1}{|F_n|} H_\phi(\alpha_g \circ \gamma : g \in H_n F) \end{aligned}$$

so we obtain the desired inequality.

■

Proposition 3.3.8 enables us to show that our non-commutative definition of entropy generalizes the classical definition. More precisely, we will show that the classical entropy of an amenable group action $\alpha : G \to (X,\mu)$ coincides with the operator-algebraic entropy of the induced group action $\beta : G \to L^\infty(X,\mu)$ given by $\beta_g(f) = f \circ \alpha_g^{-1}$ with respect to the state μ.

We begin by showing $h_\mu(\alpha) \leq h_\mu(\beta)$, i.e. that the operator algebraic entropy dominates the classical one. To that end, fix a partition $\mathcal{P}$ of (X,μ) and a positive integer n. By Proposition 2.3.3 we have

$$H(\bigvee_{s \in F_n} s^{-1}\mathcal{P}) = \sup\{\sum_{i_1,\ldots,i_n} \eta(\mu(f_{i_1,\ldots,i_n})) - \sum_{k=1}^{n} \sum_{i_k} \mu(\eta(E_{s^{-1}\mathcal{P}}(f_{i_k}^{(k)})))\}$$

$$= \sup\{\sum_{i_1,\ldots,i_n} \eta(\mu(f_{i_1,\ldots,i_n})) - \sum_{k=1}^{n}\sum_{i_k}\mu(\eta(\sum_{Z\in\mathcal{P}} \frac{\mu(f_{i_k}^{(k)}\mathbb{1}_{s^{-1}Z})}{\mu(Z)}\mathbb{1}_{s^{-1}Z}))\}$$

$$= \sup\{\sum_{i_1,\ldots,i_n} \eta(\mu(f_{i_1,\ldots,i_n})) + \sum_{k=1}^{n}\sum_{i_k}\sum_{Z\in\mathcal{P}} \mu(f_{i_k}^{(k)}\mathbb{1}_{s^{-1}Z})(\log\mu(f_{i_k}^{(k)}\mathbb{1}_{s^{-1}Z}) - \log\mu(Z))\} \tag{3.25}$$

{eq13}

Here the supremum is taken over all ordered partitions of unity $\{f_{i_1,\ldots,i_n}\}_{i_1,\ldots,i_n}$. But note that each such ordered partition of unity gives rise to a state decomposition $\mu = \sum_{i_1,\ldots,i_n} \mu(\cdot\, f_{i_1,\ldots,i_n})$ on $L^\infty(X,\mu)$. Letting $\gamma : L^\infty(X/\mathcal{P}) \to L^\infty(X,\mu)$ be the natural inclusion we recognize Equation (3.25) as $H_\gamma(\gamma; F_n)$. Dividing both sides of Equation (3.25) by $|F_n|$ and taking limits yields $\lim_{n\to\infty} \frac{1}{|F_n|} H(\bigvee_{s\in F_n} s^{-1}\mathcal{P}) \le h_\mu(\beta)$. Taking supremum over all partitions $\mathcal{P}$ we get $h_\mu(\alpha) \le h_\mu(\beta)$.

The reverse inequality relies on some machinery that, thankfully, we have already proven. Namely, let $\gamma : B \to L^\infty(X,\mu)$ be a channel and fix $\epsilon > 0$. Since (X,μ), is a standard probability space there exists an increasing sequence of finite partitions $\{\mathcal{P}_n\}_n$ such that their union generate the measurable sets of X modulo null sets (for example, if (X,μ) were the Cantor space $\{0,1\}^{\mathbb{N}}$, then $\mathcal{P}_n = \{\{x \in X : x_k = y_k \text{ for } k = 1,\ldots,n\} : y \in \{0,1\}^n\}$ would work). Then $\bigcup_n L^\infty(X/\mathcal{P}_n)$ is strongly operator dense in $L^\infty(X,\mu)$. By Proposition 3.1.7 we can then approximate γ by channels $\gamma' : B \to L^\infty(X/\mathcal{P}_n)$ in the pointwise s.o.t. topology. In particular we can, for any $b \in B$, get $\int_X |\gamma(b) - \gamma'(b)|^2 \, d\mu$ as small as desired. Use Proposition 3.3.8 to get a $\delta > 0$ such that for all $n \in \mathbb{N}$ and channels $\gamma_k, \gamma_k' : B \to A$ satisfying $\|\gamma_k - \gamma_k'\|_\mu < \delta$ for $1 \le k \le n$, we have

$$|H_\mu(\gamma_1,\ldots,\gamma_n) - H(\gamma_1',\ldots,\gamma_n')| < n\epsilon.$$

Since B is finite dimensional, we can now find a $\gamma' : B \to L^\infty(X/\mathcal{P}_n)$ such that $\|\gamma - \gamma'\|_\mu < \delta$. β is μ-preserving so clearly $\|\beta_g \circ \gamma - \beta_g \circ \gamma'\|_\mu < \delta$ for all $g \in G$. For $n \in \mathbb{N}$ we then get

$$H_\mu(\beta_g \circ \gamma; F_n) \le H(\beta_g \circ \gamma'; F_n) + |F_n|\epsilon$$

In particular we have,

$$\lim_{n\to\infty} \frac{1}{|F_n||} H_\mu(\gamma; F_n) \le \lim_{n\to\infty} \frac{1}{|F_n|} H(\gamma'; F_n) + \epsilon \tag{3.26}$$

{eq14}

Note that $\beta_g \circ \gamma'$ is a channel $B \to L^\infty(X/(g\mathcal{P}_n))$ so trivially it factorizes through the inclusion map $i_g : L^\infty(X/(g\mathcal{P}_n)) \to L^\infty(X,\mu)$, i.e. $\beta_g \circ \gamma = i_g \circ \beta_g \circ \gamma$. By Proposition 3.3.4 (i),

$$H_\mu(\beta_g \circ \gamma'; F_n) \le H_\mu(i_g; g \in F_n) \tag{3.27}$$

Computing the quantity $H_\mu(i_g; g \in F_n)$ is easy because any state decomposition $\mu = \sum_{i_1,\ldots,i_n} \phi_{i_1,\ldots,i_n}$ is of the form $\{\mu(\cdot\, f_{i_1,\ldots,i_n})\}$ for some partition of unity $\{f_{i_1,\ldots,i_n}\}_{i_1,\ldots i_n}$. Hence

$$H_\mu(\beta_g \circ \gamma'; g \in F_n) \le H_\mu(i_g; g \in F_n)$$

$$= \sup\{\sum_{i_1,\dots,i_n} \eta(\mu(f_{i_1,\dots,i_n})) + \sum_{k=1}^n \sum_{i_k} \sum_{Z \in \mathcal{P}} \mu(f_{i_k}^{(k)} \mathbb{1}_{s^{-1}Z})(\log \mu(f_{i_k}^{(k)} \mathbb{1}_{s^{-1}Z}) - \log \mu(Z))\}$$

By Equation (3.25) this is the same as $H(\bigvee_{s \in F_n} s^{-1}\mathcal{P})$. Hence we obtain

$$\lim_{n\to\infty} \frac{1}{|F_n|} H_\mu(\beta_g \circ \gamma'; g \in F_n) \le \lim_{n\to\infty} \frac{1}{|F_n|} H(\bigvee_{s \in F_n} s^{-1}\mathcal{P}).$$

Combining this with Equation (3.26), this completes the proof that $h_\mu(\beta) = h_\mu(\alpha)$. Using Proposition 3.4.3(ii) "backwards" we see that the action $(\beta_0)_g(f) = f \circ \alpha_g^{-1}$ on $C(X)$ satisfies $h_{\beta_0}(\mu) = h_\alpha(\mu)$ also.

More precisely, treat μ as a state on $C(X)$ and consider the Hilbert space H_μ and representation π_μ in the μ's GNS-construction. H_μ can be identified with $L^2(X,\mu)$ and π_ϕ as multiplication, i.e. $\pi_\mu(f)(g) = fg$, $g \in L^2(X,\mu)$. Then $\pi_\mu(C(X))'' \simeq L^\infty(X,\mu)$. Under this identification the action $\overline{\beta_0}$ on $\pi_\mu(C(X))''$ as in Proposition 3.4.3(ii) corresponds to β. Then Proposition 3.4.3(ii) tells us that $h_{\beta_0}(\mu) = h_\beta(\mu) = h_\alpha(\mu)$.

Let us return to general entropy. Intuitively, the entropy should be small when the action doesn't "move" A much around. The next result tells us that if the action is compact in the $\|\cdot\|_\phi$ topology, its entropy is 0.

Proposition 3.4.4. *Let (A,ϕ,α) be a C*-dynamical system. Suppose that for each $a \in A$ the orbit $Ga := \{\alpha_g(a) : g \in G\}$ is precompact when A is given the topology induced by the seminorm $\|\cdot\|_\phi$. Then $h_\phi(\alpha) = 0$.*

Proof. The result is essentially a corollary of Proposition 3.3.8. Fix a channel $\gamma : B \to A$ and $\epsilon > 0$ and find $\delta > 0$ such that for any $n \in \mathbb{N}$ and channels $\gamma_1,\dots,\gamma_n,\gamma_1',\dots,\gamma_n'$ on C*-algebras of dimension no greater than $\dim B$ we have that

$$\|\gamma_k - \gamma_k'\|_\phi < \delta,\ 1 \le k \le n, \to |H_\phi(\gamma_1,\dots,\gamma_n) - H_\phi(\gamma_1',\dots,\gamma_n')| < n\epsilon.$$

B is finite dimensional and precompactness just means total boundedness so surely we can find a finite set $F \subset G$ such that for all $g \in G$ there is an $s \in F$ satisfying $\|\alpha_g \circ \gamma - \alpha_s \circ \gamma\|_\phi < \delta$. We then obtain

$$\frac{1}{|F_n|} H_\phi(\alpha_g \circ \gamma; g \in F_n) \le \frac{1}{|F_n|} H_\phi(\alpha_s \circ \gamma : s \in F) + \epsilon \le \frac{1}{|F_n|} |F| H_\phi(\gamma).$$

Equation (3.9) tells us that

$$H_\phi(\gamma) = \sup\{\sum_i \phi_i(1) S(\hat{\phi}_i \circ \gamma, \phi \circ \gamma)\} = \sup\{S(\phi \circ \gamma) - \sum_i \phi_i(1) S(\hat{\phi}_i \circ \gamma)\}$$
$$\le S(\phi \circ \gamma) \le \log \dim B$$

Hence,

$$\frac{1}{|F_n|} H_\phi(\alpha_g \circ \gamma; g \in F_n) \le \frac{|F|}{|F_n|} \log \dim B \to 0$$

Hence $h_\phi(\alpha) = 0$. ∎

Note the similarity between the above proposition and Theorem 2.4.9. Indeed, if the C*-dynamical system in the above proposition is just $(L^\infty(X,\mu),\mu,\alpha)$ where α is induced by some action $\beta : G \curvearrowright (X,\mu)$, then the $\|\cdot\|_\mu$ norm is just the L^2-norm and the compactness assumption in the above proposition boils down to the compactness assumption in Theorem 2.4.9. Since the latter is about classical sofic entropy in its full generality, and the former is about amenable dynamical entropy neither result generalizes the other.

At this point we need to add an assumption on our C*-dynamical systems to get interesting results. Our assumption will, as in the proposition above, make use of the $\|\cdot\|_\phi$ seminorm. More precisely, for points $a \in A$ we define

$$\|a\|_\phi := \phi(a^* a)^{1/2}.$$

For channels $\gamma, \gamma' : B \to A$ we define

$$\|\gamma - \gamma'\|_\phi := \sup_{b\in B, \|b\|\le 1} \phi(((\gamma-\gamma')(b))^*((\gamma-\gamma')(b)))^{1/2}$$

as before.

The assumption we will have on our C*-dynamical systems is that there is a net of channels $\{\gamma_i : A_i \to A\}_{i\in I}$ and a net of unital completely positive maps $\{\theta_i : A \to A_i\}_{i\in I}$ such that for each $x \in A$ we have $\|(\gamma_i \circ \theta_i)(x) - x\|_\phi \to 0$. The net $\{\gamma_i : A_i \to A\}_{i\in I}$ will be called **ϕ-approximating** ([NS06][pp. 49]). By finite dimensionality it follows at once that for any channel $\gamma : B \to A$ we have $\|\gamma_i \circ \theta_i \circ \gamma - \gamma\|_\phi \to 0$. Proposition 3.3.8 then implies that

$$\lim_i h_\phi(\gamma_i \circ \theta_i \circ \gamma; \alpha) = h_\phi(\gamma;\alpha)$$

By Proposition 3.3.4(i) we have $h_\phi(\gamma_i \circ \theta_i \circ \gamma;\alpha) \le h_\phi(\gamma_i;\alpha) \le h_\phi(\alpha)$. Hence we get the following.

thm: approximatingnet

Theorem 3.4.5 ([NS06][pp. 49, Theorem 3.2.3]). *If $\{\gamma_i : A_i \to A\}_{i\in I}$ is a ϕ-approximating net for (A,ϕ,α), then $h_\phi(\alpha) = \lim_i h_\phi(\gamma_i;\alpha)$.*

Having a ϕ-approximating net might seem like an awkward condition and clearly it is designed ad-hoc; it is a condition that is just strong enough so that many of our proofs will work. Note that if A is a nuclear C*-algebra, any system (A,ϕ,α) has a ϕ-approximating net. Indeed, by definition of nuclearity we have nets $\{\gamma_i : A_i \to A\}_{i\in I}$ and $\{\theta_i : A \to A_i\}_{i\in I}$ such that for any $a \in A$ $\|(\gamma_i \circ \theta_i)(a) - a\| \to 0$. Hence there is a net which is ϕ-approximating for any ϕ. Another result related to the $\|\cdot\|_\phi$-seminorm is the following.

prop: liminfsubalgs

Proposition 3.4.6 ([NS06][pp. 49, Proposition 3.2.4]). *Let (A, ϕ, α) be a C^*-dynamical system, $\{A_i\}_{i \in I}$ an increasing net of α-invariant C^*-subalgebras of A such that $\bigcup_{i \in I} \pi_\phi(A_i)$ is strongly dense in $\pi_\phi(A)$. Then*

$$h_\phi(\alpha) \leq \liminf_i h_{\phi_{|A_i}}(\alpha_{|A_i}).$$

Proof. By Proposition 3.1.7 we can approximate any channel $\gamma : B \to \pi_\phi(A)''$ by a channel $\gamma' : B \to \pi_\phi(A_i)$ in the pointwise strong operator topology, for sufficiently large i. In particular we can approximate $\gamma : B \to \pi_\phi(A)''$ by a channel $\gamma' : B \to \pi_\phi(A_i)$ in the $\|\cdot\|_{\overline{\phi}}$-seminorm, for sufficiently large i. Proposition 3.3.8 then implies that to compute $h_{\overline{\phi}}(\overline{\alpha})$ it suffices to take suprema over $h_{\overline{\phi}}(\gamma; \overline{\alpha})$ for channels $\gamma : B \to \pi_\phi(A_j)$ for $j \in I$. For such a γ we have for $i \geq j$:

$$h_{\overline{\phi}}(\gamma; \overline{\alpha}) \leq h_{\overline{\phi}_{|\pi_\phi(A_i)}}(\gamma; \overline{\alpha}_{|\pi_\phi(A_i)}).$$

Taking lim inf on both sides yields,

$$h_{\overline{\phi}}(\gamma; \overline{\alpha}) \leq \liminf_i h_{\overline{\phi}_{|\pi_\phi(A_i)}}(\gamma; \overline{\alpha}_{|\pi_\phi(A_i)}).$$

Taking suprema over γ's that map into a $\pi_\phi(A_j)$ algebra yields

$$h_{\overline{\phi}}(\overline{\alpha}) \leq \liminf_i h_{\overline{\phi}_{|\pi_\phi(A_i)}}(\overline{\alpha}_{|\pi_\phi(A_i)}).$$

The LHS equals $h_\phi(\alpha)$ and by Corollary 3.1.8, $h_{\overline{\phi}_{|\pi_\phi(A_i)}}(\overline{\alpha}_{|\pi_\phi(A_i)}) = h_{\phi_{|A_i}}(\alpha_{|A_i})$. Hence

$$h_\phi(\alpha) \leq \liminf_i h_{\phi_{|A_i}}(\alpha_{|A_i}).$$

■

We will now see why the ϕ-approximating net assumption is useful.

thm: dynamicalsystemproperties

Theorem 3.4.7 ([NS06][pp. 50, Theorem 3.2.5]). *Let (A, ϕ, α) be a C^*-dynamical system having a ϕ-approximating net. Then*

(i) *If $H \subset G$ is a subgroup of G with finite index, then*

$$h_\phi(\alpha_{|H}) = [G : H]\ h_\phi(\alpha).$$

(ii) *if (B, ψ, β) is another C^*-dynamical system with B abelian, then*

$$h_{\phi \otimes \psi}(\alpha \otimes \beta) = h_\phi(\alpha) + h_\psi(\beta).$$

(iii) *if ψ is a state on a C^*-algebra B, and there exists a ψ-approximating net, then $h_{\alpha \otimes \mathrm{id}_B}(\alpha \otimes \beta) = h_\phi(\alpha)$.*

(iv) *if G contains subgroups of arbitrarily large finite index then $h_\phi(\alpha)$ is concave in ϕ; if $\phi = \lambda\psi + (1 - \lambda)\omega$, where ψ and ω are α-invariant states, then*

$$h_\phi(\alpha) \geq \lambda h_\psi(\alpha) + (1 - \lambda) h_\omega(\alpha).$$

Proof.

(i) By Proposition 3.4.3 (vi) we have $h_\phi(\alpha_{|H}) \le [G:H]h_\phi(\alpha)$. To prove the opposite inequality, let $\{\gamma_i : A_i \to A\}_{i\in I}$ be a ϕ-approximating net and $\theta_i : A \to A_i$ be as in the definition of such a net. Fix a channel $\gamma : B \to A$ and $\epsilon > 0$. Set $k = [G:H]$, let F be a set of representatives of right cosets of H and use Proposition 2.1.13 to obtain a Følner sequence $\{H_n\}_n$ of H such that $\{H_nF\}_n$ is a Følner sequence of G. Let $\delta > 0$ and choose $i \in I$ such that

$$\|\gamma_i \circ \theta_i \circ \alpha_g \circ \gamma - \alpha_g \circ \gamma\|_\phi < \delta \text{ for } g \in F$$

We then also obtain $\|\alpha_h \circ \gamma_i \circ \theta_i \circ \alpha_g \circ \gamma - \alpha_{hg} \circ \gamma\|_\phi < \delta$ for any $h \in H$ and $g \in F$. If we choose δ small enough to match ϵ and the dimension of B as in Proposition 3.3.8, we obtain, for all $n \in N$:

$$H_\phi(\alpha_{hg} \circ \gamma; h \in H_n,\ g \in F) \le$$

$$H_\phi(\alpha_h \circ \gamma_i \circ \theta_i \circ \alpha_g \circ \gamma; h \in H_n,\ g \in F) + |H_n||F|\epsilon \le$$

$$H_\phi(\alpha_h \circ \gamma; h \in H_n) + |H||F|\epsilon.$$

Divide both sides by $\dfrac{1}{|H_nF|}$ and take limits to obtain
$h_\phi(\gamma;\alpha) \le [G:F]^{-1}h_\phi(\gamma;\alpha_{|H}) + \epsilon$, i.e. $h_\phi(\alpha_{|H}) \ge [G:F]h_\phi(\alpha)$.

(ii) We first claim that the dynamical systems $(\pi_{\phi\otimes\psi}(A\otimes B)', \overline{\phi\otimes\psi}, \overline{\alpha\otimes\beta})$ and $(\pi_\phi(A)' \otimes \pi_\psi(B)', \overline{\phi}\otimes\overline{\psi}, \overline{\alpha}\otimes\overline{\beta})$ are isomorphic. Note that

$$\langle(\pi_\phi \otimes \pi_\psi)(a\otimes b)(h_\phi \otimes h_\psi), (h_\phi \otimes h_\psi)\rangle = \langle \pi_\phi(a)h_\phi, h_\phi\rangle\langle \pi_\psi(a)h_\phi, h_\psi\rangle$$
$$= (\phi\otimes\psi)(a\otimes b)$$

so by the uniqueness of the GNS-representation there is a unitary $U : H_{\phi\otimes\psi} \to H_\phi \otimes H_\psi$ mapping $\pi_{\phi\otimes\psi}(a\otimes b)h_{\phi\otimes\psi}$ to $(\pi_\phi\otimes\pi_\phi)(a\otimes b)(h_\phi\otimes h_\psi)$. It is then checked that $\overline{\phi\otimes\psi} = (\overline{\phi}\otimes\overline{\psi}) \circ \mathrm{Ad}\,U$ and $\overline{\alpha\otimes\beta} \circ \mathrm{Ad}\,U = \mathrm{Ad}\,U \circ (\overline{\alpha}\otimes\overline{\beta})$ showing that the systems are equivariant and so $h_{\overline{\phi\otimes\psi}}(\overline{\alpha\otimes\beta}) = h_{\overline{\phi}\otimes\overline{\psi}}(\overline{\alpha}\otimes\overline{\beta})$. Hence, in proving (iv) we may assume that (B, ψ, β) (and $(A.\phi, \alpha)$) are W*-dynamical systems.

Then B is an abelian inductive limit of finite dimensional C*-algebras $\{B_k\}_k$. Let $\gamma'_k : B_k \to B$ denote their inclusions into B and choose arbitrary conditional expectations $E_k : B \to B_k$. Then $\gamma'_k \circ E_k \to \mathrm{id}_B$ pointwise so $\{\gamma'_k\}_k$ is a ϕ-approximating net. Letting $\{\gamma_i\}_i$ be a ϕ-approximating net we see that $\{\gamma_i \otimes \gamma'_k\}_{(i,k)}$ is a $\phi\otimes\psi$-approximating net. $h_{\phi\otimes\psi}(\alpha\otimes\beta) = h_\phi(\alpha) + h_\psi(\beta)$ now follows from Proposition 3.3.7(iv) and Theorem 3.4.5.

(iii) We omit the proof.

(iv) By Proposition 3.3.7(ii) we have

$$h_\phi(\alpha) \ge \lambda h_\psi(\alpha) + (1-\lambda)h_\omega(\alpha) - \eta(\lambda) - \eta(1-\lambda)$$

This isn't exactly what we want, but if we apply this inequality to $\alpha_{|H}$ for any finite index subgroup H of G and applying (i) yields

$$[G:H]h_\phi(\alpha) \ge \lambda[G:H]h_\psi(\alpha) + (1-\lambda)h_\omega(\alpha) - \eta(\lambda) - \eta(1-\lambda)$$

Dividing by $[G : H]$ and choosing H so that $[G : H] \to \infty$ yields the desired inequality.

■

The assumption on G we made in (iv) in the above proposition is be satisfied when G is an infinite residually finite group. For in this case have an increasing sequence of finite indexed normal subgroups $\{N_n\}_n$ such that any non-trivial element of G is contained in at most finitely many N_n. Clearly then $[G : N_n] \to \infty$. We finish this chapter on extending states of subalgebras in an entropy increasing way.

prop:
extendingautoone

Proposition 3.4.8 ([NS06][pp. 53, Theorem 3.2.7]). *Let α be an automorphism of a unital C*-algebra A, $B \subset A$ an α-invariant C*-subalgebra and ψ an $\alpha_{|B}$-invariant state on B. Assume $(B, \phi_{|B}, \alpha_{|B})$ has a $\phi_{|B}$-approximating net. Then there exists an extension of ψ to A, say ϕ, such that $h_\phi(\alpha) \geq h_\psi(\alpha_{|B})$.*

Proof. Let $\{\gamma_i : B_i \to B\}$ be a $\psi_{|B}$-approximating net and θ_i be the corresponding maps $\theta_i : B \to B_i$. We can assume that the B_i's are full matrix algebras; if they weren't just realize them as subalgebras of $M_{\mathrm{rank}(1_{B_i})}(\mathbb{C})$ and replace γ_i with $E_i \circ \gamma_i$ where $E_i : M_{\mathrm{rank}(1_{B_i})}(\mathbb{C}) \to B_i$ is a conditional expectation. Then Arveson's extension theorem ([BO08][pp. 17, Theorem 1.6.1]) gives us completely positive unital extensions $\overline{\theta_i} : A \to B_i$ of the maps θ_i. Let $\overline{\psi}$ and $\overline{\beta}$ denote the usual extensions of ψ and $\alpha_{|B}$ to $\pi_\psi(B)''$. Consider the net of unital c.p. maps $\{\pi_\psi \circ \gamma_i \circ \overline{\theta_i} : A \to \pi_\psi(B)''\}_i$. For each $a \in A$ the net $\{(\pi_\psi \circ \gamma_i \circ \overline{\theta_i})(a)\}_i$ is bounded so Tychonoff's theorem and the weak operator compactness of the unit ball of $\pi_\psi(B)''$ imply that the net has a pointwise weak operator cluster point $\Psi : A \to \pi_\psi(B)''$.

Now let Φ be a pointwise weak operator cluster point of the net $\{|F_n|^{-1} \sum_{g \in F_n} \beta_g \circ \Psi \circ \alpha_{g^{-1}}\}_n$. For each $s \in G$, $\beta_s \circ \Phi - \Phi \circ \alpha_s$ is a pointwise weak cluster point of

$$|F_n|^{-1} \sum_{g \in F_n; sg \notin F_n} (\beta_{sg} \circ \Psi \circ \alpha_{(sg)^{-1}} - \beta_g \circ \Psi \circ \alpha_{g^{-1}}).$$

Now, since $\{g \in F_n : sg \notin F_n\} \subset F_n \cap (s^{-1}F_n)^c \subset F_n \vartriangle s^{-1}F_n$ it is clear that the maps on the right tend to 0 even in norm. In particular $\beta_s \circ \Phi = \Phi \circ \alpha_s$ for $s \in G$. We now claim that $\Phi(b)h_\psi = \pi(b)h_\psi$ for $b \in B$. Fixing $b' \in B$ and $\epsilon > 0$ it will suffice to show that

$$|\langle (\Phi - \pi)(b)h_\psi, \pi(b')h_\psi \rangle| < \epsilon. \tag{3.28}$$ {eq25}

Begin by picking an $m \in \mathbb{N}$ and $i \in I$ such that

$$|\langle (\Phi - |F_m|^{-1} \sum_{g \in F_m} \beta_g \circ \pi \circ \gamma_i \circ \overline{\theta_i} \circ \alpha_{g^{-1}})(b)h_\psi, \pi(b')h_\psi \rangle| < \epsilon/2.$$

Using that β is $\overline{\psi}$-invariant we have

$$|\langle |F_m|^{-1} \sum_{g \in F_m} \beta_g \circ \pi \circ \gamma_i \circ \overline{\theta_i} \circ \alpha_{g^{-1}}(b) - \pi(b)h_\psi, \pi(b')h_\psi \rangle|$$

$$= |F_m|^{-1} | \sum_{g \in F_m} \psi(b'^*((\gamma_i \circ \overline{\theta_i} \circ \alpha_{g^{-1}})(b) - b))| \le |F_m|^{-1} \sum_{g \in F_m} ||b'||_\psi ||(\gamma_i \circ \overline{\theta_i} \circ \alpha_{g^{-1}})(b) - b||_\psi$$

$$\le |F_m|^{-1} \sum_{g \in F_m} ||b'||_\psi (||(\gamma_i \circ \overline{\theta_i} \circ \alpha_{g^{-1}})(b) - \alpha_{g^{-1}}(b)||_\psi + ||\alpha_{g^{-1}}(b) - b||_\psi)$$

$$= |F_m|^{-1} \sum_{g \in F_m} ||b'||_\psi ||(\gamma_i \circ \overline{\theta_i})(\alpha_{g^{-1}}(b)) - \alpha_{g^{-1}}(b)||_\psi.$$

The last quantity can clearly be dominated by $\epsilon/2$ if we chose i sufficiently large. Equation (3.28) then holds, as desired. It now follows that the state $\phi := \overline{\psi} \circ \Psi$ defined on A is α-invariant and extends ψ. For the former note that $\phi \circ \alpha_s = \overline{\psi} \circ \Psi \circ \alpha_s = \overline{\psi} \circ \beta_s \circ \Psi = \overline{\psi} \circ \Psi$ using that $\Phi \circ \alpha_s = \beta_s \circ \Phi$ and that $\overline{\psi}$ is β-invariant. That Φ extends ψ follows from the fact $\Phi(b)h_\psi = \pi(b)h_\psi$ for $b \in B$.

Finally, $h_\phi(\alpha) \ge h_\psi(\alpha_{|B})$, because any positive linear functional ω on B such that $\omega \le \phi$ extends to the functional $\overline{\omega} = \langle xh_\psi, h_\psi \rangle$ for some $x \in \pi(B)'$ and then $\overline{\omega} \circ \Psi$ extends ω. Since $\overline{\omega} \circ \Psi \le \phi$ we see that any decomposition of ψ extends to a decomposition of ϕ, hence $h_\phi(\alpha) \ge h_\psi(\alpha_{|B})$. ∎

CHAPTER 4

Examples and an Alternative Definition of Entropy

for:fourth

In this final chapter we will apply the theory we have developed so far to compute the entropy of some C*-dynamical systems. First for two non-commutative analogues of the Bernoulli shift and then for type 1 W*-dynamical systems. We finish this chapter by giving an alternative definition of entropy.

4.1 Non-commutative Bernoulli Shifts

Before dealing with non-commutative Bernoulli shifts we need a lemma.

lem:maxabelianincentralizer

Lemma 4.1.1. *Suppose A and C are algebras and $\pi : A \to C$ is a surjective representation and ψ and ϕ are linear functionals on A and C respectively, such that $\phi \circ \pi = \psi$. Then for any maximal abelian subalgebra D of Z_ψ, the centralizer of ψ, $\pi(D)$ will be a maximal abelian subalgebra of Z_ϕ.*

Proof. Clearly $\pi(D)$ is abelian. It also lies in Z_ϕ for if $a \in A$ and $d \in D$ we have $\phi(\pi(a)\pi(d)) = \psi(ad) = \psi(da) = \phi(\pi(d)\pi(a))$. To show that $\pi(D)$ is maximal abelian suppose F is an abelian subalgebra with $\pi(D) \subset F \subset Z_\psi$. We immediately see that $D \subset \pi^{-1}(F)$ and if $a \in A$ and $f \in F$ then

$$\psi(a\pi^{-1}(f)) = \phi(\pi(a\pi^{-1}(f))) = \phi(\pi(a)f) = \phi(f\pi(a)) = \psi(\pi^{-1}(f)a).$$

This shows that $\pi^{-1}(F) \subset Z_\psi$ so since D is maximal abelian in Z_ψ we obtain $\pi^{-1}(F) = D$. Then by the surjectivity of π, $F = \pi(D)$. Hence $\pi(D)$ is maximal abelian in Z_ϕ. ∎

Now recall that the first action we considered in the classical case was the action β on $\{1,\dots,n\}^G$ given by $\beta((x_g)_{g\in G}) = (x_{s^{-1}g})_{g\in G}$ for $x \in \{1,\dots,n\}^G$ and $s \in G$. The operator algebraic analogue is taking a finite dimensional C*-algebra B, a state ψ on B, and consider the shift action α_0 on $B^{\otimes G}$ with respect to the state $\psi^{\otimes G}$. This is called the **non-commutative Bernoulli shift**:

Bernoulli shift 1

Recall that $B^{\otimes G}$ is simply the inductive limit of the net of C*-algebras $\{B^{\otimes F}\}_F$ where F ranges over finite subsets of G ordered under inclusion. The morphisms here are just inclusions.

The left shift action α_0 is formally defined by letting

$$(\alpha_0)_s(\otimes_{g\in F} b_g) = \otimes_{g\in sF} b_{s^{-1}g}.$$

for finite subsets $F \subset G$. To make this as clear as possible, we are mapping an elementary tensor $\otimes_{g\in F} b_g$ in $B^{\otimes F}$ to an elementary tensor in $B^{\otimes sF}$ whose g'th "tensor factor" will be $b_{s^{-1}g}$, $g \in sF$. We let F range over all finite subsets of G and then extend α_0 to an action on $B^{\otimes G}$.

For technical reasons it will be convenient to pass to the GNS-representation of $\psi^{\otimes G}$. We put $M = \pi_{\psi^{\otimes G}}(B^{\otimes G})''$, $\phi = \overline{\psi^{\otimes G}}$, $\alpha = \overline{\alpha_0}$ and $\pi = \pi_{\psi^{\otimes G}}$. We claim that $h_\phi(\alpha) = S(\psi)$, just as for the classical Bernoulli shift. We prove this in three steps.

Step 1: We begin by constructing ϕ-preserving conditional expectations $E_F : M \to \pi(B^{\otimes F})$ for finite sets $F \subset G$. For finite sets $F' \subset G$ that contain F define $E'_F(\bigotimes_{g\in F'} b_g) = (\bigotimes_{g\in F} b_g) \otimes \psi^{\otimes F'\setminus F}(\bigotimes_{g\in F'\setminus F} b_g)$ on $B^{\otimes F'}$. Letting F' range over all finite subsets of G containing F this gives a well-defined conditional expectation on $E'_F : \bigcup_{F'\subset G,\ F' \text{ is finite}} B^{\otimes F'} \to \pi(B^{\otimes F})$ that preserves $\psi^{\otimes G}$. Now define $E_F : \pi(\bigcup_{F'\subset G,\ F' \text{ is finite}} B^{\otimes F}) \to \pi(B^{\otimes F})$ by $E_F \circ \pi = \pi \circ E'_F$. It is ϕ-preserving and weakly continuous on bounded sets. Hence it can be extended to a ϕ-preserving conditional expectation on M which we continue to denote by E_F.

These maps show that the inclusions $\{i_F : \pi(B^{\otimes F}) \to M\}_F$ are ϕ-approximating. Indeed, for $x \in \pi(\bigcup_{F'\subset G,\ F' \text{ is finite}} B^{\otimes F'})$ it is obvious that

$$\phi(((i_F \circ E_F)(x) - x)^*((i_F \circ E_F)(x) - x)) = \|(E_F(x) - x)h_{\psi^{\otimes G}}\| \to 0$$

simply because for sufficiently large F we have $E_F(x) = x$. However, since E_F is strongly continuous, and $\pi(\bigcup_{F'\subset G,\ F' \text{ is finite}} B^{\otimes F})$ is strongly dense in M, the convergence will hold for *all* $x \in M$.

Step 2: We now claim that if $F \subset G$ is a finite subset and D is a maximal abelian subalgebra of Z_ψ, then $\pi(D^{\otimes F})$ is maximal abelian in the centralizer of the restriction of ϕ to $\pi(B^{\otimes F})$.

By Lemma 4.1.1 it suffices to show that $D^{\otimes F}$ is maximal abelian in $Z_{\psi^{\otimes F}}$. By induction it further suffices to verify that $D \otimes D$ is maximal abelian in $Z_{\psi\otimes\psi}$. We have $Z_{\psi\otimes\psi} = Q'_{\psi\otimes\psi} = (Q_\psi \otimes Q_\psi)' = Q'_\psi \otimes Q'_\psi = Z_\psi \otimes Z_\psi$ where the third equality follows from von Neumann's commutation theorem for tensor products. $D \otimes D$ is easily seen to be a an abelian subalgebra of $Z_\psi \otimes Z_\psi$. We have

$$\dim(D\otimes D) = \dim(D)\dim(D) = \mathrm{Tr}_B(I_{Z_\psi})\,\mathrm{Tr}_B(I_{Z_\psi}) = \mathrm{Tr}_{B\otimes B}(I_{Z_\psi}\otimes I_{Z_\psi}) = \mathrm{Tr}_{B\otimes B}(I_{Z_\psi}\otimes Z_\psi)$$

which is the dimension of a maximal abelian subalgebra of $Z_\psi \otimes Z_\psi$. Hence $D \otimes D$ is maximal abelian in $Z_\psi \otimes Z_\psi = Z_{\psi\otimes\psi}$ as desired.

Step 3: Fixing a finite subset $F \subset G$ and $n \in \mathbb{N}$, step 2 implies that for each $s \in F_n$, $\pi(D^{\otimes sF})$ is maximal abelian in $\pi(B^{\otimes sF}) = \pi(\alpha_s(B^{\otimes F}))$. Moreover, as

s ranges over F_n the algebras $\pi(D^{\otimes sF})$ are mutually commuting, and step 1 gave us ϕ-preserving conditional expectations $M \to \pi(B^{\otimes F_n F})$. Lemma 3.3.6 then implies that

$$H_\phi(\alpha_s(\pi(B^{\otimes F})) : s \in F_n) = S(\phi_{|\pi(B^{\otimes F_n F})}) = S(\psi^{\otimes F_n F}) = |F_n F| S(\psi).$$

Here the second equality holds since we can identify $\pi_{|B^{\otimes F_n F}}$ with the GNS-representation of $\psi^{\otimes F_n F}$.

We obtain

$$h_\phi(\pi(B^{\otimes F}); \alpha) = \lim_{n\to\infty} \frac{1}{|F_n|} H_\phi(\alpha_s(B^{\otimes F}) : s \in F_n) = \lim_{n\to\infty} \frac{|F_n F|}{|F_n|} S(\psi) = S(\psi).$$

As noted, the inclusions $\{i_F : \pi(B^{\otimes F}) \to M\}_F$ form a ϕ-approximating net so Theorem 3.4.5 implies that $h_\phi(\alpha) = S(\phi)$.

Bernoulli shift 2

Consider now the specific case where $B \simeq M_n(\mathbb{C})$ for some $n \geq 2$ and the state ψ on B is faithful. Define ϕ, α and M as before and again let D be a maximal abelian subalgebra in Z_ψ. From the above we have $h_\alpha(\phi) = S(\psi)$. But we claim that even if we restrict to the centralizer $N := Z_\phi$, put $\tau = \phi_{|N}$ and $\beta = \alpha_{|N}$ we have $h_\tau(\beta) = S(\psi)$. From modular theory there exists a ϕ-preserving conditional expectation $M \to N$ ([Tak13][pp. 211, Theorem 4.2]), hence $h_\tau(\beta) \leq h_\phi(\alpha) = S(\psi)$ by Proposition 3.4.3 (v).

To show the reverse inequality let C denote the von Neumann subalgebra of N that is generated by the sets $\pi(D^{\otimes F})$ for finite $F \subset G$. We claim that there exists a ϕ-preserving conditional expectation $N \to C$. Clearly τ is tracial so the existence of such a conditional expectation follows if we can show that ϕ is faithful. To see this, we need an alternate way of describing the GNS-representation $\pi_{\psi^{\otimes G}} : B^{\otimes G} \to M$.

Let $\pi_\psi : B \to H_\psi$ be the GNS-representation of ψ and h_ψ the cyclic vector. Analogous to how we constructed $B^{\otimes G}$, let $H_\psi^{\otimes G}$ denote the inductive limit of Hilbert spaces $\{H_\psi^{\otimes F}\}_F$ where F ranges over finite subsets of G and for $F_1 \subset F_2$ the morphism $i_{F_1,F_2} : H_\psi^{\otimes F_1} \to H_\psi^{\otimes F_2}$ is given by $\otimes_{g\in F} h_g \to \otimes_{g\in F} h_g \otimes \otimes_{g\in F_2\setminus F_1} (h_\psi)_g$. We treat $\bigcup_{F\subset G,\ F \text{ finite}} H_\phi^{\otimes F}$ as a dense subspace of $H^{\otimes G}$. Note that because of how we defined our morphisms, h_ψ again is a unit vector of $H^{\otimes G}$.

Define the representation $\pi' : \bigcup_{F\subset G,\ F \text{ finite}} B^{\otimes F} \to B(H^{\otimes G})$ by

$$\pi'(\bigotimes_{g\in F} b_g)(\bigotimes_{g\in F'} h_g) = \bigotimes_{g\in F} \pi_\psi(b_g)(h_g) \otimes \bigotimes_{g\in F'\setminus F'} (h_\psi)$$

whenever $F' \supset F$. One checks that this representation is isometric, hence it extends to a representation $B^{\otimes G} \to B(H^{\otimes G})$ we will continue to denote by π'.

It is easy to verify that $\psi^{\otimes G}(\cdot) = \langle \pi'(\cdot)h_\psi, h_\psi\rangle$ so we recognize $(\pi', H^{\otimes G}, h_\psi)$ as a GNS-triple for $\psi^{\otimes G}$. By the uniqueness of the GNS-representation, if we want to show that ϕ from above is faithful on M, it suffices to show that $\langle\ \cdot h_\psi, h_\psi\rangle$ is faithful on $\pi'(B^{\otimes G})''$. This is equivalent to h_ψ being cyclic for the commutant $\pi'(B^{\otimes G})'$ which is immediate: fix an elementary vector $h = \bigotimes_{g\in F'} h_g \in \bigcup_{F\subset G,\ F \text{ finite}} H_\psi^{\otimes F}$. Since ψ is faithful h_ψ is cyclic for $\pi_\psi(B)'$ so for each $g \in F'$ there exists $T_g \in \pi_\psi(B)'$ such that $T_g h_\psi = h_g$. Defining $T \in B(H_\psi^{\otimes G})$ by putting

$$T(\bigotimes_{g\in F''} h_g) = \bigotimes_{g\in F'} T_g h_g \otimes \bigotimes_{g\in F''\setminus F'} (h_\psi)$$

for finite $F'' \supset F'$. Then $T \in \pi'(B^{\otimes G})'$ and $T(h_\psi) = h$ so h_ψ is cyclic for the commutant.

Since ϕ is faithful there exists a ϕ-preserving conditional expectation $N \to C$. It follows that $h_\tau(\beta) \geq h_{\tau_{|C}}(\beta_{|C})$. But since $D \simeq \mathbb{C}^m$ for some $m \geq 2$ we can identify each $\pi(B^{\otimes F})$ algebra with $C(\{1,\dots,m\}^F)$. The latter algebras can naturally be embedded into $C(\{1,\dots,m\}^G)$, and the union of these embeddings are norm-dense as F ranges over all finite subsets of G. So $C(\{1,\dots,m\}^G)$ can be viewed as a weakly dense C*-subalgebra of C. Hence, $C \simeq L^\infty(\{1,\dots,m\}^G, \mu^{\otimes G})$ where μ denotes the normalized counting measure on $\{1,\dots,m\}$. Under this identification $\beta_{|C}$ simply corresponds to the Bernoulli shift on $L^\infty(\{1,\dots,m\}^G, \mu^{\otimes G})$ and τ corresponds to $\mu^{\otimes G}$. From the classical result we then have $h_\tau(\beta) \geq h_{\tau_{|C}}(\beta_{|C}) = S(\psi)$. This establishes that $h_\tau(\beta) = S(\psi)$.

4.2 Type 1 W*-dynamical systems

There is an interesting special case where the study of non-commutative entropy completely reduces to the classical setting, i.e. abelian systems. Namely, suppose we have a W*-dynamical system (M, ϕ, α) where M is a type 1 von Neumann algebra and $\alpha : G \to \mathrm{Aut}(G)$. That M is type 1 means that every nonzero central projection in M dominates a nonzero abelian projection in M (recall that a projection $p \in M$ is abelian if the algebra pMp is abelian)([Tak79][pp. 296, Definition 1.17]). If we also assume that G contains subgroups of arbitrarily large finite index we can actually prove that $h_\phi(\alpha) = h_{\phi_{|Z}}(\alpha_{|Z})$ where Z denotes the center of M. Before the proof we recall that type 1 von Neumann Algebras are simply direct sum of algebras $A \otimes B(H)$, where the A's are abelian v.n. algebras ([Tak79][pp. 299, Theorem 1.27]). Before proving that $h_\phi(\alpha) = h_{\phi_{|Z}}(\alpha_{|Z})$ we shall need two lemmas.

lem: irreduciblerep

Lemma 4.2.1. *Suppose A and B are unital C*-algebras with A abelian. If $\pi : A \otimes B \to B(H)$ is an irreducible representation there exists a character ϕ on A and an irreducible representation $\rho : B \to B(H)$ such that $\pi = \phi \otimes \rho$.*

Proof. Since A is abelian $\pi(A \otimes 1_B) \subset \pi(A \otimes B)' = \mathbb{C}I$ so $\phi(\cdot) = \pi(\cdot \otimes 1_B)$ defines a character on A. We claim that $\pi(1 \otimes B)' = \pi(A \otimes B)'$. Indeed, take a $T \in \pi(1 \otimes B)'$ and $a \in A$, $b \in B$. Then $T\pi(a \otimes b) = T\pi(a \otimes 1)\pi(1 \otimes b)$, but since

$\pi(a \otimes 1) \in \mathbb{C}I$ this equals $\pi(a \otimes 1)T\pi(1 \otimes b) = \pi(a \otimes 1)\pi(1 \otimes b)T = \pi(a \otimes b)T$ as desired. By linearity and continuity $T \in \pi(A \otimes B)'$. Hence $\pi(1 \otimes B)' = \mathbb{C}I$ so the representation $\rho := \pi(1 \otimes \cdot)$ on B is irreducible. Clearly $\pi = \phi \otimes \rho$. ∎

lem: conditionaltruncation

Lemma 4.2.2 ([NS06][pp. 54, Theorem 3.3.1])**.** *Suppose* (M, ϕ, α) *is a W*-dynamical system and let* $p \in M$ *be an* α*-invariant projection. Setting* $N = pMp + \mathbb{C}(1 - p)$ *we then have* $h_{|Z(N)}(\alpha_{|Z(N)}) \leq h_{|\phi_{Z(M)}}(\alpha_{|Z(M)})$ *where* $Z(M)$ *and* $Z(N)$ *denotes the center of* M *and* N*, respectively.*

Proof. If $p = 1$ there is nothing to prove so suppose $p \neq 1$. Let q denote the central support of p, i.e. the smallest projection in $Z(M)$ that dominates p. Then
$Z(N) = Zp \oplus \mathbb{C}(1-p)$ is isomorphic to $Z(M)q \oplus \mathbb{C}$. Under this isomorphism the state $\phi_{|Z(N)}$ becomes $\phi(p)\psi \oplus \phi(1-p)$ where $\psi = \phi(p)^{-1}\phi(\cdot\, q)_{|Z(M)q}$. Hence $h_{\phi_{|Z(N)}}(\alpha_{|Z(N)}) = \phi(q)h_\psi(\alpha_{|Z(M)q})$.

On the other hand, consider the state $\phi_N = \phi(q)^{-1}\phi_{|Z(M)q}$. Since $Z(M)q + \mathbb{C}(1-q)$ is an α-invariant subalgebra of $Z(M)$, we have $h_{\phi_{|Z(M)}}(\alpha_{Z(M)}) \geq \phi(q)h_{\phi_N}(\alpha_{|Z(M)q})$. Hence we have just to prove that

$$h_\psi(\alpha_{|Z(M)q}) \leq \frac{\phi(q)}{\phi(p)} h_{\phi_N}(|\alpha_{|Z(M)q}).$$

Since $\psi \leq \phi(p)^{-1}\phi(q)\phi_N$ and the function η is monotone for small t this follows from the classical defintion of entropy. ∎

Theorem 4.2.3 ([NS06][pp. 54, Theorem 3.3.1])**.** *Suppose* (M, ϕ, α) *is a W*-dynamical system with* M *a type* 1 *von Neumann algebra and that the group* G *acting on* A *via* α *has subgroups of arbitrarily large finite index. Let* $Z = Z(M)$ *be the center of* M*. If* (B, ψ, β) *is a W*-dynamical system where the group acting on* B *is also* G *has a* ψ*-approximating net, then*

$$h_{\psi\otimes\phi}(\beta \otimes \alpha) = h_\psi(\beta) + h_{\phi_{|Z}}(\alpha_{|Z}).$$

In particular, taking $B = \mathbb{C}$*, we see that*

$$h_\phi(\alpha) = h_{\phi_{|Z}}(\alpha_{|Z}).$$

Proof. Since the modular group acts trivially on the center, there exists a ϕ-preserving conditional expectation $M \to Z$ ([Tak13][pp. 211, Theorem 4.2]). Hence the inequality $\geq$ follows from Proposition 3.4.3 (iii) and (v). To prove the opposite inequality, we will first prove it in the case where $M = Z \otimes M_n(\mathbb{C})$. Let $\{\gamma_i\}_i$ be a ψ-approximating net and $\{\gamma'_k\}_k$ a $\phi_{|Z}$-approximating net (the latter exists because Z is an abelian von Neumann algebra, hence nuclear). It is then easy to see that $\{\gamma_i \otimes \gamma'_k \otimes \mathrm{id}_{|M_n(\mathbb{C})}\}_{(i,k)}$ is $\psi \otimes \phi$-approximating. By Proposition 3.3.7(v) we have

$$h_{\psi\otimes\phi}(\gamma_i \otimes \gamma'_k \otimes \mathrm{id}_{|M_n(\mathbb{C})}; \beta \otimes \alpha) \leq h_{\psi\otimes\phi_{|Z}}(\gamma_i \otimes \gamma'_k \otimes \mathrm{id}_{|M_n(\mathbb{C})}; \beta \otimes \alpha_{|Z}) + 2\log n.$$

Applying this inequality to $\alpha_{|H}$ and $\beta_{|H}$ for some finite index subgroup H of G, using Theorem 3.4.7(i), and letting $[G : H] \to \infty$ reveals that

$$h_{\psi\otimes\phi}(\gamma_i \otimes \gamma'_k \otimes \mathrm{id}_{|M_n(\mathbb{C})}; \beta \otimes \alpha) \leq h_{\psi\otimes\phi_{|Z}}(\gamma_i \otimes \gamma'_k \otimes \mathrm{id}_{|M_n(\mathbb{C})}; \beta \otimes \alpha_{|Z})$$

$$= h_\psi(\beta) + h_{\phi_{|Z}}(\alpha_{|Z}),$$

where the last equality follows from Theorem 3.4.7(iii).

We will now prove $\leq$ in the case where M is a direct sum of algebras of the form $A \otimes M_n(\mathbb{C})$ where A is abelian. By using that $\oplus$ distributes over $\otimes$ we group together the Z's corresponding to the same $M_n(\mathbb{C})$ algebra and get that $M \simeq \bigoplus_{n=1}^{\infty} Z_n \otimes M_n(\mathbb{C})$ with the Z_n's abelian. For each n, let $1_n \in M$ denote the unit of the n'th summand $Z_n \otimes M_n(\mathbb{C})$.

We claim that α_s fixes 1_n for each $s \in G$. For each $k \in \mathbb{N}$, $\alpha_s(1_n)_k$ is a central projection in $Z_k \otimes M_k(\mathbb{C})$. Identifying Z_k with $C(X)$ for a compact Hausdorff space X we have $Z_K \otimes M_k(\mathbb{C}) \simeq C(X; M_k(\mathbb{C}))$, the space of continuous $M_k(\mathbb{C})$-valued functions on X. The only projections here are functions $\sum \mathbb{1}_{A_i} p_i$ where the $A_i \in X$ are pairwise disjoint and the $p_i \in M_k(\mathbb{C})$ are projections. This is to say that there are mutually orthogonal projections q_i such that $\alpha_s(1_n)_k = \sum_i q_i \otimes p_i$. Since $\alpha_s(1_n)_k$ is central the p_is are central so we actually obtain $\alpha_s(1_n)_k = q \otimes \mathrm{id}_{M_k(\mathbb{C})}$ for a projection $q \in Z_k$.

The same argument reveals that $\alpha_s^{-1}(\alpha_s(1_n)_k)$ too is of the form $p \otimes \mathrm{id}_{M_n(\mathbb{C})}$ for some projection $p \in Z_n$. Then α_s gives an isomorphism of $Z_n p \otimes M_n(\mathbb{C})$ and $Z_k q \otimes M_k(\mathbb{C})$. For $k \neq n$ Lemma 4.2.1 implies that this is impossible unless p and q are zero: any irreducible representation on $Z_n p \otimes M_n(\mathbb{C})$ is the tensor product of a character on $Z_n p$ and an irreducible representation on $M_n(\mathbb{C})$, but these are respectively 1- and n-dimensional so their tensor product is n-dimensional. Similarly, if $q \neq 0$ any irr. rep. on $Z_k p \otimes M_n(\mathbb{C})$ is k-dimensional so they cannot possibly be isomorphic for $k \neq n$.

We conclude that $\alpha_s(1_n)_k = 0$ for $n \neq k$ implying that α_s fixes 1_n.

We will now see that, in general, if (N, ϕ, γ) is a W*-dynamical system, with $\{z_n\} \subset N$ a sequence of γ-invariant central projections summing to 1, then

$$h_\phi(\gamma) = \sum_n \lambda_n h_{\phi_n}(\gamma_{|N z_n}) \tag{4.1}$$

{eq30}

where $\lambda_n = \phi(z_n)$ and $\phi_n = \lambda_n^{-1} \phi_{|N z_n}$. Set

$$N_n = N(z_1 + \ldots + z_n) + \mathbb{C}(1 - z_1 - \ldots - z_n).$$

Since $\{z_1 + \ldots + z_n\}_n$ converges to 1 strongly and multiplication is SOT-continuous when a factor is fixed we see that $\bigcup_n N_n$ is SOT-dense in N. There exists ϕ-preserving conditional expectations $N \to M_n$ so by Proposition 3.4.3(v) we have $h_{\phi_{|M_n}}(\gamma_{|M_n}) \leq h_\phi(\gamma)$ for each n. On the other hand, by Proposition 3.4.6 we have $h_\phi(\gamma) \leq \liminf_n h_{\phi_{|M_n}}(\gamma_{|M_n})$. Hence $h_{\phi_{|M_n}}(\gamma_{|M_n}) \to h_\phi(\gamma)$ so noting that

$$h_{\phi_{|M_n}}(\gamma_{|M_n}) = \sum_{k=1}^{n} \lambda_k h_{\phi_k}(\gamma_{N z_k})$$

by Proposition 3.4.3(iv), we have established Equation (4.1). Applying Equation (4.1) to the α-invariant central projections 1_n, $n \in \mathbb{N}$, summing to 1 and noting that the formula $h_{\psi\otimes\phi}(\beta\otimes\alpha) = h_\psi(\beta) + h_{\phi_{|Z}}(\alpha_{|Z})$ holds on each summand $Z_n \otimes M_n(\mathbb{C}))$ we obtain

$$\begin{aligned} h_{\psi\otimes\phi}(\beta\otimes\alpha) = \sum_n \lambda_n(h_\psi(\beta) + h_{\phi_{|Z_n}}(\alpha_{|Z_n})) = h_\psi(\beta) + \sum_n \lambda_n h_{\phi_{|Z_n}}(\alpha_{|Z_n}) \\ = h_\psi(\beta) + h_\phi(\alpha_{Z(M)}) \end{aligned}$$

.
Hence the desired formula is established in the case where M is a direct sum of algebras of the form $A \otimes M_n(\mathbb{C})$ with A abelian.

We have now to establish the inequality

$$h_{\psi\otimes\phi}(\beta\otimes\alpha) \leq h_\psi(\beta) + h_{\phi_{|Z}}(\alpha_{|Z}) \tag{4.2}$$

{eq35}

in the general case. We first argue note that if q denotes the support of ϕ, then

$$h_{\psi\otimes\phi}(\beta\otimes\alpha) = h_{\psi\otimes\phi_{|qMq}}(\beta\otimes\alpha_{|qMq}) \tag{4.3}$$

{eq67}

. To see this it suffices to verify that if (C,φ,γ) is a C*-dynamical such that φ has a support projection p, i.e. a smallest projection p such that $\varphi(p) = 1$, then $h_\varphi(\gamma) = h_{\varphi_{|pAp}}(\gamma_{|pAp})$. Indeed, $\phi(\gamma_g(p)) = 1$ so $\gamma_g(p) \geq p$. Similarly $\gamma_{g^{-1}}(p) \geq p$ so in fact $\gamma(p) = p$ so the subalgebras pAp, $pA(1-p)$, $(1-p)Ap$ and $(1-p)A(1-p)$ are all invariant under γ. Since ϕ vanishes on the subalgebras $pA(1-p)$, $(1-p)Ap$ and $(1-p)A(1-p)$ Proposition 3.4.3(iii) now implies that $h_\varphi(\gamma) = h_{\phi_{|pAp}}(\gamma_{|pAp})$. Then Equation (4.3) is established so if we manage to prove Equation (4.2) with M replaced by qMq, we will get

$$\begin{aligned} h_{\psi\otimes\phi}(\beta\otimes\alpha) = h_{\psi\otimes\phi_{|qMq}}(\beta\otimes\alpha_{|qMq}) \leq h_\psi(\beta) + h_{\phi_{|Z(qMq)}}(\alpha_{|Z(qMq)}) \\ = h_\psi(\beta) + h_{\phi_{|Z(qMq+\mathbb{C}(1-q))}}(\alpha_{|Z(qMq+\mathbb{C}(1-q))})) \leq h_\psi(\beta) + h_{\phi_{|Z}}(\alpha_{|Z}), \end{aligned}$$

as desired. Here the last inequality is due to Lemma 4.2.2.

Hence, in proving Equation (4.2) we can replace M by pMp so we may suppose ϕ is faithful. In particular M can have at most countably many mutually orthogonal projections. Then there is a sequence of α-invariant finite projections $\{p_n\}_n \subset M_\phi$ such that $\|1 - p_n\|_\phi \to 0$. We do not prove this, see [NS06][pp. 54, Theorem 3.3.1].

Now set $M_n = p_nMp_n + \mathbb{C}(1-p_n)$. Then $E_n : M \to M_n$ defined by $E_n(x) = p_nxp_n + \phi(a(1-p_n))(1-p_n)$ is a ϕ-preserving conditional expectation. Since $\|1-p_n\|_\phi \to 0$ it then follows that $\|x - E_n(x)\|_\phi \to 0$ for all $x \in M$. Proposition 3.3.8 then implies that given a channel $\gamma : C \to B\otimes M$, $h_\phi(\gamma;\beta\otimes\alpha)$ can be approximated by $h_\phi((\mathrm{id}_B \otimes E_n)\circ\gamma;\beta\otimes\alpha)$ for $n \in N$ so in fact

$$h_{\psi\otimes\phi}(\beta\otimes\alpha) = \lim_n h_{\psi\otimes\phi_{|M_n}}(\beta\otimes\alpha_{|M_n}).$$

Note that M_n is again a type 1 algebra so it is a direct sum of $A\otimes B(H)$ algebras with A abelian, but since the unit of M_n is finite, due to p_n being finite, any H occuring in such a decomposition is finite dimensional. Letting Z_n denote the center of M_n we can then apply the previous case to get

$$h_{\psi\otimes\phi_{|M_n}}(\beta\otimes\alpha_{|M_n}) \leq h_\psi(\beta) + h_{\phi_{|Z_n}}(\alpha_{|Z_n}).$$

This finishes the proof since by Lemma 4.2.2 $h_{\phi_{|Z_n}}(\alpha_{|Z_n}) \leq h_{\phi_{|Z}}(\alpha_{|Z})$ ∎

4.3 An alternative definition of entropy via stationary couplings

In this final section we will discuss a generalized version of abelian models. Lemma 3.3.3 tells us that given a C*-dynamical system (A, ϕ, α) and a channel γ into A and a finite subset $F \subset G$, the quantity $H_\phi(\alpha_g \circ \gamma : g \in F)$ can be estimated by $H_\mu(P \circ \alpha_g \circ \gamma : g \in F)$ where P is a unital completely positive map from A into a finite dimensional abelian C*-algebra C and μ satisfies $\mu \circ P = \phi$. If in addition $\beta : G \to \mathrm{Aut}(C)$ was a group action equivariant to α in the sense that $\beta_g \circ P = P \circ \alpha_g$ for all $g \in G$ we have

$$H_\phi(\alpha_g \circ \gamma : g \in F) = H_\phi(\beta_g \circ P \circ \gamma : g \in F).$$

Thus there is the hope that letting F range over a Følner sequence and dividing by $|F_n|$, the quantity $h_\phi(\gamma; \alpha) = \lim_n \frac{1}{|F_n|} H_\phi(\alpha_g \circ \gamma : g \in F_n)$ can be approximated by quantities

$$\frac{1}{|F_n|} H_\mu(\beta_g \circ P \circ \gamma : g \in F_n) \approx h_\mu(P \circ \gamma; \beta)$$

for G-equivariant systems (C, β, μ) with C finite dimensional and abelian. Of course, this is a far cry from a proof; it is not clear that the β-action exists for a given pair (P, μ) and even if it does the above argument requires an exchange of limits and suprema. However, we will see that if we consider all abelian G-equivariant systems (C, β, μ), for also infinite dimensional abelian C, then these can be used to compute the entropy $h_\phi(\alpha)$.

First and foremost, it will be useful to deal with states instead of unital completely positive (u.c.p.) maps. More precisely, suppose $(L^\infty(X, \mu), \mu, \beta)$ is a W*-dynamical system where β is induced by an action $\beta' : G \curvearrowright (X, \mu)$, i.e. $\beta_g(f) = f \circ \beta'_g$. If $P : A \to L^\infty(X, \mu)$ is a unital completely positive map satisfying $\mu \circ P = \phi$, consider the state λ on $A \otimes L^\infty(X, \mu)$ defined by

$$\lambda(a \otimes f) = \int_X P(a) f \; d\mu \;, a \in A.$$

It satisfies $\lambda_{|A} = \phi$ and $\lambda_{|L^\infty(X,\mu)} = \mu$. In fact, any state on $A \otimes L^\infty(X, \mu)$ that restricts to ϕ on A and μ on $L^\infty(X, \mu)$ arises in this way: for such λ in $S(A \otimes L^\infty(X, \mu))$, take $a \in A^+$ and consider the positive functional $\lambda(a \otimes \cdot)$ on $L^\infty(X, \mu)$. It is dominated by $\|a\|\mu$ so by general theory it is of the form $\int_X \cdot \, g \; d\mu$ for some $g \in L^\infty(X, \mu)$. Define $P(a) = g$. This furnishes a positive unital map $P : A \to L^\infty(X, \mu)$ which is completely positive by Theorem 3.1.4.

The above establishes a bijection between u.c.p. maps $P : A \to L^\infty(X, \mu)$ such that $\mu \circ P = \phi$ and states on $A \otimes L^\infty(X, \mu)$ restricting to ϕ on A and μ on $L^\infty(X, \mu)$. These states are of course couplings which we already discussed in the previous chapter. If in addition, $P \circ \alpha_g = \beta_g \circ P$ for $g \in G$ the corresponding λ will be $\alpha \otimes \beta$-invariant and vice versa. We call such λ a **stationary coupling** of (A, ϕ, α) with (X, μ, β). Since we know more about states then c.p. maps we phrase the below definition in terms of stationary couplings and not G-equivariant maps.

def: SauvageotThouvenotEntropy

Definition 4.3.1 ([NS06][pp. 77, Definition 5.1.1])**.** Suppose (A, ϕ, α) is a C*-dynamical system (A, ϕ, α) with a ϕ-approximating net $\{\gamma_i : A_i \to A\}_{i \in I}$. The *Sauvageot-Thouvenot entropy* is then

$$h_\phi^{ST}(\alpha) = \sup\{h_\mu(\mathcal{P}; \beta) + \lim_i \sum_{A \in \mathcal{P}} S(\lambda(\gamma_i(\cdot) \otimes \mathbb{1}_A), \phi)\}$$

where the supremum is taken over all stationary couplings λ of (A, ϕ, α) with (X, μ, β) and over all finite measurable partitions $\mathcal{P}$ of X.

We have phrased the above definition in terms of stationary couplings, but as discussed above we could equivalently express it in terms of G-equivariant maps between (A, ϕ, α) and a classical system $(L^\infty(X, \mu), \mu, \beta)$. From that point of view the term $h_\mu(\mathcal{P}; \beta)$ in the definition of $h_\phi^{ST}(\alpha)$ estimates the entropy of $(L^\infty(X, \mu), \mu, \beta)$. The remaining term $\lim_i \sum_{A \in \mathcal{P}} S(\lambda(\gamma_i(\cdot) \otimes \mathbb{1}_A), \phi)$ can be considered an error term. Technically we have not showed that this limit exists. This requires defining relative entropy $S(\cdot, \cdot)$ for states on *infinite* dimensional C*-algebras. This is done in [NS06][pp. 26].

There is an alternative way of viewing Definition 4.3.1. Namely, recall that given a coupling λ of (A, ϕ, α), a partition $\mathcal{P}$ of (X, μ) and a finite subset $F \subset G$, the partition $\mathcal{P}^F$ induces a decomposition of ϕ as in Equation (3.15). Using the notation proceeding that equation we have, by Equation (3.21),

$$H_\lambda(\{\alpha_g \circ \gamma : g \in F_n\} : \mathcal{P}^{F_n}) = H_\mu(\mathcal{P}^{F_n}) - |F_n| H_\mu(\mathcal{P}) + |F_n| H_\lambda(\gamma; \mathcal{P}) \tag{4.4}$$

{eq58}

Dividing the RHS by $\frac{1}{|F_n|}$, taking the limit as $n \to \infty$ and using Equation (3.19), we get

$$\begin{aligned} &h_\mu(\mathcal{P}; \beta) - H_\mu(\mathcal{P}) + (H_\mu(\mathcal{P}) + \sum_{A \in \mathcal{P}} S(\lambda(\gamma(\cdot) \otimes \mathbb{1}_A), \phi)) \\ &= h_\mu(\mathcal{P}; \beta) + \sum_{A \in \mathcal{P}} S(\lambda(\gamma(\cdot) \otimes \mathbb{1}_A), \phi) \end{aligned}$$

which is similar to Definition 4.3.1.

As mentioned, it is possible to define relative entropy, $S(\cdot, \cdot)$, for states defined for arbitrary C*-algebras. Had we taken this route we could have defined the Sauvageot-Thouvenot entropy for arbitrary C*-dynamical systems by replacing the quantity $\lim_i \sum_{A \in \mathcal{P}} S(\lambda(\gamma_i(\cdot) \otimes \mathbb{1}_A), \phi)$ with $\sum_{A \in \mathcal{P}} S(\lambda(\cdot \otimes \mathbb{1}_A), \phi)$.

To see why stationary couplings can be useful in the first place we show how the existence of a coupling that is *not* the tensor product state $\phi \otimes \mu$ implies that the system has positive entropy. We call such couplings nontrivial.

Proposition 4.3.2 ([NS06][pp. 54, Theorem 3.3.1])**.** *Suppose (A, ϕ, α) is a C*-dynamical system where α is the action of a residually finite group G. If (A, ϕ, α) has a nontrivial stationary coupling λ with a classical Bernoulli shift, (X, μ, β), then $h_\phi(\alpha) > 0$.*

Proof. Let $\mathcal{P}$ be the standard generating partition for (X, μ, β). If for all self-adjoint elements $a \in A$ and all sets $C \in \bigcup_{F \subset G,\, F \text{ finite}} \mathcal{P}^F$ we had $\lambda(a \otimes \mathbb{1}_C) = \phi(a)\mu(C)$ then by linearity and continuity $\lambda = \phi \otimes \mu$. We have assumed that this is not the case so pick a self-adjoint element $a \in A$ and $C \in \mathcal{P}^F$, for some finite $F \subset G$, such that $\lambda(a \otimes \mathbb{1}_C) \neq \phi(a)\mu(C)$. Since G is residually finite there is an increasing sequence of finite index subgroups $\{N_n\}_n \subset G$, such that any non-trivial element lies in only finitely many of them. Hence we can find an N_n that does not contain any of the elements $f^{-1}f'$ for distinct $f, f' \in F$. This ensures that all members of F lie in distinct left cosets of N_n.

Now let H be a set of representatives of left cosets of N_n containing F. Since $\mathcal{P}^H \geq \mathcal{P}^F$ and $C \in \mathcal{P}^F$ with $\lambda(a \otimes \mathbb{1}_C) \neq \phi(a)\mu(B)$, there must be a $B \in \mathcal{P}^H$ such that $\lambda(a \otimes \mathbb{1}_B) \neq \phi(a)\mu(B)$ Clearly λ is again a stationary coupling of $(A, \phi, \alpha_{|N_n})$ with $(X, \mu, \beta_{|N_n})$. The latter system is again the Bernoulli shift and $\mathcal{P}^H$ is a generating partition. Furthermore, we know that $h_\phi(\alpha) = [G : N_n] h_\phi(\alpha_{|N_n})$ so to prove that that $h_\phi(\alpha) > 0$ it suffices to prove that $h_{\phi_{|N_n}}(\alpha_{|N_n}) > 0$. Choose an $\kappa > 0$ such that $1 - \kappa a$ is positive and define the channel $\gamma : \mathbb{C}^2 \to A$ by setting $\gamma(e_1) = \kappa a$ and $\gamma(e_2) = 1 - \kappa a$. By Equation (4.4) and Equation (3.19) we get,

$$h_\phi(\gamma; \alpha_{|N_n}) \geq H_\lambda(\gamma; \mathcal{P}^H) = \sum_{Z \in \mathcal{P}^H} \mu(Z) S(\mu(Z)^{-1} \lambda(\gamma(\cdot) \otimes \mathbb{1}_Z), \phi \circ \gamma).$$

Since $a \in \mathrm{Im}(\gamma)$ and $\lambda(a \otimes \mathbb{1}_B) \neq \phi(a)\mu(B)$ we have $\mu(B)^{-1}\lambda(\gamma(\cdot) \otimes \mathbb{1}_B) \neq \phi \circ \gamma$, which means that at least one of the terms in the above sum is non-zero. Hence $h_\phi(\alpha_{|N_n}) > 0$ as desired. ■

The above result gives some idea why the Sauvageot-Thouvenot entropy, $h_\phi^{ST}(\alpha)$, relates to the standard entropy $h_\phi(\alpha)$. In fact one can show that the two coincide, see [NS06][pp. 82, Theorem 5.1.5]. This is nice result because the Sauvageot-Thouvenot entropy is often easier to work with. For example, that $h_\phi^{ST}(\alpha) = h_\phi(\alpha)$ can be used to show the following two results:

prop: extendingautotwo

Proposition 4.3.3 ([NS06][pp. 86, Proposition 5.1.7]). *Let α be an automorphism of a nuclear C*-algebra A, $B \subset A$ an α-invariant C*-subalgebra, ψ an α-invariant state on B. Then for every $h < h_\phi(\alpha_{|B})$ there exists an α-invariant state ϕ on A such that $\phi_{|B} = \psi$ and $h_\phi(\alpha) > h$.*

Proposition 4.3.4 ([NS06][pp. 86, Proposition 5.1.8]). *Let (M, ϕ, α) be a W*-dynamical system having a ϕ-approximating net, $N \subset M$ an α-invariant von Neumann subalgebra, $E : M \to N$ a ϕ-preserving faithful normal conditional expectation commuting with α. Assume there exists a constant $c > 0$ such that $E(a) \geq ca$ for any $a \geq 0$. Then $h_\phi(\alpha) = h_{\phi_{|N}}(\alpha_{|N})$.*

We remark that Proposition 4.3.3 is very similar to 3.4.8 except for the nuclearity of A being an assumption in the former and the existence of a ψ-approximating net on B an assumption in the latter. The conclusion Proposition 4.3.3 is of course, slightly weaker.

Final remarks

In this thesis we have essentially explored two generalizations of amenable classical entropy: sofic classical entropy and amenable operator algebraic entropy. Indeed, we actually proved that the latter is a generalization and that the former is a generalization is proven in [KL10]-this proof is quite involved. A natural question then is if it is possible to develop a theory of sofic operator algebraic entropy. To this day, this is an open research problem.

Even proving that sofic classical entropy generalizes amenable classical entropy is difficult and naturally we would expect a proof that some notion of sofic operator algebraic entropy generalizes amenable operator algebraic entropy to be at least as difficult. We could expect the notion of sofic operator algebraic entropy to combine channels and a sofic approximation sequence in some clever way, though this may be too naive. However, a good way of checking that a notion of sofic operator algebraic entropy is good, once we have one, is to see whether the entropy of a non-Commutative Bernoulli shift on $B^{\otimes G}$ with respect to $\psi^{\otimes G}$ is again determined by B and ψ. This thesis also contains a number of operator algebraic results and properties we might be able to check in the sofic case.

Bibliography

BO08 [BO08] Brown, N. P. and Ozawa, N. *C*-Algebras and Finite-Dimensional Approximations*. Vol. 88. American Mathematical Soc., 2008.

BR12 [BR12] Bratteli, O. and Robinson, D. W. *Operator Algebras and Quantum Statistical Mechanics: Volume 1: C*-and W*-Algebras. Symmetry Groups. Decomposition of States*. Springer Science and Business Media, 2012.

FR03 [Fra03] Fraleigh, J. B. *A first course in abstract algebra*. Pearson Education India, 2003.

KL10 [KL10] Kerr, D. and Li, H. "Soficity, amenability, and dynamical entropy". In: *arXiv e-prints*, arXiv:1008.1429 (Aug. 2010), arXiv:1008.1429. arXiv: `1008.1429 [math.DS]`.

KL16 [KL16] Kerr, D. and Li, H. "Ergodic theory". In: *Springer Monographs in Mathematics. Springer, Cham* (2016).

MW99 [MW99] McDonald, J. N. and Weiss, N. A. *A course in real analysis*. Taylor and Francis US, 1999.

NS06 [NS06] Neshveyev, S. and Stormer, E. *Dynamical entropy in operator algebras*. Vol. 50. Springer Science and Business Media, 2006.

PE12 [Ped12] Pedersen, G. K. *Analysis now*. Vol. 118. Springer Science and Business Media, 2012.

TA13 [Tak13] Takesaki, M. *Theory of operator algebras II*. Vol. 125. Springer Science and Business Media, 2013.

TA79 [Tak79] Takesaki, M. *Theory of operator algebras I*. Vol. 125. Springer Science and Business Media, 1979.