On the family of q-entropies

Evgueni Verbitski

Vakgroep
Wiskunde

RuG

Master thesis

On the family of q-entropies

Evgueni Verbitski

Rijksuniversiteit Groningen
Vakgroep Wiskunde
Postbus 800
9700 AV Groningen

June 1996

This work is a Master's thesis performed at the department of Mathematics, Groningen University under the supervison of Prof.Dr.F.Takens.

CONTENTS

0. Introduction.

Entropy is one of the major qualitative characteristics in the determination of chaos. We call it qualitative because dynamical systems are classified by the entropy h as follows

- $h = 0$ for ordered systems,
- $h \in (0, \infty)$ for chaotic deterministic systems,
- $h = \infty$ for random systems.

We understand entropy as an indicator for chaos. Moreover, one can understand entropy as a measure for sensitive dependence on initial conditions,or as a measure of unpredictability.

We start with describing the settings. We consider a map $f : X \to X$ on some space X, describing discrete time evolution

$$x_{n+1} = f(x_n).$$

In practice, we can have 3 situations:

(1) f is given by the "exact" equations;
(2) f can be computed in every given point;
(3) we can only know one finite, but sufficiently long, orbit $\{x_0, x_1, x_2, \dots\}$. where $x_{i+1} = f(x_i)$.

The first case is almost never happens in real life. Moreover, for the first two situations one can compute the Lyapunov exponents, which can be used as indicators for chaos too, due to relations between different characteristics of dynamical systems. One has to mention in this relation the Kaplan–Yorke and Eckmann–Ruelle conjectures, and the corresponding results of F.Ledrapier [Led-81] and L.Young [Young-82].

The third case is the most problematic. After observing a seemingly chaotic signal in the laboratory, the researcher is faced with the question how to characterize the signal, how to be sure that it is chaotic, and how to quantify "how" chaotic it is.

The direct application of the classical definition of entropy leads to the computational problems (e.g., difficulties with taking the limits and low confidence of the obtained results). Therefore, methods to estimate entropy directly from a time signal can answer some of the above questions. P.Grassberger and I.Procaccia in [GraPro-83] proposed such method. The general idea is the following. We consider some family of quantities, which are related to entropy and can be estimated from a time signal. They introduced a one–parameter family of order-q Renyi entropies

$$K(q) = \lim_{\varepsilon \to 0} \lim_{k \to \infty} -\frac{1}{(q-1)k} \log\Big(\sum_{i_1, \dots, i_k} p^q(i_1, \dots, i_k) \Big), \tag{0.1}$$

where a finite dimensional space is partitioned into cubes $\{C_i\}$ of size ε, and

$$p(i_1, \dots, i_k) = \mathbb{P}(\{x : x \in C_{i_1}, \dots, f^{k-1}(x) \in C_{i_k}\}).$$

A practical application is clear, one can estimate $p(i_1,\dots,i_k)$ quite easily, and the averaging procedure should improve the confidence. The original paper by P.Grassberger and I.Procaccia had some motivation for using this family of order-q Renyi entropies and stated few conjectures without proofs. We rewrite them as

(C1) $h = \lim_{q\to 1} K(q)$.

(C2) $h_{top} = \lim_{q\to 0} K(q)$, where h_{top} is the topological entropy.

In the present work another, but very similar, family of generalized entropies is considered. This family was introduced by F.Takens and it naturally arises in the reconstruction theory. The main advantage is that the partitions of the phase space do not enter in the consideration, as it was in [GraPro-83].

Namely, let $f: X \to X$ be a continuous map on the compact metric space (X,d). For any integer $k > 0$ define a metric

$$d_k(x,y) := \max_{i=0,\dots,k-1} d(f^i x, f^i y).$$

Consider any invariant Borel probability measure μ on X. For any $q \geqslant 0$, $q \neq 1$, each $\varepsilon > 0$ and any positive integer k we introduce

$$A^{(q)}(k,\varepsilon) = -\frac{1}{(q-1)k}\log\int_X \mu(\mathcal{B}_k(x,\varepsilon))^{q-1}\,d\mu, \tag{0.2}$$

where $\mathcal{B}_k(x,\varepsilon) = \{y : d_k(x,y) < \varepsilon\}$ is the open ε–neighborhood of x in the metric d_k. We study the asymptotic behavior of (0.2) as $k \to \infty$ and $\varepsilon \to 0$,

$$H^{(q)} = \lim_{\varepsilon\to 0}\lim_{k\to\infty} A^{(q)}(k,\varepsilon). \tag{0.3}$$

The spectrum (0.3) was introduced as an approximation for the spectrum in (0.1). So, one has to expect similar properties from both families.

It turns out that the generalized spectrum of the Renyi entropies is closely related to the similar spectrum for dimensions. In general they have the same global properties, one may call this fact a "duality". For example, a phase transition was found first for the dimensions and later for the entropies, one can transfer the notion of multifractal formalism to the case of generalized entropies. Therefore, we discuss the spectrum for dimensions too.

In the last years a large number of papers has been devoted to the properties of the Renyi dimension; we refer to [Pes-93]. The equivalent mathematical theory for the Renyi entropies has not been developed. The main aim of this work is to start the investigation of the generazlized entropies on the rigurous mathematical level.

Renyi in [Ren-70] introduced a generalization of the standard Shannon's information function $I(p) = -\sum_i p_i \log p_i$. For any probability vector $p = (p_1,\dots,p_n)$ with $p_i \geqslant 0$ and $\sum_i p_i = 1$ and $q \in \mathbb{R}$, $q \neq 1$ define

$$I_q(p) = -\frac{1}{q-1}\log\Big(\sum_i p_i^q\Big).$$

One can easily see that $\lim_{q\to 1} I_q(p) = I(p)$ for any fixed p, but this convergence is not uniform in p and this fact leads to a "phase transition" phenomenon described below.

The functions I_q lead to various generalizations. We introduce the generalized spectrums for dimensions and entropies based on these information functions.

Renyi dimensions. For a probability measure μ on a $\mathbb{R}^n$ take a uniform partition by boxes of size r and define

$$\begin{aligned} D(q) &= \lim_{r\to 0} \frac{1}{q-1} \frac{\log \sum_i p_i^q}{\log r} \text{ for } q \in \mathbb{R}, q \neq 1, \\ &= \lim_{r\to 0} \frac{\sum_i p_i \log p_i}{\log r} \text{ for } q = 1. \end{aligned} \tag{0.4}$$

where p_i is the measure of the i-th cell, and the sum is taken over all indices i such that $p_i \neq 0$. Assume that all limits exist.

One can show that almost all known dimensions belong to the family of Renyi dimensions. $D(0)$ is the limit capacity, which often coincides with the Hausdorff dimension. $D(1)$ is the information dimension, which describes how entropy increases with scaling and $D(2)$ is the correlation dimension. Usually $D(q)$ is continuous at $q = 1$, but see [Beck-90] for the example where $D(q) = 1$ for $q < 1$, $D(q) = 0$ for $q > 1$ and $D(1)$ can take an arbitrary value in the interval $(0, 1)$. This discontinuity at $q = 1$ is called *a phase transition* in the physical literature.

For the spaces with a finite dimension (or, more general, with a finite covering dimension) (0.4) coincides with the following (probably, more common) definition, let μ be a Borel probability measure on some separable metric space (X, d), define

$$D(q) = \lim_{r\to 0} \frac{1}{q-1} \frac{\log \int \mu(B(x,r))^{q-1}\, d\mu}{\log r}, \text{ for } q \neq 1,$$

where $\mathcal{B}(x, r)$ is an open ball in (X, d) with the center at x and the radius r.

Renyi entropies. The standard definition of entropy can be generalized as follows. We introduce the order-q Renyi entropy of a partition ξ as

$$H_{R,\mu}^{(q)}(\xi) = -\frac{1}{q-1} \log\Big(\sum_{C\in\xi} \mu^q(C)\Big).$$

For the arbitrary partitions ξ and η denote by $\xi \vee \eta$ the partition of consisting of all nonempty sets of the form $C \cap D$, where $C \in \xi$ and $D \in \eta$. And consider the limit (if it exists)

$$h_\mu^{(q)}(f;\xi) = \lim_{k\to\infty} \frac{1}{k} H_{R,\mu}^{(q)}(\xi^{(k)}),$$

where $\xi^{(k)} = \xi \vee f^{-1}\xi \vee \ldots \vee f^{-k+1}\xi$.

The phenomenon of a phase transition has been found in the case of the Renyi entropies too, see [CsoSze2-88], [STCK-87]. It was shown that for some smooth chaotic dynamical systems, but not hyperbolic everywhere, the spectrum can have a singularity at $q = 1$. Thus, the idea of P.Grassberger and I.Procaccia of using $K(2)$ as a good approximation for the entropy, and consequencely as an indicator for the chaotic behavior does not work for some chaotic dynamical systems. In other words, the condition $K(2) > 0$ is sufficient, but not necessary for chaos.

In the present work we are starting the development of a mathematical frame for generalized entropies defined as limits of the family (0.2). Some of the results, which are valid for the Renyi entropies too, are known in the physical literature, but they do not have satisfactory mathematical proofs.

This work is organized in the following way. In Section 1 we derive the basic monotonicity properties of the family (0.2), which allow us to define the following limits

$$\overline{H}^{(q)} = \lim_{\varepsilon\to 0} \limsup_{k\to\infty} A^{(q)}(k,\varepsilon),$$
$$\underline{H}^{(q)} = \lim_{\varepsilon\to 0} \liminf_{k\to\infty} A^{(q)}(k,\varepsilon). \tag{0.5}$$

The brief exposition in the theory of entropy is given in Section 2. Theorems 2.7 (A.Katok) and 2.8 (A.Katok & M.Brin) are of extreme importance for us and will be used very often.

In Section 3 we prove the basic estimates

$$0 \leqslant \underline{H}^{(q)} \leqslant \overline{H}^{(q)} \leqslant h_{top}(f) \text{ for } 0 \leqslant q < 1,$$
$$0 \leqslant \underline{H}^{(q)} \leqslant \overline{H}^{(q)} \leqslant h_{\mu}(f) \text{ for } q > 1, \text{ providing the ergodicity of } (X,\mu,f).$$

In the next sections we study some particular cases where an explicit calculations are possible. For the homogeneous measures (Section 4)

$$\underline{H}^{(q)} = \overline{H}^{(q)} = h_{top}(f).$$

In Sections 5 and 6 we study a symbolic dynamical system (X, σ), where X is the set of all infinite sequences in the finite alphabet Ω and σ is a left shift. For the Bernoulli measure $\mu = \mu(p)$ generated by a probability vector $p = (p_1, \cdots, p_n)$ we have

$$\underline{H}^{(q)} = \overline{H}^{(q)} = -\frac{1}{q-1}\log\Big(\sum_{i=1}^{m} p_i^q\Big).$$

In the case of Gibbs measure $\mu = \mu_\phi$ (Section 6) corresponding to the function ϕ the answer is given in terms of the topological pressure

$$\underline{H}^{(q)} = \overline{H}^{(q)} = \frac{qP(\phi) - P(q\phi)}{q-1}.$$

The conjectures (C1) and (C2) are full filed in these examples.

In section 7 we improve our lower estimate for $0 \leqslant q < 1$. For an ergodic dynamical system (X, μ, f)

$$\underline{H}^{(q)} \geqslant h_{\mu}(f).$$

In Section 8 we prove that our families $\{\underline{H}^{(q)}\}$ and $\{\overline{H}^{(q)}\}$ are continuous for $0 < q < 1$ and $q > 1$. We show that for the spectrums whcih are continuous at the critical point $q = 1$ the conjecture (C1) is fulfiled.

We discuss the phenomenon of a phase transition in Section 9. The duality between the spectrums of generalized dimensions and entropies is demonstrated on the example of an expanding maps. Due to the examples of the pahse transition the conjecture (C1) does not hold for all the chaotic dynamical systems. Nevertheless, a weaker version of (C1) seems to be true

(C1′) $h = \lim_{q \to 1-0} H(q)$.

In order to analyse the behavior for $0 \leqslant q < 1$ we obtain a new estimate for the spectrum

$$A^{(q)}(k,\varepsilon) \leqslant \frac{1}{k} H^{(q)}_{R,\mu}(\xi^{(k)}) \leqslant -\frac{1}{k} \frac{\sum_i \mu^q(\Delta_i) \log \mu(\Delta_i)}{\sum_i \mu^q(\Delta_i)} := \frac{1}{k} H^{(q)}_{K,\mu}(\xi^k)$$

for any partition ξ with $\operatorname{diam}(\xi) < \varepsilon$. The last expression is known as the Kapur entropy in the information theory, see [AscDar-75]. We suppose that this quantity will be easier to analyse in relation with the conjecture (C1′). In general, it can be another source of generalized entropies in the theory of dynamical systems.

Acknowledgments. The problem, discussed in this report, has been investigated over few months under the supervision of prof. F.Takens. The author is grateful to prof. F.Takens for propounding an interesting problem, giving his time and attention to regular discussions and his valuable advices. The author would like to thank prof. B.M.Gurevich (The Moscow State University) for the fruitful discussion of the problem. The author gratefully acknowledges MRI and RUG for their financial support and hospitality.

1. Definition and basic properties.

Let (X, d) be a compact metric space and $f : X \to X$ be a continious mapping preserving a Borel probability measure μ. For any $k \in \mathbb{N}$ and any $\varepsilon > 0$ we define

$$\mathcal{B}_k(x,\varepsilon) = \left\{ y \in X : \ d(f^i x, f^i y) < \varepsilon \ \text{ for all } \ i = 0, \dots, k-1 \right\}.$$

For any $q \geqslant 0, q \neq 1$, any $\varepsilon > 0$ and $k \in \mathbb{N}$ define

$$A^{(q)}(k,\varepsilon) = -\frac{1}{(q-1)k} \log \int_X \mu \left\{ \mathcal{B}_k(x,\varepsilon) \right\}^{q-1} \, d\mu. \tag{1.1}$$

Furthermore we assume an integration without limits as an integration over the space X. Our main goal is to study an asymptotic behavior of the family (1.1) as $k \to \infty$ and $\varepsilon \to 0$. We start with the following simple observation

Proposition 1.1. *The family (1.1) has the following monotonicity properties*

(1) $A^{(q)}(k, \varepsilon_1) \geqslant A^{(q)}(k, \varepsilon_2)$ *for any* $0 < \varepsilon_1 < \varepsilon_2$;

(2) $A^{(q_1)}(k, \varepsilon) \geqslant A^{(q_2)}(k, \varepsilon)$ *for any* $q_1 < q_2$, $q_1, q_2 \neq 1$.

Proof: (1) By definition of $\mathcal{B}_k(x,\varepsilon)$ for any $0 < \varepsilon_1 < \varepsilon_2$ we have $\mathcal{B}_k(x,\varepsilon_1) \subset \mathcal{B}_k(x,\varepsilon_2)$. Hence, $\mu\{\mathcal{B}_k(x,\varepsilon_1)\} \leqslant \mu\{\mathcal{B}_k(x,\varepsilon_2)\}$. To prove (1) we consider cases $0 \leqslant q < 1$ and $q > 1$ separately.

For $q > 1$, $\mu\{\mathcal{B}_k(x,\varepsilon_1)\}^{q-1} \leqslant \mu\{\mathcal{B}_k(x,\varepsilon_2)\}^{q-1}$. Integration preserves this inequality. Since the logarithm is a monotonic function and $k \in \mathbb{N}$

$$\frac{1}{k}\log \int \mu\{\mathcal{B}_k(x,\varepsilon_1\}^{q-1}\, d\mu \leqslant \frac{1}{k}\log \int \mu\{\mathcal{B}_k(x,\varepsilon_2\}^{q-1}\, d\mu.$$

Finaly, since $q > 1$ we have $A^{(q)}(k,\varepsilon_1) \geqslant A^{(q)}(k,\varepsilon_2)$.

For $0 \leqslant q < 1$, $\mu\{\mathcal{B}_k(x,\varepsilon_1)\}^{q-1} \geqslant \mu\{\mathcal{B}_k(x,\varepsilon_2)\}^{q-1}$. Integration preserves this inequality. Since the logarithm is a monotonic function and $k \in \mathbb{N}$

$$\frac{1}{k}\log \int \mu\{\mathcal{B}_k(x,\varepsilon_1\}^{q-1}\, d\mu \geqslant \frac{1}{k}\log \int \mu\{\mathcal{B}_k(x,\varepsilon_2\}^{q-1}\, d\mu.$$

And finaly, since $q < 1$ we have $A^{(q)}(k,\varepsilon_1) \geqslant A^{(q)}(k,\varepsilon_2)$.

(2) We shall use the Lyapunov inequality [Shir84]. For any function $\xi : X \to \mathbb{R}$ and any $0 < s < t$

$$\left(\int |\xi|^s d\mu\right)^{\frac{1}{s}} \leqslant \left(\int |\xi|^t d\mu\right)^{\frac{1}{t}}.$$

Then (2) becomes obvious for the cases $0 \leqslant q_1 < q_2 < 1$ and $1 < q_1 < q_2$. To finish the proof we have to show (2) in the case $q_1 < 1 < q_2$. Let $\delta := \min\{1 - q_1, q_2 - 1\}$. Then for $q' = 1 - \delta$ and $q'' = 1 + \delta$

$$A^{(q')}(k,\varepsilon) - A^{(q'')}(k,\varepsilon) = \frac{1}{\delta k}\log\left[\int \mu\{\mathcal{B}_k(x,\varepsilon)\}^{-\delta}\, d\mu \times \int \mu\{\mathcal{B}_k(x,\varepsilon)\}^{\delta}\, d\mu\right]$$

$$\geqslant \frac{1}{\delta k}\log\left[\int \mu\{\mathcal{B}_k(x,\varepsilon)\}^{-\delta/2}\mu\{\mathcal{B}_k(x,\varepsilon)\}^{\delta/2}\, d\mu\right]^2 = 0. \tag{1.2}$$

by the Cauchy–Schwartz–Bunyakowskii inequality. Hence $A^{(q')}(k,\varepsilon) \geqslant A^{(q'')}(k,\varepsilon)$. Finally, by the choice of δ we have $q_1 \leqslant q' < 1$ and $1 < q'' \leqslant q_2$. We have already shown (2) in the cases where q_1 and q_2 both lie in $[0,1)$ or $(1,\infty)$. Hence

$$A^{(q_1)}(k,\varepsilon) \geqslant A^{(q')}(k,\varepsilon) \geqslant A^{(q'')}(k,\varepsilon) \geqslant A^{(q_2)}(k,\varepsilon).$$

This finishes the proof. ∎

These monotonicity properties allow us to formulate the folowing statement concerning the limits on k and ε.

PROPOSITION 1.2. *For any $q \geqslant 0$, $q \neq 1$, there exist limits*

$$\overline{H}^{(q)} := \lim_{\varepsilon\to 0}\limsup_{k\to\infty} A^{(q)}(k,\varepsilon), \quad \underline{H}^{(q)} := \lim_{\varepsilon\to 0}\liminf_{k\to\infty} A^{(q)}(k,\varepsilon).$$

And for every $q_1 < q_2$

$$\overline{H}^{(q_1)} \geqslant \overline{H}^{(q_2)}, \quad \underline{H}^{(q_1)} \geqslant \underline{H}^{(q_2)}.$$

Analysis of the same type as in Propostion 1.1 shows that $\underline{H}^{(q)}$ is non-negative for every q.

If for some $q \geqslant 0$, $q \neq 1$, $\underline{H}^{(q)} = \overline{H}^{(q)}$, then we say that there exists $H^{(q)} := \underline{H}^{(q)} = \overline{H}^{(q)}$.

2. Entropy.

In this section we give definitions and state without proofs of all the results about the entropies which we shall use later. We start with the definition of a measure–theoretic (or Kolmogorov–Sinai) entropy [Sin95]. We give two different equivalent definitions of the topological entropy, formulate the variational principle and present two important theorems due to Katok [Kat80] and Brin, Katok [Br&Ka81].

Let $(X, \mathcal{F}, \mu)$ be a probability space, and $f : X \to X$ a measure–preserving map, i.e.

$$\mu\{A\} = \mu\{f^{-1}A\} \text{ for any } A \in \mathcal{F}.$$

The dynamical system $(X, \mathcal{F}, \mu, f)$ is called ergodic if

$$\forall A \in \mathcal{F} : \mu\{A \vartriangle f^{-1}A\} = 0 \Rightarrow \mu\{A\} = 0 \text{ or } \mu\{A\} = 1.$$

Let ξ be a finite or countable partition of X into subsets $\{\Delta_i\}$.

DEFINITION 2.1. The entropy of the parttion ξ is defined as the value $H(\xi) = -\sum_i \mu\{\Delta_i\} \log \mu\{\Delta_i\}$. For all other partitions $H(\xi) = \infty$.

Denote by $\xi^{(k)}$ the partition of X into all non-empty sets of the form

$$\Delta_{i_0,\dots,i_{k-1}} := \Delta_{i_0} \cap f^{-1}\Delta_{i_1} \cap \dots \cap f^{-k+1}\Delta_{i_{k-1}}.$$

The partition $\xi^{(k)}$ is called the k-th iterate of ξ. One can show that for any countable partition ξ with $H(\xi) < \infty$, the limit

$$h_\mu(f, \xi) := \lim_{k\to\infty} \frac{1}{k} H(\xi^{(k)})$$

exists. This limit is called the measure–theoretic entropy of f with respect to ξ.

DEFINITION 2.2. The measure–theoretic entropy is defined as

$$h_\mu(f) := \sup_{\xi:\, H(\xi)<\infty} h_\mu(f, \xi).$$

The following theorem by Shannon, McMillan and Breiman [Sin95] is an extremely useful tool in proving the convergence to the entropy.
[Sin95].

THEOREM 2.3. *For any $x \in X$, denote by $\xi^{(k)}(x)$ the unique element of the partition $\xi^{(k)}$ containing x. Assume $(X, \mathcal{F}, \mu, f)$ is an ergodic dynamical system. Then for μ-a.s. x*

$$\lim_{k\to\infty} -\frac{1}{k} \log \mu\{\xi^{(k)}(x)\} = h_\mu(f, \xi).$$

The statement also shows that for most elements of $\xi^{(k)}$ their measures are in some weak sence the same. Using the fact that convergence almost sure implies the convergence in probability, the previous statement can be rewritten in the following form

COROLLARY 2.4. **(Theorem on asymptotic uniform distribution.)** *Assume $(X, \mathcal{F}, \mu, f)$ is an ergodic dynamical system. Then for any partition ξ and for each $\delta > 0$ there exists $K = K(\xi, \delta)$ such that for any $k > K$ one can choose a $\xi^{(k)}$–measurable set G_k such that*

(1) $\mu\{G_k\} \geqslant 1 - \delta$,
(2) *If $x \in G_k$ then* $\exp\big(-(h_\mu(f) + \delta)k\big) \leqslant \mu\{\xi^{(k)}(x)\} \leqslant exp\big(-(h_\mu(f) - \delta)k\big)$.

We give two equivalent definitions of the topological entropy. Both of them will be used later. The first one is due to Bowen. The second one was given by Adler, Konheim and McAndrew.

Consider the compact metric space (X, d) with a map $f : X \to X$. We say that a subset $S \in X$ is a (k, ε)–generator if for every $x \in X$ there exists $y \in S$ such that

$$d_k(x, y) := \max_{i=0,\dots,k-1} d(f^i x, f^i y) < \varepsilon.$$

Let $N(k, \varepsilon)$ be the least number of points in a (k, ε)–generator. Then there exists the limit

$$h_{top}(f) := \lim_{\varepsilon \to 0} \limsup_{k \to \infty} \frac{1}{k} \log N(k, \varepsilon).$$

One can show that

$$h_{top}(f) = \lim_{\varepsilon \to 0} \liminf_{k \to \infty} \frac{1}{k} \log N(k, \varepsilon).$$

We say that a set E is (k, ε)–separated if for very $x, y \in E$ there exists $0 \leqslant i < k$ such that $d(f^i x, f^i y) > \varepsilon$. Let $S(k, \varepsilon)$ be the maximal cardinality of a (k, ε)–separated set. Then one can show

$$\begin{aligned} h_{top}(f) &= \lim_{\varepsilon \to 0} \limsup_{k \to \infty} \frac{1}{k} \log S(k, \varepsilon) \\ &= \lim_{\varepsilon \to 0} \liminf_{k \to \infty} \frac{1}{k} \log S(k, \varepsilon). \end{aligned}$$

Let $\mathcal{U}$ be an open cover of X. We denote by $r(\mathcal{U})$ the smallest number of elements of $\mathcal{U}$ necessary to cover X. If $\mathcal{U}_1$ and $\mathcal{U}_2$ are open covers, we denote by $\mathcal{U}_1 \vee \mathcal{U}_2$ the cover formed by open sets $U \cap V$ with $U \in \mathcal{U}_1$ and $V \in \mathcal{U}_2$. Then one can show that the following limit exists

$$h(f, \mathcal{U}) = \lim_{k \to \infty} \frac{1}{k} \log r \left(\bigvee_{i=0}^{k-1} f^{-i} \mathcal{U}. \right)$$

This limit $h(f, \mathcal{U})$ is called the entropy of f with respect to $\mathcal{U}$. One can show that

$$h_{top}(f) = \sup_{\mathcal{U}} h(f, \mathcal{U}).$$

The basic relation between topological and metric entropies can be formulated as follows.

PROPOITION 2.5. *(Variatinal principle). Let X be a compact metric space and $f : X \to X$ a contionuous map. Then the toplogical entropy of f satisfies*

$$h_{top}(f) = \sup_{\mu \in \mathcal{M}_f(X)} h_\mu(f),$$

where $\mathcal{M}_f(X)$ is the set of all f-invariant Borel probability measures on X.

This fact leads to the following definiton.

DEFINITION 2.6. If $h_{top}(f) = h_\mu(f)$ for some $\mu \in \mathcal{M}_f(X)$, then μ is called a measure with a maximal entropy.

A.Katok in [Kat80] showed that we can define the metric entropy in terms similar to the definiton of the topological entropy. Namely,

THEOREM 2.7. *Let X be a compact metric space and $f : X \to X$ a contionuous map. Assume μ is an ergodic f-invariant probability measure on X. Then for every $\delta > 0$*

$$h_\mu(f) = \lim_{\varepsilon \to 0} \limsup_{k \to \infty} \frac{1}{k} \log N(k, \varepsilon, \delta),$$

where $N(k, \varepsilon, \delta)$ is the least number of ε-balls in the d_k-metric which cover the set of measure more than or equal to $1 - \delta$.

The following theorem by M.Brin and A.Kator [Br&Ka81] may be considered as a topological version of the Shannon–McMillan–Breiman theorem.

THEOREM 2.8. *Let X be a compact metric space and $f : X \to X$ be a continuous map preserving a non-atomic Borel probability measure μ. Assume that $h_\mu(f) < \infty$. Then for μ-a.s. $x \in X$*

(1) $\lim_{\varepsilon \to 0} \limsup_{k \to \infty} -\frac{1}{k} \log \mu\{\mathcal{B}_k(x, \varepsilon)\} = \lim_{\varepsilon \to 0} \liminf_{k \to \infty} -\frac{1}{k} \log \mu\{\mathcal{B}_k(x, \varepsilon)\} = h_\mu(f, x)$;
(2) *$h_\mu(f, x)$ is f-invariant;*
(3) $\int_X h_\mu(f, x) = h_\mu(f)$.

For the ergodic dynamical system (X, μ, f) every invariant function is a constant almost sure. Therefore we have

COROLLARY 2.8. *If (X, μ, f) is an ergodic dynamical system then $h_\mu(f, x) = h_\mu(f)$ for μ-almost every x.*

3. Upper estimates.

From this moment we assume that our dynamical system (X, f) has finite topological entropy. In this section we give basic estimates in the terms of the topological and measure–theoretic entropies. We separate cases $q < 1$ and $q > 1$ again.

PROPOSITION 3.1. *Assume that (X, μ, f) is ergodic and let $q > 1$. Then for any $\delta > 0$ and for any finite partition ξ of X such that*

$$\operatorname{diam}(\xi) := \max_{\Delta \in \xi} \operatorname{diam}(\Delta) < \varepsilon \tag{3.1}$$

there exists $K = K(\xi, \delta) > 0$ *such that for any integer* $k > K$

$$A^{(q)}(k, \varepsilon) \leqslant h_\mu(f, \xi) + 2\delta,$$

Proof: Take an arbitrary $\delta > 0$ and let ξ be a finite partition satisfying (3.1). Then for every x $\xi(x) \subset \mathcal{B}(x, \varepsilon)$. Hence

$$\xi^{(k)}(x) = \bigcap_{i=0}^{k-1} f^{-i}\xi(f^i x) \subset \bigcap_{i=0}^{k-1} f^{-i}\mathcal{B}(f^i x, \varepsilon) = \mathcal{B}_k(x, \varepsilon).$$

Therefore we can estimate

$$\begin{aligned} A^{(q)}(k, \varepsilon) &= -\frac{1}{(q-1)k} \log \int \mu\left\{\mathcal{B}_k(x, \varepsilon)\right\}^{q-1} d\mu \\ &\leqslant -\frac{1}{(q-1)k} \log \int \mu\left\{\xi^{(k)}(x)\right\}^{q-1} d\mu \\ &= -\frac{1}{(q-1)k} \log \sum_{\Delta \in \xi^{(k)}} \mu\left\{\Delta\right\}^q. \end{aligned}$$

Choose $K = K(\xi, \delta)$ as in the Corollary 2.4 on asymptotic uniform distribution. Then for every $k > K(\xi, \delta)$ there is a set $G_k \in \xi^{(k)}$ with the correspondent properties and we can continue our estimate as follows

$$\begin{aligned} A^{(q)}(k, \varepsilon) &\leqslant -\frac{1}{(q-1)k} \log \sum_{\Delta \in G_k} \mu\left\{\Delta\right\}^q \\ &\leqslant -\frac{1}{(q-1)k} \log \left(\exp\{-(h_\mu(f) + \delta)(q-1)k\} \sum_{\Delta \in G_k} \mu\left\{\Delta\right\} \right) \\ &\leqslant h_\mu(f) + \delta - \frac{1}{(q-1)k} \log(1-\delta) \leqslant h_\mu(f) + 2\delta. \end{aligned}$$

for suffiently large k. ■

Because of the arbitrary choice of $\delta > 0$ in Proposition 3.1 we have immediately

CORROLARY 3.2. *For an ergodic dynamical system* (X, μ, f) *and any* $q > 1$

$$\overline{H}^{(q)} \leqslant h_\mu(f).$$

Now we consider the case $0 \leqslant q < 1$.

PROPOSITION 3.3. *For any dynamical system* (X, μ, f) *and any* $0 \leqslant q < 1$

$$\overline{H}^{(q)} \leqslant h_{top}(f).$$

Proof: Again take an arbitrary finite partion ξ of X with a diameter less then ε. As in Proposition 3.1 we have

$$\begin{aligned} A^{(q)}(k,\varepsilon) &= \frac{1}{(1-q)k}\log\int\frac{1}{\mu\left\{\mathcal{B}_k(x,\varepsilon)\right\}^{1-q}}\,d\mu \\ &\leqslant \frac{1}{(1-q)k}\log\int\frac{1}{\mu\{\xi^{(k)}(x)\}^{1-q}}\,d\mu \\ &= \frac{1}{(1-q)k}\log\sum_{\Delta\in\xi^{(k)}}\mu\left\{\Delta\right\}^q \\ &\leqslant \frac{1}{(1-q)k}\log\operatorname{card}(\xi^{(k)}). \end{aligned}$$

As we have seen in Section 2 $h_{top}(f) = \sup_{\mathcal{U}}\lim_{k\to\infty}\frac{\log \mathrm{r}(\mathcal{U}^{(k)})}{k}$, where supremum is taken over all finite open covers of X. Let us take any finite open cover $\mathcal{U}$ such that for any $\Delta \in \xi$ there is $U \in \mathcal{U}$: $\Delta \subset U$.[1] Then $\mathrm{r}(\xi^{(k)}) \leqslant \mathrm{r}(\mathcal{U}^{(k)})$ and hence

$$\overline{H}^{(q)} \leqslant \frac{1}{1-q}h_{top}(f).$$

By the monotonicity properties from Proposition 1.2 we have for every $0 \leqslant q < 1$

$$\overline{H}^{(q)} \leqslant \overline{H}^{(0)} \leqslant h_{top}(f).$$

This completes the proof. ■

4. Homogeneous measures

In this section we give an example where our family may be computed explicitly. We restrict ourselves to the case of homogeneous measures. We give the definition in the case of a compact metric space. For a general definition see [Ward94].

DEFINITION 4.1. Let f be a continuous mapping on the compact metric space (X, d). A Borel measure μ on X is said to be f-homogeneous if for each $\varepsilon > 0$ there exist $\delta > 0$ and $c > 0$ such that

$$\mu\{\mathcal{B}_k(y,\delta)\} \leqslant c\mu\{\mathcal{B}_k(x,\varepsilon)\}$$

for all $k > 0$ and $x, y \in X$, where $\mathcal{B}_k(x,\varepsilon) = \{z : \max_{i=0,\dots,k-1} d(f^i x, f^i z) < \varepsilon\}$.

This condition is rather strong. But on the other hand, these measures have a lot of good properties. For example, the Theorem 7.6 from [Ward94] in the case of a compact space takes the form

[1] For example, if for any $\Delta \in \xi$ we take its open ε-neighbourhood and consider all these sets as an open cover $\mathcal{U}$, then $\mathcal{U}$ obviously satisfy this condition.

THEOREM 4.2. *Let μ be a f-homogeneous measure. For every $x \in X$ define*

$$h_*(\mu, f) = \lim_{\varepsilon \to 0} \limsup_{k \to \infty} -\frac{1}{k} \log \mu\{\mathcal{B}_k(x, \varepsilon)\}.$$

Then

(1) *This definition does not depend on the point $x \in X$.*
(2) $h_{top}(f) = h_*(\mu, f)$.
(3) *If μ is a f-invariant probability measure then $h_\mu(f) = h_*(\mu, f)$.*

Remark. Originaly in [Ward-94], there was an additional assumption of finiteness of covering dimension in (3), but it can be omitted using Theorem 2.8, where X is assumed to be a compact space only. Indeed, let X be a compact space. Then by (2) we have that $h_*(\mu, f)$ is a constant on X. And from Theorem 2.8 one has

$$h_\mu(f) = \int h_*(\mu, f) d\mu = h_*(\mu, f).$$

The characteristic property of the homogeneous measures allows to get exact values of $\underline{H}^{(q)}$ and $\overline{H}^{(q)}$. Let us assume that μ is a f-homogeneous invariant Borel probability measure on X. Take an arbitrary $\varepsilon > 0$ and choose the corresponding $\delta > 0$ and $c > 0$ as in definiton 4.1. Consider $q > 1$. Then for every $0 < \gamma < \delta$ and any $x, y \in X$ we have

$$\mu\{\mathcal{B}_k(y, \gamma) \leqslant \mu\{\mathcal{B}_k(y, \delta)\} \leqslant c\mu\{\mathcal{B}_k(x, \varepsilon)\}.$$

Since the monotonicity properties from Proposition 1.1 we have

$$\begin{aligned} A^{(q)}(k, \gamma) \geqslant A^{(q)}(k, \delta) &= -\frac{1}{(q-1)k} \log \int \mu\{\mathcal{B}_k(y, \delta)\}^{q-1} \, d\mu(y) \\ &\geqslant -\frac{1}{(q-1)k} \log \left(c^{q-1} \mu\{\mathcal{B}_k(x, \varepsilon)\}^{q-1}\right) \\ &= -\frac{1}{k} \left(\log c + \log \mu\{\mathcal{B}_k(x, \varepsilon)\}\right). \end{aligned}$$

It is clear that $\delta \to 0$ as $\varepsilon \to 0$. Hence after taking the limits and applying the first part of Theorem 2.7 we have

$$\begin{aligned} \underline{H}^{(q)} &\geqslant \lim_{\varepsilon \to 0} \liminf_{k \to \infty} -\frac{1}{k} \left(\log c + \log \mu\{\mathcal{B}_k(x, \varepsilon)\}\right) \\ &= \lim_{\varepsilon \to 0} \liminf_{k \to \infty} -\frac{1}{k} \log \mu\{\mathcal{B}_k(x, \varepsilon)\} \\ &= \lim_{\varepsilon \to 0} \limsup_{k \to \infty} -\frac{1}{k} \log \mu\{\mathcal{B}_k(x, \varepsilon)\} = h_*(\mu, f) \end{aligned}$$

For any $q \geqslant 0$, $q \neq 1$, we have shown that $\overline{H}^{(q)} \leqslant h_{top}(f)$. Since μ is a homogeneous measure and for these measures $h_{top}(f) = h_*(\mu, f)$ one can conclude that

$$h_{top}(f) = h_*(\mu, f) \geqslant \overline{H}^{(q)} \geqslant \underline{H}^{(q)} \geqslant h_*(\mu, f).$$

It means that for homogeneous invariant measures and $q > 1$ we have

$$H^{(q)} := \overline{H}^{(q)} = \underline{H}^{(q)} = h_*(\mu, f) = h_{top}(f) = h_\mu(f).$$

We proceed in a completely similar way for $0 \leqslant q < 1$. As in the previous case let $\varepsilon > 0$, take $\delta > 0$ and $c > 0$ from the definition of a homogeneous measure. Then for every $0 < \gamma < \delta$

$$\begin{aligned} A^{(q)}(k,\gamma) \geqslant A^{(q)}(k,\delta) &= \frac{1}{(1-q)k} \log \int \frac{1}{\mu\{\mathcal{B}_k(y,\delta)\}^{1-q}} \, d\mu(y) \\ &\geqslant \frac{1}{(1-q)k} \log \frac{1}{c^{1-q}\mu\{\mathcal{B}_k(x,\varepsilon)\}^{1-q}} \\ &= -\frac{1}{k}\left(\log c + \log \mu\{\mathcal{B}_k(x,\varepsilon)\}\right). \end{aligned}$$

Because of the same reasons we have

$$h_{top}(f) = h_*(\mu, f) \geqslant \overline{H}^{(q)} \geqslant \underline{H}^{(q)} \geqslant h_*(\mu, f)$$

for $0 \leqslant q < 1$.

Therefore we proved:

THEOREM 4.3. *Let μ be a f-homogeneous invariant Borel probability measure on X. Then for any $q \geqslant 0$, $q \neq 1$*

$$H^{(q)} := \overline{H}^{(q)} = \underline{H}^{(q)} = h_*(\mu, f) = h_{top}(f) = h_\mu(f).$$

And we can define

$$H^{(1)} = \lim_{q \to 1} H^{(q)}.$$

5. Symbolic dynamics.

Let $\Omega = \{1, \dots, m\}$ be a finite alphabet. Let $X = \{\mathbf{x} = \{x_i\}_{i=-\infty}^{\infty} : \ x_i \in \Omega\}$ and σ be a left shift:

$$\sigma(\mathbf{x})_i = x_{i+1}.$$

We define a metric d on X as follows

$$d(\mathbf{x}, \mathbf{y}) = 2^{-N},$$

where $N = \max\{n \in \mathbb{N} : \ x_i = y_i \ \forall |i| < n\}$. The triple (X, μ, σ) is called a symbolic dynamical system.

For any s, t, $s \leqslant t$, and any set $\{a_s, \dots, a_t\}$, $a_i \in \Omega$, we define a cylinder

$$C_s^t(a_s, \dots, a_t) = \{\mathbf{x} \in X : \ x_i = a_i \text{ for } i = s, \dots, t\}.$$

Let now $\mathbf{x}$ be an arbitrary point and $\varepsilon = \frac{1}{2^{n+1}}$. If $\mathbf{y} \in \mathcal{B}_k(\mathbf{x}, 1/2^{n+1})$ then by the definition of d and d_k one has

$$d(\mathbf{x},\mathbf{y}) < \frac{1}{2^{n+1}} \quad \Rightarrow \quad x_i = y_i \text{ for all } i = -n, \dots, n$$
$$d(\sigma\mathbf{x},\sigma\mathbf{y}) < \frac{1}{2^{n+1}} \quad \Rightarrow \quad x_i = y_i \text{ for all } i = -n+1, \dots, n+1$$
$$\dots$$
$$d(\sigma^{k-1}\mathbf{x},\sigma^{k-1}\mathbf{y}) < \frac{1}{2^{n+1}} \quad \Rightarrow \quad x_i = y_i \text{ for all } i = -n+k-1, \dots, n+k-1$$

Combining all together one has

$$\mathbf{y} \in \mathcal{B}_k(\mathbf{x}, 1/2^{n+1}) \quad \Rightarrow \quad x_i = y_i \text{ for all } i = -n, \dots, n+k-1.$$

Using the notion of a cylindric set we can write

$$\mathcal{B}_k(\mathbf{x}, 1/2^{n+1}) \subset C_{-n}^{n+k-1}(x_{-n}, \dots, x_{n+k-1}). \tag{5.1}$$

Because of the similar reasonings we have

$$\mathcal{B}_k(\mathbf{x}, 1/2^{n+1}) \supset C_{-n-1}^{n+k-1}(x_{-n-1}, \dots, x_{n+k-1}). \tag{5.2}$$

We denote $I(k,n) = \int \mu\{\mathcal{B}_k(\mathbf{x}, 1/2^{n+1})\}^{q-1}\, d\mu$. Let $\sum_{C\in(s,t)}$ denotes the sum over all possible cylinders C starting at s and ending at t. The situation in the case of a symbolic dynamics becomes simple because the integral $I(k,n)$ can be etimated through the finite sums. Namely, for every s, t

$$\begin{aligned} I(k,n) &= \int \mu\{B_k(\mathbf{x}, 1/2^{n+1})\}^{q-1}\, d\mu \\ &= \sum_{C\in(s,t)} \int_C \mu\{\mathcal{B}_k(\mathbf{x}, 1/2^{n+1})\}^{q-1}\, d\mu. \end{aligned} \tag{5.3}$$

Now setting $t = n+k-1$, $s = -n$ and $s = -n-1$ and taking into account our approximation (5.1), (5.2) we have

$$0 \leqslant q < 1: \qquad \sum_{C'\in(-n,n+k-1)} \mu\{C'\}^q \leqslant I(k,n) \leqslant \sum_{C''\in(-n-1,n+k-1)} \mu\{C''\}^q, \tag{5.4}$$

$$q > 1: \qquad \sum_{C''\in(-n-1,n+k-1)} \mu\{C''\}^q \leqslant I(k,n) \leqslant \sum_{C'\in(-n,n+k-1)} \mu\{C'\}^q. \tag{5.5}$$

In the symbolic case we can calculate quantities $\overline{H}^{(q)}, \underline{H}^{(q)}$ at $q = 0$ precisely.

PROPOSITION 5.1. *For a dynamical system (X, μ, σ), where X is the space of all infinite sequences over a finite alphabet Ω,σ is a left shift and μ is an $\sigma-$ invariant measure, positive on open sets (or if $spt(\mu) = X$),*

$$\overline{H}^{(0)} = \underline{H}^{(0)} = \log m = h_{top}(\sigma),$$

where $m = card(\Omega)$.

Proof: For $q = 0$ the inequality (5.4) takes form

$$m^{2n+k} \leqslant I(k,n) \leqslant m^{2n+k+1}.$$

Our estimates are just the numbers of cylinders of lenghth $2n + k$ and $2n + k + 1$ in an alphabet of size m. Hence our limit quantities are

$$\overline{H}^{(q)} = \lim_{n\to\infty} \limsup_{k\to\infty} \frac{1}{k} \log I(k,n) \leqslant \lim_{n\to\infty} \limsup_{k\to\infty} \frac{2n+k+1}{k} \log m = \log m,$$

$$\underline{H}^{(q)} = \lim_{n\to\infty} \liminf_{k\to\infty} \frac{1}{k} \log I(k,n) \geqslant \lim_{n\to\infty} \liminf_{k\to\infty} \frac{2n+k}{k} \log m = \log m.$$

And using the first definition of a topological entropy it is not very difficult to show that $h_{top}(\sigma)$ is exactly $\log m$. ■

For studying the behavior of $\overline{H}^{(q)}$ and $\underline{H}^{(q)}$ at $q = 1$ we have to specify the measure μ. We give two examples where explicit calculations are possible.

Bernoulli shift.

Let $\mathbf{p} = \{p_1, \ldots, p_m\}$ be a probability vector, i.e $p_i \geqslant 0$ for any i and $\sum p_i = 1$. We define a measure $\rho = \rho(\mathbf{p})$ on Ω. By definition $\rho(\{i\}) = p_i$. And let $\mu = \mu(\mathbf{p})$ be a product measure on X. Then for any s, t, $s \leqslant t$, and any set $\{a_s, \ldots, a_t\}$, the measure of a corresponding cylinder is given by

$$\mu\left\{C_s^t(a_s, \ldots, a_t)\right\} = \prod_{i=s}^{t} p_{a_i}.$$

For this measure the Kolmogorov–Sinai entropy is given by formula $h_\mu(\sigma) = -(p_1 \log p_1 + \ldots + p_m \log p_m)$ [Bil]. Estimates (5.4), (5.5) take the form

(6.4a) $$0 \geqslant q < 1: \ (p_1^q + \ldots + p_m^q)^{2n+k} \leqslant I(k,n) \leqslant (p_1^q + \ldots + p_m^q)^{2n+k+1},$$

(6.5a) $$q > 1: \ (p_1^q + \ldots + p_m^q)^{2n+k+1} \leqslant I(k,n) \leqslant (p_1^q + \ldots + p_m^q)^{2n+k}.$$

Taking all the limits we have

$$\overline{H}^{(q)} = \underline{H}^{(q)} = -\frac{1}{q-1} \log(p_1^q + \ldots + p_m^q) =: H^{(q)}.$$

There exists a limit as $q \to 1$

$$\begin{aligned} H^{(1)} := \lim_{q\to 1} H^{(q)} &= \lim_{q\to 1} -\frac{1}{q-1} \log(p_1^q + \ldots + p_m^q) \\ &= \lim_{q\to 1} -\frac{p_1^q \log p_1 + \ldots + p_m^q \log p_m}{p_1^q + \ldots + p_m^q} \\ &= -(p_1 \log p_1 + \ldots + p_m \log p_m) = h_\mu(\sigma) \end{aligned}$$

6. Gibbs measures and Thermodynamical Formalism.

In this section we study a family of Gibbs invariant measures for the symbolic dynamical system (X, σ). In our presentation we follow [Bow].

Let $\phi : X \to \mathbb{R}$ be a continuous function. We define

$$var_k(\phi) := \sup\{|\phi(\mathbf{x}) - \phi(\mathbf{y})| : \ x_i = y_i \ \forall |i| < k\}.$$

The main result on the existence of Gibbs measures is

THEOREM 6.1. *Suppose that for* $\phi : X \to \mathbb{R}$ *there are constants* $c > 0$ *and* $\alpha \in (0,1)$ *such that* $var_k(\phi) \leqslant c\alpha^k$. *Then there is a unique invariant Borel probability measure* μ *for which one can find constants* $c_1, c_2 > 0$ *and* P *such that*

$$c_1 \leqslant \frac{\mu\{\mathbf{y} : \ y_i = x_i \ \forall i = 0, \dots, k-1\}}{\exp\{-Pk + \sum_{i=0}^{k-1} \phi(\sigma^i \mathbf{x})\}} \leqslant c_2. \tag{6.1}$$

for all $\mathbf{x} \in X$ *and* $k > 0$.

This measure μ is denoted by μ_ϕ and called the Gibbs measure corresponding to ϕ.

Let ϕ be a continuous function on X. Define

$$\sup\nolimits_{a_0, \dots, a_{k-1}} S_k(\phi) := \sup\Big\{\sum_{i=0}^{k-1} \phi(\sigma^i \mathbf{x}) : \ x_i = a_i \text{ for } 0 \leqslant i < k-1\Big\},$$

$$Z_k(\phi) := \sum_{a_0, \dots, a_{k-1}} \exp\{\sup\nolimits_{a_0, \dots, a_{k-1}} S_k(\phi)\}.$$

LEMMA 6.2. *For any continuous function* ϕ *there is a well defined limit*

$$P(\phi) = \lim_{k \to \infty} \frac{1}{k} \log Z_k(\phi).$$

Moreover, if ϕ *satisfies the condition of Theorem 6.1, then* $P(\phi)$ *equals to the constant* P *in the definition (6.1) of* μ_ϕ. *This value* $P(\phi)$ *is called the topological preasure of* ϕ.

PROPOSITION 6.3. *(Variatinal Principle) (1) Suppose,* ϕ *is a continuous function and* μ *is an invariant measure then*

$$h_\mu(\sigma) + \int \phi \, d\mu \leqslant P(\phi),$$

(2) The measure $\mu = \mu_\phi$ *from Theorem 6.1 is a unique invariant measure such that*

$$h_\mu(\sigma) + \int \phi \, d\mu = P(\phi).$$

This measure μ_ϕ *is called the equilibrium state.*

THEOREM 6.4. *Let (X,μ,σ) be a symbolic dynamical system, where $\mu = \mu_\phi$ is the equilibrium state for the function ϕ with an exponential decaying variation (6.1). Then for each $q \geqslant 0$, $q \neq 1$.*

$$\underline{H}^{(q)} = \overline{H}^{(q)} = \frac{P(q\phi) - qP(\phi)}{1-q}.$$

Proof: Since our measure μ is σ–invariant and it is an equlibrium state, we can continue estimates (5.4) and (5.5) as folows

$$\begin{aligned} 0 \leqslant q < 1: \quad I(k,n) &\geqslant \sum_{a_{-n},\dots,a_{n+k-1}} \mu\{C_{-n}^{n+k-1}(a_{-n},\dots,a_{n+k-1})\}^q \\ &= \sum_{a_{-n},\dots,a_{n+k-1}} \mu\{C_{0}^{2n+k-1}(a_{-n},\dots,a_{n+k-1})\}^q \\ &\geqslant c_1^q \exp\{-P(2n+k)q\} Z_{2n+k}(q\phi). \end{aligned}$$

Similary,

$$I(k,n) \leqslant c_2^q \exp\{-P(2n+k+1)q\} Z_{2n+k+1}(q\phi).$$

Hence, after taking all the limits one has

$$\underline{H}^{(q)} = \overline{H}^{(q)} = \frac{P(q\phi) - qP(\phi)}{1-q}.$$

The case $q > 1$ can be proved in a similar way. ■

7. Lower estimates

In this section we give the lower estimate of $\underline{H}^{(q)}$ for $0 \leqslant q < 1$. We use the notion of separated sets from Section 2 and results of Katok (Theorem 2.7) and Brin, Katok (Theorem 2.8).

PROPOSITION 7.1. *Let (X,μ,f) be an ergodic dynamical system. Then the following estimate holds for $0 \leqslant q < 1$*

$$\underline{H}^{(q)} \geqslant h_\mu(f).$$

Proof: Recall the definition from Section 2: a set $E \subset X$ is called (k,ε)–separated if for any $x, y \in X$ there exists i, $0 \leqslant i < k$, such that $d(f^i x, f^i y) > \varepsilon$.

One can make two simple observations.

(1) Let E be an arbitrary (k,ε)–separated set in X. Then for each $x, y \in E$, $x \neq y$, $\mathcal{B}_k(x,\varepsilon/2) \cap \mathcal{B}_k(y,\varepsilon/2) = \emptyset$.

(2) Let $y \in \mathcal{B}_k(x,\varepsilon/2)$ then $\mathcal{B}_k(y,\varepsilon/2) \subset \mathcal{B}_k(x,\varepsilon)$.

The proofs are simple. (1) Suppose $\mathcal{B}_k(x,\varepsilon/2) \cap \mathcal{B}_k(y,\varepsilon/2) \neq \emptyset$. Choose any z in this intersection. Since d_k is a metric on X, we have a triangle inequality $d_k(x,y) \leqslant d_k(x,z) + d_k(z,y) < \varepsilon/2 + \varepsilon/2 = \varepsilon$. We arrived at contradiction with the definition of a (k,ε)–separated set.

(2) Consider any $z \in \mathcal{B}_k(y, \varepsilon/2)$. As in (1) since d_k is a metric we have $d_k(x,z) \leqslant d_k(x,y) + d_k(y,z) < \varepsilon$. Hence $z \in \mathcal{B}_k(x,\varepsilon)$ and $\mathcal{B}_k(y,\varepsilon/2) \subset \mathcal{B}_k(x,\varepsilon)$.

Consider $0 \leqslant q < 1$ and any (k,ε)–separated set E in X. Then using the previous observations one can have

$$\begin{aligned}
A^{(q)}(k, \frac{\varepsilon}{2}) &= \frac{1}{(1-q)k} \log \int \frac{d\mu}{\mu\{\mathcal{B}_k(x, \frac{\varepsilon}{2}\}^{1-q}} \\
&\geqslant \frac{1}{(1-q)k} \log \sum_{x_j \in E} \int_{\mathcal{B}_k(x_j, \varepsilon/2)} \frac{d\mu}{\mu\{\mathcal{B}_k(x, \frac{\varepsilon}{2}\}^{1-q}} \\
&\geqslant \frac{1}{(1-q)k} \log \sum_{x_j \in E} \frac{\mu\{\mathcal{B}_k(x_j, \frac{\varepsilon}{2})\}}{\mu\{\mathcal{B}_k(x_j, \varepsilon)\}^{1-q}}.
\end{aligned} \tag{7.1}$$

Now we have to estimate the last expression form bellow. For this we intoduce

$$f^{(q)}(x,k,\varepsilon) := \frac{\mu\{\mathcal{B}_k(x, \frac{\varepsilon}{2})\}}{\mu\{\mathcal{B}_k(x, \varepsilon)\}^{1-q}}.$$

Assume that μ is an erodic invariant measure. Then using Theorem 2.8 one can show that

$$\lim_{\varepsilon \to 0} \limsup_{k \to \infty} \frac{\log f^{(q)}(x,k,\varepsilon)}{k} = \lim_{\varepsilon \to 0} \liminf_{k \to \infty} \frac{\log f^{(q)}(x,k,\varepsilon)}{k} = -q h_\mu(f) \tag{7.2}$$

for μ almost all $x \in X$. Indeed,

$$\begin{aligned}
\lim_{\varepsilon \to 0} \limsup_{k \to \infty} \frac{\log f^{(q)}(x,k,\varepsilon)}{k} &\leqslant \lim_{\varepsilon \to 0} \limsup_{k \to \infty} \frac{\log \mu\{\mathcal{B}_k(x, \varepsilon/2)\}}{k} \\
&+ \lim_{\varepsilon \to 0} \liminf_{k \to \infty} -(1-q) \frac{\log \mu\{\mathcal{B}_k(x,\varepsilon)\}}{k} \\
&= -h_\mu(f) + (1-q) h_\mu(f) = -q h_\mu(f).
\end{aligned}$$

And similary

$$\begin{aligned}
\lim_{\varepsilon \to 0} \liminf_{k \to \infty} \frac{\log f^{(q)}(x,k,\varepsilon)}{k} &\geqslant \lim_{\varepsilon \to 0} \liminf_{k \to \infty} \frac{\log \mu\{\mathcal{B}_k(x, \varepsilon/2)\}}{k} \\
&+ \lim_{\varepsilon \to 0} \limsup_{k \to \infty} -(1-q) \frac{\log \mu\{\mathcal{B}_k(x,\varepsilon)\}}{k} \\
&= -h_\mu(f) + (1-q) h_\mu(f) = -q h_\mu(f).
\end{aligned}$$

Combinig two previous inequalities we arrive at (7.2). We can delete a set of measure 0 where (7.2) does not converge from our consideration. So, we can assume that (7.2) holds for any $x \in X$. Take any $\delta > 0$. Then

(1) $\exists \varepsilon_0 = \varepsilon_0(x, \delta)$ such that for any $0 < \varepsilon < \varepsilon_0$

$$-q h_\mu(f) - \delta < \liminf_{k \to \infty} \frac{\log f^{(q)}(x,k,\varepsilon)}{k} \leqslant \limsup_{k \to \infty} \frac{\log f^{(q)}(x,k,\varepsilon)}{k} < -q h_\mu(f) + \delta.$$

(2) $\exists K = K(x, \varepsilon, \delta)$ such that for any $k > K$

$$-q h_\mu(f) - 2\delta < \frac{\log f^{(q)}(x,k,\varepsilon)}{k} < -q h_\mu(f) + 2\delta.$$

Consider the set $E(\varepsilon, K) := \{x : \varepsilon_0(x,\delta) > \varepsilon,\ K(x,\varepsilon,\delta) < K\}$. Since we have convergence for almost all $x \in X$

$$\lim_{\varepsilon\to 0}\lim_{K\to\infty} \mu\{E(\varepsilon, K)\} = 1.$$

Choose $\varepsilon > 0$ and K such that $\mu\{E(\varepsilon, K)\} > 1-\delta$. For every $k > K$ let S_k be the maximal (k,ε)–separated set in $E(\varepsilon, K)$. It means that $S_k \subset E(\varepsilon, K)$ and S_k is a (k,ε)–separated set. Continuing (7.1) for S_k we find

$$\begin{aligned} A^{(q)}(k,\frac{\varepsilon}{2}) &\geqslant \frac{1}{(1-q)k}\log\sum_{x_j\in S_k}\frac{\mu\{\mathcal{B}_k(x_j,\frac{\varepsilon}{2})\}}{\mu\{\mathcal{B}_k(x_j,\varepsilon)\}^{1-q}} \\ &\geqslant \frac{1}{(1-q)k}\log\Big(\mathrm{card}(S_k)\exp\{-q(h_\mu(f)+2\delta)k\}\Big) \\ &\geqslant \frac{1}{(1-q)k}\log\Big(N(\varepsilon,k,\delta)\exp\{-q(h_\mu(f)+2\delta)k\}\Big), \end{aligned} \tag{7.3}$$

where $N(\varepsilon, k, \delta)$ is the minimal number of ε–balls in the d_k metric which cover the set of measure at least $1-\delta$. We have to show that $N(\varepsilon,k,\delta) \leqslant \mathrm{card}(S_k)$. Indeed, S_k has been chosen as a maximal (k,ε)–separated set in $E(\varepsilon, k)$. It means that for every $x \in E(\varepsilon,k)$ there is $x_j \in S_k$ such that $d_k(x,x_j) < \varepsilon$. This shows that $E(\varepsilon,k) \subset \bigcup_{x_j\in S_k} \mathcal{B}_k(x_j,\varepsilon)$. By the way $E(\varepsilon,k)$ was constructed we have

$$\mu\{\bigcup_{x_j\in S_k}\mathcal{B}_k(x_j,\varepsilon)\} \geqslant \mu\{E(\varepsilon,k)\} \geqslant 1-\delta.$$

Taking the limits on the both sides of inequality (7.3) and using Theorem 2.8 we have

$$\underline{H}^{(q)} \geqslant \frac{1}{(1-q)}(h_\mu(f) - qh_\mu(f) - 2\delta) = h_\mu(f) - \frac{2}{1-q}\delta.$$

And since the arbitrary choice of $\delta > 0$ finaly for any $0 \leqslant q < 1$ we have

$$\underline{H}^{(q)} \geqslant h_\mu(f).$$

■

Remark. We can improve our estimate at $q = 0$ in some cases. At $q = 0$ the limit in (7.2) equals to 0. And suppose

$$f^{(0)}(x,k,\varepsilon) = \frac{\mu\{\mathcal{B}_k(x,\varepsilon/2)\}}{\mu\{\mathcal{B}_k(x,\varepsilon)\}} \geqslant g(\varepsilon) > 0$$

for some function g, for all $k > K$, for all sufficiently small $\varepsilon > 0$ and all $x \in X$. We can conclude from (7.1) that

$$A^{(q)}(k,\frac{\varepsilon}{2}) \geqslant \frac{1}{k}\log(card(E)g(\varepsilon)),$$

where E is any (k,ε)–separated set. Let E be the maximal (k,ε)–separated set. Then taking the limits and using the definition of the topological entropy we have $\underline{H}^{(0)} \geqslant h_{top}(f)$. Hence one can define $H^{(0)}$

$$H^{(0)} := \underline{H}^{(0)} = \overline{H}^{(0)} = h_{top}(f).$$

Remark. Another way to prove the estimate consists of applying the Iensen inequality, Fatu's lemma and Theorem 2.8. This immediately gives a required estimate. However, one loses the control over convergence at $q = 0$ as it was descibed above.

8. Continuity on q.

In this section we prove that our families $\{\underline{H}^{(q)}\}$ and $\{\overline{H}^{(q)}\}$ are continuous on q for $0 < q < 1$ and $q > 1$. We consider $q > 1$ and $0 \leqslant q < 1$ separately.

We start with $q > 1$. For this we make a simple estimate. Let now $1 < q_1 < q_2$. Then

$$\begin{aligned}\frac{q_2-1}{q_1-1}A^{(q_2)}(k,\varepsilon) &= -\frac{q_2-1}{q_1-1}\frac{1}{(q_2-1)k}\log\int \mu\{\mathcal{B}_k(x,\varepsilon)\}^{q_2-1}d\mu \\ &\geqslant -\frac{1}{(q_1-1)k}\log\int \mu\{\mathcal{B}_k(x,\varepsilon)\}^{q_1-1}d\mu = A^{(q_1)}(k,\varepsilon).\end{aligned}$$

Combining this with the monotonic properties from Section 1, we have

$$A^{(q_2)}(k,\varepsilon) \leqslant A^{(q_1)}(k,\varepsilon) \leqslant \frac{q_2-1}{q_1-1}A^{(q_2)}(k,\varepsilon)$$

for any $1 < q_1 < q_2$. This immediately yields to the inequalities for $\underline{H}^{(q)}$ and $\overline{H}^{(q)}$.

PROPOSITION 8.1. *For any $1 < q_1 < q_2$ the following inequalities hold*

$$\underline{H}^{(q_2)} \leqslant \underline{H}^{(q_1)} \leqslant \frac{q_2-1}{q_1-1}\underline{H}^{(q_2)}, \tag{8.1}$$

$$\overline{H}^{(q_2)} \leqslant \overline{H}^{(q_1)} \leqslant \frac{q_2-1}{q_1-1}\overline{H}^{(q_2)}. \tag{8.2}$$

Now we can easily prove

PROPOSITION 8.2. *Families $\{\underline{H}^{(q)}\}$ and $\{\overline{H}^{(q)}\}$ are continuous for $q > 1$.*

Proof: We prove this Proposition only for the family $\{\underline{H}^{(q)}\}$. For the family $\{\overline{H}^{(q)}\}$ the proof is completely similar. Consider any $q > 1$ and take any $\varepsilon > 0$. Let $\delta = (q-1)\dfrac{\varepsilon}{\varepsilon + h_{top}(f)}$. Then for every $q' > 1$ such that $|q - q'| < \delta$ the folllowing inequality holds

$$|\underline{H}^{(q)} - \underline{H}^{(q')}| < \varepsilon. \tag{8.3}$$

Indeed, consider first all $q' > q$ with $q' - q < \delta$. Then using the Proposition 8.1 we have

$$0 \leqslant \underline{H}^{(q)} - \underline{H}^{(q')} \leqslant \frac{q'-q}{q-1}\underline{H}^{(q')} \leqslant \frac{q'-q}{q-1}h_{top}(f)$$
$$< \frac{\delta}{q-1}h_{top}(f) = \frac{h_{top}(f)}{\varepsilon + h_{top}(f)}\varepsilon < \varepsilon.$$

Similary for any $1 < q' < q$ with $q - q' < \delta$ we have

$$0 \leqslant \underline{H}^{(q')} - \underline{H}^{(q)} \leqslant \frac{q-q'}{q'-1}\underline{H}^{(q)} \leqslant \frac{q-q'}{q'-1}h_{top}(f)$$
$$< \frac{\delta}{q-\delta-1}h_{top}(f) = \frac{(q-1)}{(q-1)}\frac{\varepsilon}{h_{top}(f)}h_{top}(f) = \varepsilon.$$

Hence our family $\{\underline{H}^{(q)}\}$ is continuous at each point q, $q > 1$. ∎

For the case $0 \leqslant q < 1$ the situation is not so good as for $q > 1$. One can not prove inequalities similar to (8.1) and (8.2). But one can get more complicated ones. For this we need Cauchy-Bunyakowskii inequality [Shir].

PROPOSITION 8.3. *Let ξ, η be any measurable functions on (X, μ). If $\int |\xi|^2 d\mu$, $\int |\eta|^2 d\mu < \infty$ then $\int |\xi\eta| d\mu < \infty$ and*

$$\left(\int |\xi\eta| d\mu\right)^2 \leqslant \int |\xi|^2 d\mu \int |\eta|^2 d\mu.$$

PROPOSITION 8.4. *Families $\{\underline{H}^{(q)}\}$ and $\{\overline{H}^{(q)}\}$ are continuous for $0 < q < 1$.*

Proof: Take any $0 < q < 1$ and for any $\delta \geqslant 0$, such that $0 \leqslant q - \delta$ and $q + \delta \neq 1$, we can apply the Cauchy–Bunyakowskii inequality to the functions

$$\xi(x) := \mu\{\mathcal{B}_k(x,\varepsilon)\}^{(q-\delta-1)/2}, \quad \eta(x) := \mu\{\mathcal{B}_k(x,\varepsilon)\}^{(q+\delta-1)/2}$$

Then we have

$$\left(\int \mu\{\mathcal{B}_k(x,\varepsilon)\}^{q-1} d\mu\right)^2 \leqslant \left(\int \mu\{\mathcal{B}_k(x,\varepsilon)\}^{q-\delta-1} d\mu\right)\left(\int \mu\{\mathcal{B}_k(x,\varepsilon)\}^{q+\delta-1} d\mu\right).$$

This immediately yields to the following inequality for $A^{(q)}(k,\varepsilon)$.

$$A^{(q)}(k,\varepsilon) = \frac{1}{(1-q)k}\log \int \mu\{\mathcal{B}_k(x,\varepsilon)\}^{q-1} d\mu$$
$$\leqslant \frac{1}{2(1-q)k}\left(\log \int \mu\{\mathcal{B}_k(x,\varepsilon)\}^{q-\delta-1} d\mu + \log \int \mu\{\mathcal{B}_k(x,\varepsilon)\}^{q+\delta-1} d\mu\right)$$
$$= \frac{1-q+\delta}{2(1-q)}A^{(q-\delta)}(k,\varepsilon) + \frac{1-q-\delta}{2(1-q)}A^{(q+\delta)}(k,\varepsilon). \tag{8.4}$$

Furthermore in this section we write $A^{(q)}$ instead of $A^{(q)}(k,\varepsilon)$. Then, using (8.4), we can estimate

$$0 \leqslant A^{(q)} - A^{(q+\delta)} \leqslant \frac{1}{2}\frac{1-q+\delta}{1-q}\left(A^{(q-\delta)} - A^{(q+\delta)}\right). \tag{8.5}$$

Take $0 < t < 1$ and any $n \in \mathbb{N}$. Let $\delta = t/2^n$. Using the previous formula several times we can show by induction

$$\begin{aligned} 0 \leqslant A^{(t)} - A^{(t+\delta)} &\leqslant \frac{1}{2}\frac{1-t+\delta}{1-t}\left(A^{(t-\delta)} - A^{(t+\delta)}\right) \\ &\leqslant \frac{1}{4}\frac{1-t+\delta}{1-t}\frac{1-t+3\delta}{1-t+\delta}\left(A^{(t-3\delta)} - A^{(t+\delta)}\right) \\ &\dots \\ &\leqslant \frac{1}{2^n}\frac{1-t+(2^n-1)\delta}{1-t}\left(A^{(t-(2^n-1)\delta)} - A^{(t+\delta)}\right) \\ &\leqslant \frac{1}{2^n}\frac{1}{1-t}A^{(0)} = \frac{\delta}{t(1-t)}A^{(0)}. \end{aligned} \tag{8.6}$$

Taking the relative limits on the both sides of (8.6) for $\delta = t/2^n$ we have

$$0 \leqslant \underline{H}^{(t)} - \underline{H}^{(t+t/2^n)} \leqslant \frac{1}{2^n(1-t)}h_{top}(f), \tag{8.7}$$

$$0 \leqslant \overline{H}^{(t)} - \overline{H}^{(t+t/2^n)} \leqslant \frac{1}{2^n(1-t)}h_{top}(f). \tag{8.8}$$

Now we are able to show that $\{\underline{H}^{(q)}\}$ and $\{\overline{H}^{(q)}\}$ are continuous on q for $0 < q < 1$. Again as in Proposition 8.2 we do it only for $\{\underline{H}^{(q)}\}$. Fix any $q \in (0,1)$ and take any $\varepsilon > 0$. Take any sufficiently large n such that $\frac{1}{2^n} < (1-q)\frac{\varepsilon}{h_{top}(f)+\varepsilon}$. And let $\delta = \frac{q}{2^n}$. Consider any $q' > q$, $q' \neq 1$, such that $q' - q < 0.5 * \delta$. Then using monotonic properties of $\{\underline{H}^{(q)}\}$ and (8.7) we have

$$0 \leqslant \underline{H}^{(q)} - \underline{H}^{(q')} \leqslant \underline{H}^{(q)} - \underline{H}^{(q+\delta)} \leqslant \frac{1}{2^n(1-q)}h_{top}(f) < \frac{h_{top}(f)\varepsilon}{h_{top}(f)+\varepsilon} < \varepsilon.$$

Now consider any $q' < q$ such that $q - q' < 0.5 * \delta$. Then using (8.7) for $t = q'$. Then first of all

$$q' + \frac{q'}{2^n} \geqslant q - \frac{\delta}{2} + \frac{q-\delta/2}{2^n} = q - \frac{q}{2^{n+1}} + \frac{q}{2^n} - \frac{q}{2^{2n+1}} = q + \frac{q}{2^{n+1}} - \frac{q}{2^{2n+1}} \geqslant q.$$

Since the monotonicity properties of $\{\underline{H}^{(q)}\}$ and (8.7) for $t = q'$ we have

$$\begin{aligned} 0 \leqslant \underline{H}^{(q')} - \underline{H}^{(q)} \leqslant \underline{H}^{(q')} - \underline{H}^{(q'+q'/2^n)} &\leqslant \frac{1}{2^n(1-q')}h_{top}(f) \\ &\leqslant \frac{1}{2^n(1-q)}h_{top}(f) < \frac{h_{top}(f)\varepsilon}{h_{top}(f)+\varepsilon} < \varepsilon. \end{aligned}$$

Combinig two cases together we can conclude that for every $0 < q < 1$ and each $\varepsilon > 0$ there exists $\delta > 0$ such that for every $0 \leqslant q', q' \neq 1$: $|q' - q| < \delta \Rightarrow |\underline{H}^{(q')} - \underline{H}^{(q)}| < \varepsilon$. ∎

Since we have proven the continuity of the families $\{\underline{H}^{(q)}\}$ and $\{\overline{H}^{(q)}\}$ we can define the left and right limits at $q = 1$

$$\overline{H}^{(1-0)} := \lim_{q \to 1-0} \overline{H}^{(q)}, \quad \underline{H}^{(1-0)} := \lim_{q \to 1-0} \underline{H}^{(q)},$$
$$\overline{H}^{(1+0)} := \lim_{q \to 1+0} \overline{H}^{(q)}, \quad \underline{H}^{(1+0)} := \lim_{q \to 1+0} \underline{H}^{(q)}.$$

By the continuity we can extend all our estimates to these limits,

$$\overline{H}^{(1-0)} \geqslant \underline{H}^{(1-0)} \geqslant h_\mu(f),$$
$$\underline{H}^{(1+0)} \leqslant \overline{H}^{(1+0)} \leqslant h_\mu(f).$$

Now it is easy to see that $\overline{H}^{(1-0)} = \overline{H}^{(1+0)}$ then there exists

$$\overline{H}^{(1)} = \lim_{q \to 1} \overline{H}^{(q)} = h_\mu(f).$$

In other words, if the family $\{\overline{H}^{(q)}\}$ is continuous at $q = 1$ then $\overline{H}^{(1)} = h_\mu(f)$, The same is true for $\{\underline{H}^{(q)}\}$.

9. Singularity at $q = 1$.

At the end of the previous section we have discussed that the continuity at $q = 1$ implies that $H^{(1)} = h_\mu(f)$. All the examples considered in this paper are continuous at $q = 1$. However, there are examples where one has a singularity at $q = 1$. In this section we discuss the nature of such examples. To show the duality between the spectrums of the Renyi dimensions and entropies we consider so-called expanding homeomorphisms for which spectrums are equalivalent, i.e. they have the same continuous or singular behavior at $q = 1$. At the end we introduce the Kapur entropy, which can be useful for the analysis for $q < 1$.

Consider $q > 1$. Take any finite partition ξ with the diameter less then ε. Then

$$\begin{aligned} A^{(q)}(k,\varepsilon) &= -\frac{1}{(q-1)k} \log \int \mu\left\{\mathcal{B}_k(x,\varepsilon)\right\}^{q-1} d\mu \\ &\leqslant -\frac{1}{(q-1)k} \log \int \mu\left\{\xi^{(k)}(x)\right\}^{q-1} d\mu = -\frac{1}{(q-1)k} \log \sum_{\Delta \in \xi^{(k)}} \mu(\Delta)^q \\ &\leqslant -\frac{q}{(q-1)k} \log \max_{\Delta \in \xi^{(k)}} \mu(\Delta). \end{aligned} \tag{9.1}$$

Theorem on asymptotic uniform distribution guarantees that the majority of the intervals decay exponentialy fast. However, there can be intervals which have a polynomial decay rate. Although their total measure is small, they are the leading terms in the asymptotics. One can construct such examples. In [STCK-87] the following example has been studied. Let $f : [0,1] \to [0,1]$ be defined as follows

$$f(x) = 1 - |x^r - (1-x)^r|^{1/r}, \text{ with } r > 1.$$

The invariant density is given by $P(x) = r(1-x)^{r-1}$. For this family of maps the leftmost interval $\Delta_1^{(k)}$ of the k-th iteration of an arbitrary partition exhibits a power-law behavior

$$\mu(\Delta_1^{(k)}) \propto k^{-s},$$

where $s > 0$. Therefore $\underline{H}^{(q)} = \overline{H}^{(q)} = 0$ for $q > 1$. Since, it is known that $h_\mu(f) = 0.5$ for $r = 2$ and $\overline{H}^{(q)} \geqslant \underline{H}^{(q)} \geqslant h_\mu(f) > 0$ for $0 \leqslant q < 1$, we have a singularity in the spectrum at $q = 1$.

In this example one has $f'(0) = 1$, and in general, the non-hyperbolic systems can be a source for a such kind of examples.

To discuss the duality between spectrums for dimensions and entropies we consider the *expanding* dynamical systems [PesWe-95],[Bar-95].

Definition 9.1. A continuous map $f : X \to X$ on a compact metric space (X, d) is expanding if it is a local homeomorphism at every point, and there exist constans $a \geqslant b > 1$ and $\varepsilon_0 > 0$, such that

$$\mathcal{B}(f(x), b\varepsilon) \subset f(\mathcal{B}(x,\varepsilon)) \subset \mathcal{B}(f(x), a\varepsilon), \tag{9.2}$$

for each $x \in X$ and $0 < \varepsilon < \varepsilon_0$.

Let f be an expanding map. We can rewrite the characteristic property (9.2) as

$$\mathcal{B}(x, \varepsilon/a) \subset f^{-1}\mathcal{B}(f(x),\varepsilon) \subset \mathcal{B}(x, \varepsilon/b).$$

Since

$$\mathcal{B}_k(x,\varepsilon) = \mathcal{B}(x,\varepsilon) \cap f^{-1}\mathcal{B}(f(x),\varepsilon) \cap \ldots \cap f^{-k+1}\mathcal{B}(f^{-k+1}(x),\varepsilon),$$

we have the following approximation

$$\mathcal{B}(x, \varepsilon/a^{-k+1}) \subset \mathcal{B}_k(x,\varepsilon) \subset \mathcal{B}(x, \varepsilon/b^{-k+1}).$$

Using the generalized spectrum for dimensions [Pes-93] defined for $q \neq 1$ as

$$\overline{D}(q) = \limsup_{\varepsilon \to 0} \frac{1}{q-1} \frac{1}{\log \varepsilon} \log \int \mu\{\mathcal{B}(x,\varepsilon)\}^{q-1} d\mu,$$
$$\underline{D}(q) = \liminf_{\varepsilon \to 0} \frac{1}{q-1} \frac{1}{\log \varepsilon} \log \int \mu\{\mathcal{B}(x,\varepsilon)\}^{q-1} d\mu.$$

we obtain the following estimates for any $q \neq 1$

$$\overline{D}(q)\log b \leqslant \overline{H}^{(q)} \leqslant \overline{D}(q)\log a,$$
$$\underline{D}(q)\log b \leqslant \underline{H}^{(q)} \leqslant \underline{D}(q)\log a.$$

Remark. This can be considered as a generalization of a relation between entropy and dimension for Cantor sets, which was found long time ago, see [Bill-65].

In [STCK-87] authors argue the smooth behavior on q up to $q = 1$. This is still an open question. Now, we introduce a Kapur entropy, which can be useful for the analysis at $0 \leqslant q < 1$.

Consider $0 \leqslant q < 1$. Take any finite partition ξ with the diameter less then ε. Then

$$A^{(q)}(k,\varepsilon) = \frac{1}{(1-q)k}\log\int\frac{1}{\mu\left\{\mathcal{B}_k(x,\varepsilon)\right\}^{1-q}}d\mu$$
$$\leqslant \frac{1}{(1-q)k}\log\int\frac{1}{\mu\left\{\xi^{(k)}(x)\right\}^{1-q}}d\mu = \frac{1}{(1-q)k}\log\sum_{\Delta\in\xi^{(k)}}\mu(\Delta)^q. \tag{9.2}$$

We use the following inequality from [HLP-52]. Let $\{p_i\}_{i=1}^n$ and $\{a_i\}_{i=1}^n$ are non-negative sequences then

$$\frac{\sum p_i\log a_i}{\sum p_i} \leqslant \log\frac{\sum p_i a_i}{\sum p_i} \leqslant \frac{\sum p_i a_i\log a_i}{\sum p_i a_i}. \tag{9.3}$$

Let $0 \leqslant q < 1$ and $p_i = \mu\{\Delta_i\}$, $a_i = \mu\{\Delta_i\}^{-(1-q)}$, where $\xi^{(k)} = \{\Delta_i\}$. Then combining the left-hand side of (9.3) with the estimate (9.2) for $A^{(q)}$ one has

$$A^{(q)}(k,\varepsilon) \leqslant \frac{1}{(1-q)k}\log\Big(\sum_{\Delta_i\in\xi^{(k)}}\mu(\Delta_i)^q\Big) \leqslant -\frac{1}{k}\frac{\sum_{\Delta_i\in\xi^{(k)}}\mu(\Delta_i)^q\log\mu(\Delta_i)}{\sum_{\Delta_i\in\xi^{(k)}}\mu(\Delta_i)^q} \tag{9.4}$$

for any partition ξ with the diameter less than ε.

Now we discuss some properties of the quantaties from (9.4) more detailed. For simplicity, we change the notation. We call $p = (p_1, \dots, p_n)$ a *probability vector* if $p_i \geqslant 0$ for any $i = 1, \dots, n$ and $\sum_i p_i = 1$. We introduce the following quantaties

(a) $H(p) = -\sum_{i=1}^{n} p_i\log p_i$ (Standard entropy),

(b) $H_{REN}^{(q)}(p) = \frac{1}{1-q}\log\Big(\sum_{i=1}^{n} p_i^q\Big)$, $0 \leqslant q < 1$, (Renyi entropy of type q),

(c) $H_{KAP}^{(q)}(p) = -\frac{\sum_{i=1}^{n} p_i^q\log p_i}{\sum_{i=1}^{n} p_i^q}$, $0 \leqslant q < 1$, (Kapur entropy of type q).

We use a standard agreement $0 \log 0 = 0$. The basic relation between these entropies is given by the following inequality

$$0 \leqslant H(p) \leqslant H^{(q)}_{REN}(p) \leqslant H^{(q)}_{KAP}(p)$$

for any probability vector p.

It is known, that the standard entropy function satisfies the Shannon's inequality, namely for any probablity vector $p = (p_1, \dots, p_n)$

$$H(p) \leqslant \log n. \tag{9.5}$$

One can easily show that the Renyi entropy satisfies the Shannon's inequality for any probability vector p and any $q, 0 \leqslant q < 1$. The situation is a little bit more complicated in the case of the Kapur entropy. In general it does satisfy the Shannon's inequality for an arbitrary probability vector $p = (p_1, \dots, p_n)$ and any q, but only for $q > q_0(n)$. For the discussion on the asymptotic of $q_0(n)$ as $n \to \infty$ see [Clau-83]. But for the purposes of the analysis at $q = 1$ the following simple proposition should be enough.

PROPOSTION 9.1. *Let $p = (p_1, \dots, p_n)$ be a probability vector, and $0 < q < 1$. Then*

$$H^{(q)}_{KAP}(p) \leqslant \frac{1}{q} \log n, \quad H^{(q)}_{REN}(p) \leqslant \log n. \tag{9.6}$$

Proof: The Shannon's inequality for the Renyi entropy can be obtained easily by the Lagrange's multiplier rule. We leave it without the proof. Let $p = (p_1, \dots, p_n)$ be a probability vector. Define $s_i = \dfrac{p_i^q}{\sum_i p_i^q}$. Then $s = (s_1, \dots, s_n)$ is a probability vector again. Sometimes it is called an escort probability vector, [BecSch-93].

$$\begin{aligned} H(s) &= -\sum_{i=1}^{n} s_i \log(s_i) = -\frac{\sum_{i=1}^{n} p_i^q \log p_i^q}{\sum_{i=1}^{n} p_i^q} + \frac{\sum_{i=1}^{n} p_i^q \log(\sum_{j=1}^{n} p_j^q)}{\sum_{i=1}^{n} p_i^q} \\ &= -q \frac{\sum_{i=1}^{n} p_i^q \log p_i^q}{\sum_{i=1}^{n} p_i^q} + (1-q)\frac{1}{1-q} \log\Big(\sum_{j=1}^{n} p_j^q\Big) \\ &= q H^{(q)}_{KAP}(p) + (1-q) H^{(q)}_{REN}(p) \geqslant q H^{(q)}_{KAP}(p). \end{aligned}$$

Combining the previous inequality and the Shannon's inequality to $H(s)$ one has

$$H^{(q)}_{KAP}(p) \leqslant \frac{1}{q} H(s) \leqslant \frac{1}{q} \log n.$$

This finishes the proof. ∎

It is easy to see that the families of generalized entropy functions ($H^{(q)}_{KAP}$ and $H^{(q)}_{REN}$) contain the standard entropy at $q = 1$. The defintion of the Kapur entropy can be extended to $q = 1$ without any problems, and for the Renyi entropy one has to apply l'Hôpital's rule. Unfortunately, the generalized entropies do not have an subadditivity property if $q \neq 1$, i.e. the basic property of the standard entropy

$$H(\xi \vee \eta) \leqslant H(\xi) + H(\eta),$$

where ξ and η are partitions, is not true for the Renyi and Kapur entropy. This makes the theory for the generalized entropies much more complicated. Nevertheless, we believe that it is possible to develop a rigorous theory for the generalized entropies along the lines of a classical ergodic theory for the standard entropy, see [Sin-95]. Such theory should have some interesting properties. Since our estimates (9.4) are almost the best possible, the knowledge of the properties of the generalized entopries would be useful for the analysis of our spectrum at $q = 1$.

References

[AczDar-75] J.Aczel, Z.Daroczy, *On Measures of Information and Their Characterizations*, Academic Press, N.Y, 1975.

[Bar-95] L.M.Barreira, *The dimension theory of hyperbolic dynamics*, preprint.

[Beck-90] C.Beck, *Upper and lower bounds on the Renyi dimensions and the uniformity of multifractals*, Physica **41D** (1990), 67-78.

[BecSch-93] C.Beck, P.Schlögl, *Thermodynamics of chaotic systems. An Introduction.*, Cambridge, 1993.

[Bill-65] P.Billingsley, *Ergodic Theory and Information*, John Wiley & Sons, 1965.

[Br&Ka-81] M.Brin,A.Katok, *On local entropy*, Geometric Dynamics, Lecture Notes in Mathematics, vol. 1007, Springer-Verlag, New York, 1983, pp. 30–38.

[Clau-83] A.Clausing, *Type t entropy and majorization*, SIAM J.Math. Anal. **14** (1983), 203–208.

[CsoSze1-88] A.Csordas, P.Szepfalusy, *Singularities in Renyi information as phase transitions in chaotic systems*, Phys.Rev A **39** (1989), 4767–4777.

[CsoSze2-88] A.Csordas, P.Szepfalusy, *Generalized entropy decay rate of one–dimensional maps*, Phys.Rev A **38** (1989), 2582-2587.

[Fuji-83] H.Fujisaka, *Statistical dynamics generated by fluctations of local Lyapunov exponents*, progress in Theor. Phys **70** (1983), 1264–1275.

[GraPro-83] P.Grassberger, I.Procaccia, *Estimation of the Kolmogorov entropy from a chaotic signal*, Phys. Rev A **28** (1983), 2591-2593.

[HLP-52] C.Hardy, T.Littlewood, G.Polya, *Inequalities*, Second edition, Cambridge, 1952.

[Kat-80] A.Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Publication Mathematique I.H.E.S. **51** (1980), 137–173.

[Led-81] F.Ledrappier, *Some Relations Between Dimension and Lyapunov exponents*, Commun. Math. Phys. **81** (1981), 229-238.

[Pes-93] Y.Pesin, *On Rigurous Mathematical Definitions of Correlation Dimension and Generalized Spectrum for Dimensions*, Journal of Stat. Phys. **7** (1993).

[PesWe-95] Y.Pesin, H.Weiss, *A multifractal analysis of equilibrium measures for conformal expanding maps and Moran-like geometric constructions*, preprint.

[Ren-70] A.Renyi, *Probability Theory*, North-Holland, Amsterdam, 1970.

[Shir-84] A.N.Shiriayev, *Probability*, Springer-Verlag, New York, 1984.

[Sinai-95] Ya.G.Sinai, *Topics in ergodic theory*, Princeton University Press, Princeton, N.J., 1995.

[STCK-87] P.Szepfalusy, T.Tel, A.Csordas, Z.Kovacs, *Phase transitions associated with dynamical properties of chaotic systems*, Phys.Rev A **36** (1987), 3525-3528.

[Ward-94] T.Ward, *Lecture Notes: Entropy of Compact Group Automorphisms*, Available at http://www.mth.uea.ac.uk/~h720/lecture_notes/lecture_notes.html.

[Young-82] L.Young, *Dimension, entropy and Lyapunov exponents*, Ergod. Th. & Dynam. Sys. 2 (1982), 109–124.