Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

Website: http://AIMsciences.org

pp. **X–XX**

ENTROPY AND VARIATIONAL PRINCIPLES FOR HOLONOMIC PROBABILITIES OF IFS

Artur O. Lopes

Instituto de Matemática, UFRGS, 91509-900, Porto Alegre, Brasil

Elismar R. Oliveira

Instituto de Matemática, UFRGS, 91509-900, Porto Alegre, Brasil

(Communicated by the associate editor name)

Abstract. An IFS (iterated function system), $([0,1],\tau_i)$, on the interval $[0,1]$, is a family of continuous functions $\tau_0,\tau_1,...,\tau_{d-1}:[0,1]\to[0,1]$.

Associated to a IFS one can consider a continuous map $\hat{\sigma}:[0,1]\times\Sigma\to[0,1]\times\Sigma$, defined by $\hat{\sigma}(x,w)=(\tau_{X_1(w)}(x),\sigma(w))$ were $\Sigma=\{0,1,...,d-1\}^{\mathbb{N}}$, $\sigma:\Sigma\to\Sigma$ is given by $\sigma(w_1,w_2,w_3,...)=(w_2,w_3,w_4...)$ and $X_k:\Sigma\to\{0,1,...,n-1\}$ is the projection on the coordinate k.

A ρ-weighted system, $\rho\geq 0$, is a weighted system $([0,1],\tau_i,u_i)$ such that there exists a positive bounded function $h:[0,1]\to\mathbb{R}$ and a probability ν on $[0,1]$ satisfying $P_u(h)=\rho h, \quad P_u^*(\nu)=\rho\nu$.

A probability $\hat{\nu}$ on $[0,1]\times\Sigma$ is called holonomic for $\hat{\sigma}$, if, $\int g\circ\hat{\sigma}\,d\hat{\nu}=\int g\,d\hat{\nu}$, $\forall g\in C([0,1])$. We denote the set of holonomic probabilities by $\mathcal{H}$.

Via disintegration, holonomic probabilities $\hat{\nu}$ on $[0,1]\times\Sigma$ are naturally associated to a ρ-weighted system. More precisely, there exist a probability ν on $[0,1]$ and $u_i, i\in\{0,1,2,..,d-1\}$ on $[0,1]$, such that is $P_u^*(\nu)=\nu$.

We consider holonomic ergodic probabilities and present the corresponding Ergodic Theorem (which is just an adaptation of a previous result by J. Elton).

For a holonomic probability $\hat{\nu}$ on $[0,1]\times\Sigma$ we define the entropy $h(\hat{\nu})=\inf_{f\in\mathbb{B}^+}\int\ln(\frac{P_\psi f}{\psi f})d\hat{\nu}\geq 0$, where, $\psi\in\mathbb{B}^+$ is a fixed (any one) positive potential.

Finally, we analyze the problem: given $\phi\in\mathbb{B}^+$, find solutions of the maximization problem

$$p(\phi)=\sup_{\hat{\nu}\in\mathcal{H}}\{\,h(\hat{\nu})+\int\ln(\phi)d\hat{\nu}\,\}.$$

We show an example where a holonomic not-$\hat{\sigma}$-invariant probability attains the supremum value.

In the last section we consider maximizing probabilities, sub-actions and duality for potentials $A(x,w)$, $(x,w)\in[0,1]\times\Sigma$, for IFS.

2000 *Mathematics Subject Classification.* Primary: 37A30, 37A35; Secondary: 37A50.

Key words and phrases. Iterated Function System, holonomic probability, entropy pressure, maximizing probability.

The first author is partially supported by CNPq, PRONEX – Sistemas Dinâmicos, Instituto do Milênio, Projeto Universal - CNPq and beneficiary of CAPES financial support and the second by CNPq.

1. **IFS and holonomic probabilities.** We want to analyze, in the setting of holonomic probabilities [13] associated to an IFS, the concepts of entropy and pressure.

We point out that this is a different problem from the usual one among invariant probabilities (see remarks 3 and 4 in section 7).

The present work is part of the PhD thesis of the second author [21].

Our main point of view is the following: the study of the holonomic probabilities allows one to understand all the transference operators P_u and the associated stationary states when the IFS is considered as a realization of a Stochastic Process.

Definition 1.1. An IFS (iterated function system), $([0,1],\tau_i)$, on the interval $[0,1]$, is a family of continuous functions $\tau_0,\tau_1,...,\tau_{d-1}:[0,1]\to[0,1]$.

Associated to a IFS one can consider a continuous map $\hat{\sigma}:[0,1]\times\Sigma\to[0,1]\times\Sigma$, defined by

$$\hat{\sigma}(x,w)=(\tau_{X_1(w)}(x),\sigma(w)),$$

were $\Sigma=\{0,1,...,d-1\}^{\mathbb{N}}$, $\sigma:\Sigma\to\Sigma$ is given by $\sigma(w_1,w_2,w_3,...)=(w_2,w_3,w_4...)$ and $X_k:\Sigma\to\{0,1,...,n-1\}$ is the projection on the coordinate k. In this way one can see such system as a Stochastic Process [5] [17] [3] [9] [20] [8] [15] [16] [22].

If we consider a IFS as a multiple dynamical systems (several maps) then, for a single point x there exists several combinations of "orbits" on the IFS (using different τ_i). Considering the map $\hat{\sigma}$ one can describe the global behavior of iterates of x. Moreover, one can think the IFS, as a branching process with index in Σ. More precisely, we define the n-branch from $x\in[0,1]$ by $w\in\Sigma$, as

$$Z_n(x,w)=\tau_{X_n(w)}\circ\tau_{X_{n-1}(w)}\circ...\tau_{X_1(w)}(x).$$

With this notation, we have

$$\hat{\sigma}^n(x,w)=(Z_n(x,w),\sigma^n(w)).$$

Definition 1.2. A weighted system (see [24], Pg. 6) is a triple, $([0,1],\tau_i,v_i)$, were $([0,1],\tau_i)$ is a IFS where the v_i's, $i\in\{0,1,...,d-1\}$, are measurable and nonnegative bounded maps. The condition $\sum_{i=0}^{d-1}v_i(x)=1$ is not required.

In some examples the u_i, $i\in\{0,1,...,d-1\}$, come from a measurable bounded potential $\phi:[0,1]\to[0,+\infty)$, that is,

$$u_i(x)=\phi(\tau_i(x)),\ \forall i=0,...,d-1\,.$$

The function ϕ is called *weight function*, (in the literature this function is also called g-function, see [18] for example). Note that ϕ can attain the value 0. This is useful for some applications of IFS to wavelets [7] [8]. We do not assume in this general definition that $\sum_{i=0}^{d-1}u_i(x)=1,\quad\forall x\in[0,1]$.

Definition 1.3. A IFS with probabilities, $([0,1],\tau_i,u_i)$, $i\in\{0,1,...,d-1\}$, is a IFS with a vector of measurable nonnegative bounded functions on $[0,1]$,

$$u(x)=(u_0(x),u_1(x),...,u_{d-1}(x)),$$

such that, $\sum_{i=0}^{d-1}u_i(x)=1,\quad\forall x\in[0,1]$.

Definition 1.4. A IFS with probabilities $([0,1],\tau_i,u_i)$ is called "uniform normalized" if there exists a weight function ϕ such that $u_i(x)=\phi(\tau_i(x))$, $\forall i=0,...,d-1$ and

$$\sum_{i=0}^{d-1}\phi(\tau_i(x))=1.$$

In this case, we write the IFS as $([0,1],\tau_i,u_i)=([0,1],\tau_i,\phi)$.

The above definition is a strong restriction in the weighted system. Several problems in the classical theory of Thermodynamic Formalism for the shift or for a d to 1 continuous expanding transformations $T:S^1\to S^1$ can be analyzed via a IFS with a weight function ϕ (see [23]). In this case the $\tau_i, i\in\{0,2,..,d-1\}$, are the inverse branches of T.

We will consider later the pressure problem for a weight function ϕ which is not necessarily uniformly normalized.

We now return to the general case.

Definition 1.5. Given a weighted system, $([0,1],\tau_i,u_i)$, we will define de Transference Operator (or Ruelle Operator) P_u by

$$P_u(f)(x)=\sum_{i=0}^{d-1}u_i(x)f(\tau_i(x)),$$

for all $f:[0,1]\to\mathbb{R}$ bounded Borel measurable functions.

A function $h:[0,1]\to\mathbb{R}$ will be called P_u-harmonic if $P_u(h)=h$ [17] [5]. A probability ν on $[0,1]$ will be called P_u-invariant if $P_u^*(\nu)=\nu$, where P_u^* is defined by equality

$$\int P_u(f)(x)d\hat{\nu}=\int f(x)dP_u^*\hat{\nu},$$

for all $f:[0,1]\to\mathbb{R}$ continuous.

The correct approach to analyze an IFS [9] [5] [17] [15] [16] with probabilities $([0,1],\tau_i,u_i)$, is to consider for each $x\in[0,1]$, the sequence of random variables $(Z_n(x,.):\Sigma\to[0,1])_{n\in\mathbb{N}}$ as a realization of the Markov process associated to the Markov chain with initial distribution δ_x and transitions of probability P_u. Moreover, we have a probability $\mathbb{P}_x$ in the space of paths, Σ, given by

$$\int_\Sigma g(w)d\mathbb{P}_x=\sum_{i_1,\dots,i_n}u_{i_1}(x)u_{i_2}(\tau_{i_1}(x))...u_{i_n}(\tau_{i_n}...\tau_{i_1}(x))g(i_1,...,i_n),$$

when $g=g(x,w)$ depends only of the n first coordinates (see [17], for a proof of the Komolgorov consistence condition).

The probability on path space and the transference operator are connected by

$$\int_\Sigma g(x,w)d\mathbb{P}_x(w)=P_u^n(f)(x),$$

when $g(x,w)=f(\tau_{w_n}...\tau_{w_1}(x))$ for some continuous f.

Definition 1.6. A ρ-weighted system, $\rho\geq 0$, is a weighted system $([0,1],\tau_i,u_i)$ such that there exists a positive bounded function $h:[0,1]\to\mathbb{R}$ and ν probability satisfying

$$P_u(h)=\rho h,\quad P_u^*(\nu)=\rho\nu.$$

Note that a IFS with probabilities is a 1-weighted system (see [8], [24] [20] [27] or [3] for the existence of P_u-invariant probabilities and [24], Theorem 4, or [26] for non-uniqueness of this probabilities). Also, a weighted system, $([0,1],\tau_i,u_i)$ were $u_0=...=u_{d-1}=k$ (constant) is a $d\,k$-weighted system. Thus the set of ρ-weighted systems is as big class of weighted systems.

Examples of nontrivial ρ-weighted system (and non-probabilistic) can be found in [12] and [24] Corollary 2.

Moreover, from a ρ-weighted system $([0,1],\tau_i,u_i)$ one can get a normalization $([0,1],\tau_i,v_i)$, in the following way

$$v_i(x)=\frac{u_i(x)h(\tau_i(x))}{\rho h(x)},\quad \mu=h\nu.$$

Then $P_v(1)=1$ and $P_v^*(\mu)=\mu$.

We thanks an anonymous referee for some comments on a previous version of the present paper. We would like to point out that there exists some similarities of sections 1, 2, 3 and 5 of our paper with some results in [15] and [16]. We would like to stress that we consider here the holonomic setting which can not be transfer by some coding to the usual shift case (see remark 4 in section 7). We introduce for such class of probabilities in IFS (which is different from the set of invariant probabilities for $\hat{\sigma}$) the concept of entropy and pressure. It is not the same same concept of entropy as for a measure invariant for the shift $\hat{\sigma}$ (see remarks 3 and 4 in section 7). Also, in our setting, it is natural to consider the all set of possible potentials u. In this way our results are of different nature than the ones in [15] [16] where the dynamical concepts are mainly consider for the shift $\hat{\sigma}$ acting on $[0,1]\times\Sigma$.

In sections 1 to 6 we consider the basic definitions and results. In sections 7 and 8 we introduce entropy and pressure for holonomic probabilities of IFS. In section 9 we consider maximizing probabilities for IFS.

2. **Holonomic probabilities.** For IFS we introduce the concept of *holonomic* probability on $[0,1]\times\Sigma$ (see [13] for general definitions and properties in the setting of symbolic dynamics of the two-sided shift in $\Sigma\times\Sigma$). Several results presented in [13] can be easily translated for the IFS setting. In [13] the main concern was maximizing probabilities. Here we are mainly interested in the variational principle of pressure.

By the other hand, some of the new results we presented here can also be translated to that setting.

Definition 2.1. A holonomic probability $\hat{\nu}$ on $[0,1]\times\Sigma$ is a probability such that

$$\int f(\tau_{X_1(w)}(x))d\hat{\nu}=\int f(x)d\hat{\nu},$$

for all $f:[0,1]\to\mathbb{R}$ continuous.

Then the set of holonomic probabilities can be viewed as the set of probabilities on $[0,1]\times\Sigma$ such that

$$\int g\circ\hat{\sigma}d\hat{\nu}=\int gd\hat{\nu},\quad \forall g\in C([0,1]).$$

From this point of view it is clear that the set of holonomic probabilities is bigger (see section 4) than the set of $\hat{\sigma}$-invariant probabilities (because $C([0,1])$ can be viewed as a subset of $C([0,1]\times\Sigma)$).

3. **Characterization of holonomic probabilities.** Disintegration of probabilities for IFS have been previously consider but for a different purpose [7], [8], [16].

Definition 3.1. A Hausdorff space is a Radon space if all probabilities in this space is Radon (See [25]).

Theorem 3.2. *([25], Prop. 6, Pg. 117) All compact metric space is Radon.*

Therefore, all spaces considered here are Radon spaces.

Theorem 3.3. *([6], Pg 78, (70-III) or [1], Theorem 5.3.1) Let $\mathbb{X}$ and X be a separable metric Radon spaces, $\hat{\mu}$ probability on $\mathbb{X}$, $\pi : \mathbb{X} \to X$ Borel mensurable and $\mu = \pi_*\hat{\mu}$. Then there exists a Borel family of probabilities $\{\hat{\mu}\}_{x\in X}$ on $\mathbb{X}$, uniquely determined μ-a.e, such that,*
1) $\hat{\mu}_x(\mathbb{X}\backslash\pi^{-1}(x)) = 0$, μ-a.e;
2) $\int g(z)d\hat{\mu}(z) = \int_X \int_{\pi^{-1}(x)} g(z)d\hat{\mu}_x(z)d\mu(x)$.

This decomposition is called the *disintegration* of the probability $\hat{\mu}$.

Theorem 3.4. *(Holonomic Disintegration) Consider a holonomic probability $\hat{\nu}$ on $[0,1]\times\Sigma$. Let*

$$\int g(x,w)d\hat{\nu} = \int_{[0,1]}\int_{\{y\}\times\Sigma} g(x,w)d\hat{\nu}_y(x,w)d\nu(y), \quad \forall g,$$

be the disintegration given by Theorem 3.3. Then ν is P_u-invariant for the IFS with probabilities $([0,1],\tau_i,u_i)_{i=0..d-1}$, were the u_i's are given by,

$$u_i(y) = \hat{\nu}_y(y,\bar{i}), \quad i = 0,...,d-1.$$

Proof. Consider a continuous function $f:[0,1]\to\mathbb{R}$ and defines $I_1 = \int f(\tau_{X_1(w)}(x))d\hat{\nu}$ and $I_2 = \int f(x)d\hat{\nu}$. As $\hat{\nu}$ is holonomic we have, $I_1 = I_2$.

Now applying the disintegration for both integrals we get
$I_1 = \int_{[0,1]}\int_{\{y\}\times\Sigma} f(\tau_{X_1(w)}(x))d\hat{\nu}_y(x,w)d\nu(y) =$
$= \sum_{i=0}^{d-1}\int_{[0,1]}\int_{\{y\}\times\bar{i}} f(\tau_{X_1(w)}(x))d\hat{\nu}_y(x,w)d\nu(y) =$
$= \sum_{i=0}^{d-1}\int_{[0,1]} f(\tau_{\bar{i}}(y))\hat{\nu}_y(\{y\}\times\bar{i})d\nu(y) = \int_{[0,1]} P_u(f)(y)d\nu(y)$
when $u_i(y) = \hat{\nu}_y(y,\bar{i}), \quad i = 0,...,d-1$.

On the other hand
$I_2 = \int_{[0,1]}\int_{\{y\}\times\Sigma} f(x)d\hat{\nu}_y(x,w)d\nu(y) = \int_{[0,1]} f(y)d\nu(y)$.

Then, $\int_{[0,1]} P_u(f)(y)d\nu(y) = \int_{[0,1]} f(y)d\nu(y)$ for all continuous function $f:[0,1]\to\mathbb{R}$, that is, ν is P_u-invariant.

□

4. **Invariance of Holonomic probabilities on IFS.** As we said before, holonomic probabilities are not necessarily invariant for the map $\hat{\sigma}$. On the other hand all $\hat{\sigma}$-invariant probability is holonomic. Now we show an example of holonomic probability which is not $\hat{\sigma}$-invariant (see [13] for the case of the two sided shift).

Suppose that $x_0\in[0,1]$, is such that $Z_n(x_0,\bar{w}) = x_0$, for some $\bar{w}\in\Sigma$, $n\in\mathbb{N}$. Then, one can obtain a holonomic probability in the following way

$$\hat{\nu} = \frac{1}{n}\sum_{j=0}^{n-1}\delta_{\sigma^j(\bar{w})}\times\delta_{Z_j(x_0,\bar{w})}.$$

Then,

$$\int g(x,w)d\hat{\nu} = \frac{1}{n}\sum_{j=0}^{n-1} g(Z_j(x_0,\bar{w}),\sigma^j(\bar{w})),$$

Note that this probability is holonomic but not $\hat{\sigma}$-invariant.

In fact, it is enough to see that

$$\int g\circ\hat{\sigma}(x,w)d\hat{\nu} = \frac{1}{n}\sum_{j=0}^{n-1} g\circ\hat{\sigma}(Z_j(x_0,\bar{w}),\sigma^j(\bar{w})) =$$

$$= \frac{1}{n}\sum_{j=0}^{n-1} g(Z_{j+1}(x_0,\bar{w}),\sigma^{j+1}(\bar{w})).$$

Thus, $\int g\circ\hat{\sigma}(x,w)d\hat{\nu} - \int g(x,w)d\hat{\nu} = \frac{1}{n}g(x_0,\sigma^n(\bar{w})) - g(x_0,\bar{w}))$, and it is clearly not identical to 0, $\forall g$.

However, $\hat{\nu}$ is holonomic because given any continuous function $f:[0,1]\to\mathbb{R}$ we have

$$\int f(\tau_{X_1(w)}(x))d\hat{\nu} = \frac{1}{n}\sum_{j=0}^{n-1} f(\tau_{X_1(\sigma^j(\bar{w}))}(Z_j(x_0,\bar{w})) =$$

$$= \frac{1}{n}\sum_{j=0}^{n-1} f(Z_{j+1}(x_0,\bar{w})) = \int f(x)d\hat{\nu},$$

because, $Z_n(x_0,\bar{w}) = x_0$.

5. **Ergodicity of holonomic probabilities.**

Given a holonomic probability $\hat{\nu}$, we can associate, by holonomic disintegration, a unique IFS with probabilities $([0,1],\tau_i,u_i)$ such that $P_u^*(\nu)=\nu$ and $\nu=\pi_*\hat{\nu}$.

Let $Z_n(\cdot):[0,1]\hookleftarrow,\ n\in\mathbb{N}$, be a sequence of random variables on $[0,1]$. Then, we obtain a Markov process with transition of probabilities P_u and initial distribution ν, that we will denote by (Z_n,P_u,ν).

This process is a stationary process by construction, thus does make sense to ask if (Z_n,P_u,ν) is ergodic([9] for details of this process and definition of ergodicity).

Definition 5.1. A holonomic probability $\hat{\nu}$ is called ergodic, if the associated Markov process (Z_n,P_u,ν) is an ergodic process.

Lemma 5.2. *(Elton,[9]) Let $\hat{\nu}$ be a holonomic probability with holonomic disintegration $([0,1],\tau_i,u_i)$. If $\pi_*\hat{\nu}$ is the unique P_u-invariant probability, then $\hat{\nu}$ is ergodic.*

Theorem 5.3. *(Elton,[9]) Let $([0,1],\tau_i)$ be a contractive IFS (contractiveness means that τ_i is a contraction for all i) and $\hat{\nu}$ be a ergodic holonomic probability with holonomic disintegration $([0,1],\tau_i,u_i)$. Suppose that $u_i\geq\delta>0$, $\forall i=0,...,d-1$. Then, $\forall x\in[0,1]$ there exists $G_x\subseteq\Sigma$ such that $\mathbb{P}_x(G_x)=1$ and for each $w\in G_x$*

$$\frac{1}{N}\sum_{i=0}^{N-1} f(Z_n(x,w)) = \frac{1}{N}\sum_{i=0}^{N-1} f(\hat{\sigma}^i(x,w)) \to \int f d\hat{\nu} = \int f(x)d\hat{\nu}(x,w),$$

for all $f\in C([0,1])$.

Proof. The proof is a straightforward modification of the one presented in Elton's ergodic theorem (see [9] [16]). In fact, the contractiveness of $([0,1],\tau_i)$ its stronger that Dini condition that appear in Elton's proof (see [9] and [3]) and the ergodicity of $\hat{\nu}$ can replace the uniqueness of the initial distribution in the last part of the argument. The other parts of the proof are the same as in [9]. $\square$

We point out that Elton's Theorem is not the classical ergodic theorem for $\hat{\sigma}$. The claim is: $\forall x \in [0,1]$ there exists $G_x \subseteq \Sigma$ such that $\mathbb{P}_x(G_x) = 1$ and for each $w \in G_x$ ... Moreover, $f : [0,1] \to \mathbb{R}$.

This theorems fits well for holonomic probabilities in the IFS case. We just mention it in order to give to the reader a broader perspective of the holonomic setting. We do not use it in the rest of the paper.

6. **Construction of holonomic probabilities for ρ-weighted systems.** Given a ρ-weighted system $([0,1],\tau_i,u_i)$, that is,

$$P_u(h) = \rho h, \quad P_u^*(\nu) = \rho\nu,$$

consider the normalization $([0,1],\tau_i,v_i)$, then $P_v(1) = 1$ and $P_v^*(\mu) = \mu$.

Its easy to see that the probability on $[0,1]\times\Sigma$ given by

$$\int g(x,w)d\hat{\mu} = \int_{[0,1]}\int_{\Sigma} g(x,w)d\mathbb{P}_x(w)d\mu(x)$$

is holonomic if $\mathbb{P}_x$ is given from v (see [7][15][16] for disintegration of projective measures on IFS). The probability $\hat{\mu}$ will be called the *holonomic lifting* of μ.

Remark 1. We point out that that the holonomic lifting $\hat{\mu}$ of a given μ (as above) is a $\hat{\sigma}$-invariant probability (one can see that by taking functions that depends only of finite symbols and applying de definition of a P_x probability). So

$$\pi_*\{\textit{Holonomic probabilities}\} = \pi_*\{\hat{\sigma} - \textit{invariant probabilities}\}.$$

We will consider in the next sections the concept of pressure. The value of pressure among holonomic or among invariant will be the same. However one cannot reduce the study of variational problems involving holonomic probabilities to the study of $\hat{\sigma}$-invariant probabilities. This will be explained in remark 4 in example 3 after Theorem 7.7.

One can reverse the process, starting from a IFS with probabilities (a 1-weighted system) $([0,1],\tau_i,v_i)$, that is, $P_v(1) = 1$ and $P_v^*(\nu) = \nu$, and consider the associated $\hat{\nu}$, the holonomic lifting of ν.

$$\int g(x,w)d\hat{\nu} = \int_{[0,1]}\int_{\Sigma} g(y,w)d\mathbb{P}_y(w)d\nu(y).$$

By holonomic disintegration (Theorem 3.4), one can represents the probability $\hat{\nu}$ as

$$\int g(x,w)d\hat{\nu} = \int_{[0,1]}\int_{\{y\}\times\Sigma} g(x,w)d\hat{\nu}_y(x,w)d\nu_0(y).$$

Then, ν_0 is P_u-invariant for the IFS with probabilities $([0,1],\tau_i,u_i)_{i=0..d-1}$, were the u_i's are given by,

$$u_i(y) = \hat{\nu}_y(y,\bar{i}), \quad i = 0,...,d-1\,.$$

We point out that $\nu_0 = \nu$ (it is a straightforward calculation), moreover we can rewrite

$$\int g(x,w)d\hat{\nu} = \int_{[0,1]}\int_{y\times\Sigma} g(x,w)d(\delta_y(x)\times\mathbb{P}_y(w))d\nu(y).$$

By the uniqueness in Theorem 3.3 we get,

$$\hat{\nu}_y = \delta_y \times \mathbb{P}_y, \quad \nu - a.e\,.$$

Then, we have

$$u_i(y) = \hat{\nu}_y(y,\bar{i}) = (\delta_y \times \mathbb{P}_y)(y,\bar{i}) = \mathbb{P}_y(\bar{i}) = v_i(x), \quad \nu - a.e\,.$$

From this argument we get the following proposition

Proposition 1. *Let $([0,1],\tau_i,v_i)$ be a 1-weighted system and $\hat{\nu}$ the holonomic lifting of the invariant probability P_v-invariant ν. If $([0,1],\tau_i,u_i)$ is the 1-weighted system obtained from holonomic disintegration of $\hat{\nu}$, then $u = v$, $\quad \nu - a.e$, where $\nu = \pi_* \hat{\nu}$.*

7. **Entropy and a variational principle for ρ-weighted systems.** Let us consider a ρ-weighted system, $([0,1],\tau_i,u_i)$. Denote by $\mathbb{B}^+$ the set of all positive bounded Borel functions on $[0,1]$ and by $\mathcal{H}$ the set of all holonomic probabilities on $[0,1]\times\Sigma$ for $\hat{\sigma}$.

The central idea in this section is to consider a generalization of the definition of entropy for the case of holonomic probabilities via the concept naturally suggested by Theorem 4 in [19]. We will show that under such point of view the classical results in Thermodynamic Formalism are also true.

Given $\hat{\nu} \in \mathcal{H}$ we can define the functional $\alpha_{\hat{\nu}} : \mathbb{B}^+ \to \mathbb{R}\cup\{-\infty\}$ by

$$\alpha_{\hat{\nu}}(\psi) = \inf_{f\in\mathbb{B}^+} \int \ln(\frac{P_\psi f}{\psi f}) d\hat{\nu}.$$

Let $\alpha_{\hat{\nu}}$ be the functional defined above. Observe that $\alpha_{\hat{\nu}}$ doesn't depend of ψ.

In fact, by taking $\psi_1,\psi_2 \in \mathbb{B}^+$ and $f \in \mathbb{B}^+$, define $g \in \mathbb{B}^+$ as being $g = \frac{\psi_1}{\psi_2} f \in \mathbb{B}^+$, then

$$\int \ln(\frac{P_{\psi_2} g}{\psi_2 g}) d\hat{\nu} = \int \ln(\frac{P_{\psi_1} f}{\psi_1 f}) d\hat{\nu}.$$

Thus, $\alpha_{\hat{\nu}}(\psi_2) = \alpha_{\hat{\nu}}(\psi_1)$.

Definition 7.1. Given $\hat{\nu} \in \mathcal{H}$ we define the Entropy of $\hat{\nu}$ by

$$h(\hat{\nu}) = \alpha_{\hat{\nu}}.$$

From above we get

$$h(\hat{\nu}) = \inf_{f\in\mathbb{B}^+} \int \ln(\frac{P_\psi f}{\psi f}) d\hat{\nu},$$

$\forall \psi \in \mathbb{B}^+$.

The above definition agrees with the usual one for invariant probabilites when it is consider a transformation of degree d, its d-branches and the naturally associated IFS (see [19]).

We point out that this way to define entropy it's also natural to consider for C^*-Algebras [4].

Lemma 7.2. *Given $\beta \geq 1+\alpha$ and numbers $a_i \in [1+\alpha,\beta]$, $i = 0,...,d-1$ there exists $\varepsilon \geq 1$ such that*

$$\ln(\varepsilon \sum_{i=0}^{d-1} a_i) = \sum_{i=0}^{d-1} \ln(\varepsilon a_i).$$

This lemma follows from the choice $\varepsilon = exp(\frac{1}{d-1}(\ln(\sum_{i=0}^{d-1} a_i))/(\sum_{i=0}^{d-1}\ln(a_i))$ and the fact that $a_i \geq 1+\alpha$.

Lemma 7.3. *Given $f \in \mathbb{B}^+$ and $\hat{\nu} \in \mathcal{H}$ then*
$$\sum_{i=0}^{d-1} \int f(\tau_i(x))d\hat{\nu} \geq \int f(x)d\hat{\nu}.$$
.

Proof. As $\hat{\nu}$ is holonomic, then we have
$$\int f(\tau_{X_1(w)}(x))d\hat{\nu} = \int f(x)d\hat{\nu}.$$
This can be written as
$$\int f(\tau_{X_1(w)}(x))d\hat{\nu} = \sum_{i=0}^{d-1} \int_{[0,1]\times\bar{i}} f(\tau_{X_1(w)}(x))d\hat{\nu} = \sum_{i=0}^{d-1} \int_{[0,1]\times\bar{i}} f(\tau_i(x))d\hat{\nu}.$$
Note that for each i
$$\int f(\tau_i(x))d\hat{\nu} = \sum_{j=0}^{d-1} \int_{[0,1]\times\bar{j}} f(\tau_i(x))d\hat{\nu} \geq \int_{[0,1]\times\bar{i}} f(\tau_i(x))d\hat{\nu}.$$
Thus,
$$\sum_{i=0}^{d-1} \int f(\tau_i(x))d\hat{\nu} \geq \sum_{i=0}^{d-1} \int_{[0,1]\times\bar{i}} f(\tau_i(x))d\hat{\nu} = \int f(x)d\hat{\nu}.$$
□

Proposition 2. *Consider $\hat{\nu} \in \mathcal{H}$. Then*
$$0 \leq h(\hat{\nu}) \leq \ln(d).$$

Proof. Initially, consider $\psi = 1$. We know that $h(\hat{\nu}) = \inf_{f\in\mathbb{B}^+} \int \ln(\frac{P_1 f}{f})d\hat{\nu} \leq \int \ln(\frac{P_1 1}{1})d\hat{\nu} = \ln(d)$.

Now, in order to prove the inequality
$$h(\hat{\nu}) = \inf_{f\in\mathbb{B}^+} \int \ln(\sum_{i=0}^{d-1} \frac{f\circ\tau_i}{f})d\hat{\nu} \geq 0,$$
consider $I = \int \ln(\sum_{i=0}^{d-1} \frac{f\circ\tau_i}{f})d\hat{\nu}$ and suppose, without lost of generality, that $1+\alpha \leq f \leq \beta$ (because this integral is invariant under the projective function $f \to \lambda f$). Then, we can write this integral as
$$I = \int \ln(\sum_{i=0}^{d-1} \frac{\varepsilon f\circ\tau_i}{\varepsilon f})d\hat{\nu} = \int \ln(\varepsilon \sum_{i=0}^{d-1} f\circ\tau_i)d\hat{\nu} - \int \ln(\varepsilon f)d\hat{\nu}, \qquad (1)$$
In order to use the inequality obtained from Lemma 7.2, denote (for each fixed x)
$$a_i = f\circ\tau_i(x).$$
Then, we get
$$\varepsilon(x) = exp(\frac{1}{d-1} \cdot \frac{\ln(\sum_{i=0}^{d-1} f\circ\tau_i)}{(\sum_{i=0}^{d-1} \ln(f\circ\tau_i))}) \geq \varepsilon_0 \geq 1,$$
by the compactness of $[0,1]$. From this choice we get
$$\ln(\varepsilon_0 \sum_{i=0}^{d-1} f\circ\tau_i) \geq \sum_{i=0}^{d-1} \ln(\varepsilon_0 f\circ\tau_i). \qquad (2)$$

Using (2) in (1) we obtain

$$I \geq \sum_{i=0}^{d-1} \int \ln(\varepsilon_0 f \circ \tau_i) d\hat{\nu} - \int \ln(\varepsilon_0 f) d\hat{\nu}.$$

Moreover, using the inequality from Lemma 7.3 applied to the function $\ln(\varepsilon f)$ (note that $\ln(\varepsilon_0 f) \in \mathbb{B}^+$, because $\varepsilon_0 \geq 1$), we get

$$I \geq \int \ln(\varepsilon f) d\hat{\nu} - \int \ln(\varepsilon f) d\hat{\nu} = 0.$$

□

Definition 7.4. Given $\phi \in \mathbb{B}^+$ we define the Topological Pressure of ϕ by

$$p(\phi) = \sup_{\hat{\nu} \in \mathcal{H}} \{h(\hat{\nu}) + \int \ln(\phi) d\hat{\nu}\}.$$

Remark 2. From Remark 1 it follows that

$$\sup_{\hat{\nu} \in \mathcal{H}} \{h(\hat{\nu}) + \int \ln(\phi) d\hat{\nu}\} =$$

$$\sup_{\hat{\nu} \text{ invariant for } \hat{\sigma}} \{h(\hat{\nu}) + \int \ln(\phi) d\hat{\nu}\}.$$

Using the formula for entropy we get a characterization of topological pressure as

$$p(\phi) = \sup_{\hat{\nu} \in \mathcal{H}} \{ \inf_{f \in \mathbb{B}^+} \int \ln(\frac{P_\phi f}{\phi f}) d\hat{\nu} + \int \ln(\phi) d\hat{\nu}\} = \sup_{\hat{\nu} \in \mathcal{H}} \inf_{f \in \mathbb{B}^+} \int \ln(\frac{P_\phi f}{f}) d\hat{\nu}.$$

In this way the pressure can be seen as a minimax problem.

Definition 7.5. A holonomic measure $\hat{\nu}_{eq}$ such that

$$p(\phi) = h(\hat{\nu}_{eq}) + \int \ln(\phi) d\hat{\nu}_{eq}$$

will be called an equilibrium state for ϕ.

Remark 3. Example 3 bellow shows that in IFS there exist examples of holonomic equilibrium states for ϕ which are not invariant for $\hat{\sigma}$.

In the next theorem we do not assume $\sum_{i=0}^{d-1} \phi(\tau_i(x)) = 1$.

Theorem 7.6. *Let us consider $\phi \in \mathbb{B}^+$ such that $([0,1], \tau_i, \phi)$ is a ρ-weighted system, for some $\rho \geq 0$. Then, $p(\phi) = \ln(\rho)$. In particular, the transference operator P_ϕ has a unique positive eigenvalue.*

Proof. Note that,

$$h(\hat{\nu}) = \inf_{f \in \mathbb{B}^+} \int \ln(\frac{P_\phi f}{\phi f}) d\hat{\nu} \leq$$

$$\int \ln(\frac{P_\phi h}{\phi h}) d\hat{\nu} = \int \ln(\frac{\rho}{\phi}) d\hat{\nu} = -\int \ln(\phi) d\hat{\nu} + \ln(\rho),$$

so,

$$h(\hat{\nu}) + \int \ln(\phi) d\hat{\nu} \leq \ln(\rho), \ \forall \hat{\nu},$$

thus, $p(\phi) \leq \ln(\rho)$.

Remember that

$$p(\phi) = \sup_{\hat{\nu} \in \mathcal{H}} \inf_{f \in \mathbb{B}^+} \int \ln(\frac{P_\phi f}{f}) d\hat{\nu}.$$

Let $\hat{\nu}_0$ be a fixed holonomic probability such that the normalized dual operator verifies $P_u^*(\pi_* \hat{\nu}_0) = \pi_* \hat{\nu}_0$ (always there exists if P_u is the normalization of P_ϕ), were $\pi(x, \omega) = x$. Thus we can write

$$p(\phi) \geq \inf_{f \in \mathbb{B}^+} \int \ln(\frac{P_\phi f}{f}) d\hat{\nu}_0.$$

Note that, from the normalization property we get

$$P_\phi(f) = P_u(\frac{f}{h})\rho h, \quad \forall f.$$

Moreover, we know that $\ln(P_u g) \geq P_u \ln(g), \forall g$, by concavity of logarithmic function.

Now, considering an arbitrary $f \in \mathbb{B}^+$, we get

$$\int \ln(\frac{P_\phi f}{f}) d\hat{\nu}_0 = \int \ln(\frac{P_u(f/h)\rho h}{f}) d\hat{\nu}_0 = \int \ln(\frac{P_u(f/h)}{f/h}) d\hat{\nu}_0 + \ln(\rho) \geq$$

$$\geq \int P_u \ln(f/h) d\hat{\nu}_0 - \int \ln(f/h) d\hat{\nu}_0 + \ln(\rho) = \ln(\rho).$$

So, $\inf_{f \in \mathbb{B}^+} \int \ln(\frac{P_\phi f}{f}) d\hat{\nu}_0 \geq \ln(\rho)$, that is, $p(\phi) \geq \ln(\rho)$. From this we get $p(\phi) = \ln(\rho)$.

In order to obtain the second part of the claim it is enough to see that $p(\phi) = \ln(\rho)$, for all ρ, thus the eigenvalue is unique. □

Theorem 7.7. *(Variational principle) Consider $\phi \in \mathbb{B}^+$ such that $([0,1], \tau_i, \phi)$ is a ρ-weighted system, for $\rho = e^{p(\phi)} \geq 0$. Then, any holonomic probability $\hat{\nu}_0$ such that the normalized dual operator verifies $P_u^*(\pi_* \hat{\nu}_0) = \pi_* \hat{\nu}_0$ is an equilibrium state.*

Proof. Note that, from the definition of pressure, we get $\ln(\rho) = p(\phi) \geq h(\hat{\nu}_0) + \int \ln(\phi) d\hat{\nu}_0$. As, $P_u^*(\pi_* \hat{\nu}_0) = \pi_* \hat{\nu}_0$, for an arbitrary $f \in \mathbb{B}^+$, we obtain

$$h(\hat{\nu}_0) + \int \ln(\phi) d\hat{\nu}_0 = \int \ln(\frac{P_\phi f}{f}) d\hat{\nu}_0 \geq \ln(\rho).$$

Thus, $\ln(\rho) = h(\hat{\nu}_0) + \int \ln(\phi) d\hat{\nu}_0$.

□

Note that, if $\hat{\nu}_0$ is a equilibrium state, such that the IFS with probabilities $([0,1], \tau_i, v_i)$ associated to $\hat{\nu}_0$ by holonomic disintegration, is uniform, then $P_u^*(\pi_* \hat{\nu}_0) = \pi_* \hat{\nu}_0$.

In fact, we know that $\ln(\rho) = h(\hat{\nu}_0) + \int \ln(\phi) d\hat{\nu}_0$, and $P_v^*(\pi_* \hat{\nu}_0) = \pi_* \hat{\nu}_0$. Then, we can write

$$\ln(\rho) = h(\hat{\nu}_0) + \int P_v \ln(\phi) d\hat{\nu}_0 = h(\hat{\nu}_0) + \int \sum_{i=0}^{d-1} v_i \ln(\phi(\tau_i)) d\hat{\nu}_0.$$

Remember that the normalization of ϕ is given by

$$u_i(x) = \frac{\phi(\tau_i(x)) h(\tau_i(x))}{\rho h(x)},$$

thus,

$$0 = h(\hat{\nu}_0) + \int \sum_{i=0}^{d-1} v_i \ln(u_i) d\hat{\nu}_0. \qquad (3)$$

As $([0,1], \tau_i, v_i)$ is uniform, there exists a weight function ψ such that $v_i(x) = \phi(\tau_i(x))$, $\forall i = 0, ..., d-1$. Moreover, $p(\psi) = 0$ and $\hat{\nu}_0$ is clearly a equilibrium state for ψ. Thus

$$0 = h(\hat{\nu}_0) + \int \ln(\psi) d\hat{\nu}_0$$

Using $P_v^*(\pi_* \hat{\nu}_0) = \pi_* \hat{\nu}_0$ we get

$$0 = h(\hat{\nu}_0) + \int \sum_{i=0}^{d-1} v_i \ln(v_i) d\hat{\nu}_0. \qquad (4)$$

It is well known that

$$-\sum_{i=0}^{d-1} a_i \ln(a_i) + \sum_{i=0}^{d-1} a_i \ln(b_i) \leq 0. \qquad (5)$$

where $\sum_{i=0}^{d-1} a_i = 1 = \sum_{i=0}^{d-1} b_i$ and $b_i \geq 0$, with equality only if $a_i = b_i$ (see [23] for a proof). From (3) and (4) we get,

$$u_i(x) = v_i(x), \quad \nu - a.e.$$

Then, it follows that $P_u^*(\pi_* \hat{\nu}_0) = \pi_* \hat{\nu}_0$.

Examples:

1) For $\phi = 1$, we have $P_\phi(1) = d \cdot 1$. Then, for all equilibrium states $\hat{\nu}_{eq}$ we get

$$\ln(d) = h(\hat{\nu}_{eq}) + \int \ln(1) d\hat{\nu}_{eq} = \sup_{\hat{\nu} \in \mathcal{H}} \inf_{f \in \mathbb{B}^+} \int \ln(\frac{P_1 f}{f}) d\hat{\nu} = \sup_{\hat{\nu} \in \mathcal{H}} h(\hat{\nu})$$

2) If $P_\phi(1) = 1$, that is, the case of IFS with probabilities, then $p(\phi) = 0$. Therefore, $h(\hat{\nu}) + \int \ln(\phi) d\hat{\nu} \leq 0$, for all holonomic probabilities. Moreover, any equilibrium state $\hat{\nu}_{eq}$ satisfies

$$h(\hat{\nu}_{eq}) = -\int \ln(\phi) d\hat{\nu}_{eq}$$

3) Consider the IFS given by $([0,1], \tau_0(x) = x, \tau_1(x) = 1 - x)$ and the potential $\phi(x) = 2 + \cos(2\pi x)$. Is clear that

$$0 \leq \int \ln(\phi) d\nu \leq \ln 3, \ \forall \nu.$$

We will consider a especial holonomic probability $\hat{\nu}_0$ constructed in the following way (similar to the one presented in section 4):

Consider fixed $\bar{w} \neq (111111...) \in \Sigma$ and $x_0 = 0$. The holonomic probability $\hat{\nu}_0$ is the average of delta of Dirac distributions at the points $(x_0, 11\bar{w})$ and $\hat{\sigma}(x_0, 11\bar{w})$ in $[0,1] \times \Sigma$, more precisely, for any g

$$\int g(x,w) d\hat{\nu} = \frac{1}{2} \sum_{j=0}^{1} g(Z_j(x_0, \bar{w}), \sigma^j(\bar{w})) = \frac{1}{2}(g(0, 11\bar{w}) + g(1, 1\bar{w}))$$

This probability is not $\hat{\sigma}$-invariant by construction, and has the interesting property:

$$\ln(2) = h(\hat{\nu}_0) = \sup_{\hat{\nu}\in\mathcal{H}} h(\hat{\nu})$$

Indeed, since $h(\hat{\nu}) \leq \ln(2), \forall \hat{\nu}$, it is enough to see that $h(\hat{\nu}_0) = \ln(2)$. Remember that $h(\hat{\nu}_0) = \inf_{f\in\mathbb{B}^+} \int \ln(\frac{P_1 f}{f}) d\hat{\nu}_0$, so

$$h(\hat{\nu}_0) = \inf_{f\in\mathbb{B}^+} \frac{1}{2}(\log\frac{P_1 f(0)}{f(0)} + \log\frac{P_1 f(1)}{f(1)}) = \inf_{f\in\mathbb{B}^+} \frac{1}{2}[\ln(1+\frac{f(0)}{f(1)}) + \ln(1+\frac{f(1)}{f(0)})]$$

Taking $\frac{f(0)}{f(1)} = \lambda > 0$, we get

$$h(\hat{\nu}_0) = \inf_{\lambda>0} \frac{1}{2}[\ln(1+\lambda) + \ln(1+\frac{1}{\lambda})] = \ln 2$$

Remark 4. This shows that such $\hat{\nu}_0$ is a maximal entropy holonomic probability which is not $\hat{\sigma}$-invariant. This also shows that the holonomic setting can not be reduced, via coding, to the analysis of $\hat{\sigma}$-invariant probabilities in a symbolic space. Otherwise, in the symbolic case a probability with support in two points would have positive entropy.

Also, one can calculate $m = \sup_{\hat{\nu}\in\mathcal{H}} \int \ln(\phi) d\hat{\nu} \leq \ln 3$. We claim that $m = \ln 3$. In fact,

$$\int \ln(\phi) d\hat{\nu}_0 = \frac{1}{2}[\ln(2+\cos(2\pi 0)) + \ln(2+\cos(2\pi 1))] = \ln 3.$$

Finally, we point out that $\hat{\nu}_0$ is also a equilibrium state. Indeed,

$$p(\phi) = \sup_{\hat{\nu}\in\mathcal{H}} \{h(\hat{\nu}) + \int \ln(\phi) d\hat{\nu}\} \leq \ln 2 + \ln 3.$$

As, $h(\hat{\nu}_0) + \int \ln(\phi) d\hat{\nu}_0 = \ln 2 + \ln 3$, we get

$$\ln(6) = p(\phi) = h(\hat{\nu}_0) + \int \ln(\phi) d\hat{\nu}_0.$$

From this example one can see that there exists equilibrium states which are not $\hat{\sigma}$-invariant probabilities.

Definition 7.8. Two functions $\psi_1, \psi_2 \in \mathbb{B}^+$ will be called holonomic-equivalent (or, co-homologous) if there exists a function $h \in \mathbb{B}^+$, such that

$$\psi_1(x) = \psi_2(x) \cdot \frac{h(\tau_{X_1(w)}(x))}{h(x)}, \ \forall (x, \omega).$$

It is clear that, two holonomic equivalent potentials ψ_1, ψ_2 will have the same equilibrium states.

8. **An alternative point of view for the concept of entropy and pressure for IFS.**

Definition 8.1. Given $\hat{\nu} \in \mathcal{H}$ and let $([0,1], \tau_i, v_i)$ be the IFS with probabilities arising in the holonomic disintegration of $\hat{\nu}$ (see Theorem 3.4). We can also define the Entropy of $\hat{\nu}$ by

$$h(\hat{\nu}) = - \sup_{\sum_{i=0}^{d-1} u_i = 1} \int \sum_{i=0}^{d-1} v_i \ln(u_i) d\hat{\nu}.$$

Proposition 3. *Consider $\hat{\nu} \in \mathcal{H}$. Then,*

$$0 \leq h(\hat{\nu}) \leq \ln(d).$$

Proof. Firstly consider $u_i^0 = \frac{1}{d}$, $i = 0, ..., d-1$ then $\sum_{i=0}^{d-1} u_i^0 = 1$ and

$$-h(\hat{\nu}) = \sup_{\sum_{i=0}^{d-1} u_i = 1} \int \sum_{i=0}^{d-1} v_i \ln(u_i) d\hat{\nu} \geq \int \sum_{i=0}^{d-1} v_i \ln(\frac{1}{d}) d\hat{\nu} = -\ln(d),$$

so, $h(\hat{\nu}) \leq \ln(d)$.

On the other hand, $u_i \leq 1$ so $\ln(u_i) \leq 0$, and then

$$\sup_{\sum_{i=0}^{d-1} u_i = 1} \int \sum_{i=0}^{d-1} v_i \ln(u_i) d\hat{\nu} \leq 0.$$

Thus, $0 \leq h(\hat{\nu})$. □

Lemma 8.2. *(Existence of equilibrium states) Consider $\phi \in \mathbb{B}^+$ such that $([0,1], \tau_i, \phi)$ is a ρ-weighted system, for some $\rho > 0$ (remember that there exists $h > 0$, such that $P_\phi(h) = \rho h$). Denote P_v the normalization of P_ϕ, that is, $([0,1], \tau_i, v_i)$ is a 1-weighted system, such that $P_v(1) = 1$ and $P_v^*(\nu) = \nu$. Let $\hat{\nu}$ be the holonomic lifting of ν. Then,*

$$h(\hat{\nu}) + \int \ln(\phi) d\hat{\nu} = \ln(\rho)$$

Proof. Let $\hat{\nu}$ be the holonomic lifting of ν. By Proposition 1, we know that the 1-weighted system associated to its holonomic disintegration is $([0,1], \tau_i, v_i)$. Then, from the definition of entropy we get

$$h(\hat{\nu}) = -\sup_{\sum_{i=0}^{d-1} u_i = 1} \int \sum_{i=0}^{d-1} v_i \ln(u_i) d\hat{\nu} = -\int \sum_{i=0}^{d-1} v_i \ln(v_i) d\hat{\nu}$$

from the logarithmic inequality (5) above.

Remember that the normalization of ϕ is given by

$$v_i(x) = \frac{\phi(\tau_i(x)) h(\tau_i(x))}{\rho h(x)}.$$

Replacing this expression in the equation for entropy we get

$$h(\hat{\nu}) = -\int \sum_{i=0}^{d-1} v_i \ln(\frac{\phi(\tau_i(x)) h(\tau_i(x))}{\rho h(x)}) d\hat{\nu} =$$

$$= -\int \sum_{i=0}^{d-1} v_i \ln(\phi(\tau_i(x)) d\hat{\nu} - \int \sum_{i=0}^{d-1} v_i \ln(\frac{h(\tau_i(x))}{h(x)}) d\hat{\nu} + \ln(\rho) =$$

$$= -\int \ln(\phi(x)) d\hat{\nu} + \ln(\rho)$$

□

Now, we use the concept introduced in the present section.

Definition 8.3. Given $\phi \in \mathbb{B}^+$, we define the Topological Pressure of ϕ by

$$p(\phi) = \sup_{\hat{\nu} \in \mathcal{H}} \{h(\hat{\nu}) + \int \ln(\phi) d\hat{\nu}\}$$

Theorem 8.4. *Let us consider $\phi \in \mathbb{B}^+$ such that $([0,1], \tau_i, \phi)$ is a ρ-weighted system, for some $\rho \geq 0$. Then $p(\phi) = \ln(\rho)$. In particular, the transference operator P_ϕ has a unique positive eigenvalue.*

Proof. Let P_u be the normalization of P_ϕ. Then,

$$h(\hat{\nu}) + \int \ln(\phi) d\hat{\nu} = h(\hat{\nu}) + \int \ln(\phi) d\hat{\nu} =$$

$$= -\int \sum_{i=0}^{d-1} v_i \ln(v_i) d\hat{\nu} + \int P_v \ln(\phi) d\hat{\nu} =$$

$$= -\int \sum_{i=0}^{d-1} v_i \ln(v_i) d\hat{\nu} + \int \sum_{i=0}^{d-1} v_i \ln(\phi(\tau_i)) d\hat{\nu} =$$

$$= -\int \sum_{i=0}^{d-1} v_i \ln(v_i) d\hat{\nu} + \int \sum_{i=0}^{d-1} v_i \ln(\frac{\phi(\tau_i) h(\tau_i)}{\rho h} \frac{\rho h}{h(\tau_i)}) d\hat{\nu} =$$

$$= -\int \sum_{i=0}^{d-1} v_i \ln(v_i) d\hat{\nu} + \int \sum_{i=0}^{d-1} v_i \ln(u_i \frac{\rho h}{h(\tau_i)}) d\hat{\nu} =$$

$$= \int -\sum_{i=0}^{d-1} v_i \ln(v_i) + \sum_{i=0}^{d-1} v_i \ln(u_i) d\hat{\nu} + \int \sum_{i=0}^{d-1} v_i \ln(\frac{\rho h}{h(\tau_i)}) d\hat{\nu} \leq$$

$$= \ln(\rho) + \int \sum_{i=0}^{d-1} v_i \ln(h) d\hat{\nu} - \int \sum_{i=0}^{d-1} v_i \ln(h(\tau_i)) d\hat{\nu} =$$

$$= \ln(\rho) + \int \ln(h) d\hat{\nu} - \int P_v \ln(h) d\hat{\nu} = \ln(\rho)$$

The equality follows from the Lemma 3 □

From Theorem 8.4 and Lemma 3, it follows that there exists equilibrium states, more precisely, given a ρ-weighted system, all holonomic liftings of the normalized probability, are equilibrium states.

The Variational principle in the formulation of the present section is stronger than the formulated in the first part. The change in the definition of entropy allow us to get a characterization of the equilibrium states as holonomic liftings of the P_u-invariant probabilities of the normalized transference operator. This point will become clear in the proof (of the "if, and only if," part) of the next theorem.

Theorem 8.5. *(Alternative Variational principle) Let us consider $\phi \in \mathbb{B}^+$ such that $([0,1], \tau_i, \phi)$ is a ρ-weighted system, for $\rho = e^{p(\phi)} \geq 0$. Then, a holonomic probability $\hat{\nu}_0$ is a equilibrium state, if and only if, the projection $\hat{\nu}_0$ by disintegration, is invariant for the normalized dual operator of P_ϕ(that is, $P_u^*(\pi_* \hat{\nu}_0) = \pi_* \hat{\nu}_0$).*

Proof. By Lemma 3, it follows that: if $\hat{\nu}_0$ is the holonomic lifting of the normalized probability $\pi_* \hat{\nu}_0$, then $\hat{\nu}_0$ is an equilibrium state.

The converse is also true. In fact, suppose that $\hat{\nu}_0$ is a equilibrium state, that is,

$$h(\hat{\nu}_0) + \int \ln(\phi) d\hat{\nu}_0 = \ln(\rho)$$

Using the normalization we get,

$$-\int\sum_{i=0}^{d-1} v_i\ln(v_i)d\hat{\nu}_0+\int\ln(\phi)d\hat{\nu}_0=\ln(\rho)$$

where $([0,1],\tau_i,v_i)$ is the 1-weighted system associated to holonomic disintegration of $\hat{\nu}_0$. From the invariance of P_v we have

$$-\int\sum_{i=0}^{d-1} v_i\ln(v_i)d\hat{\nu}_0+\int\sum_{i=0}^{d-1} v_i\ln(\phi(\tau_i))d\hat{\nu}_0=\ln(\rho)$$

Finally, from the relations of normalization P_u of P_ϕ

$$u_i(x)=\frac{\phi(\tau_i(x))h(\tau_i(x))}{\rho h(x)}\Leftrightarrow \phi(\tau_i(x))=\frac{u_i(x)\rho h(x)}{h(\tau_i(x))},$$

we get

$$-\int\sum_{i=0}^{d-1} v_i\ln(v_i)d\hat{\nu}_0+\int\sum_{i=0}^{d-1} v_i\ln(\frac{u_i(x)\rho h(x)}{h(\tau_i(x))})d\hat{\nu}_0=\ln(\rho).$$

This is equivalent to

$$\int\sum_{i=0}^{d-1} v_i\ln(\frac{u_i}{v_i})d\hat{\nu}_0=0.$$

From,

$$\sum_{i=0}^{d-1} v_i\ln(\frac{u_i}{v_i})\leq\ln(\sum_{i=0}^{d-1} v_i\frac{u_i}{v_i})=\ln(\sum_{i=0}^{d-1} u_i)=\ln(1)=0,$$

it follows that $u_i=v_i,\quad \pi_*\hat{\nu}_0-a.e.$

As, $\pi_*\hat{\nu}_0$ is P_v-invariant, we get $P_u^*(\pi_*\hat{\nu}_0)=\pi_*\hat{\nu}_0$

□

9. Maximizing probabilities for IFS.

Definition 9.1. Given a continuous potential $A:\hat{\Sigma}\to\mathbb{R}$, we denote by $m(A)$ the maximal value of the integral of A in $\hat{\sigma}$-holonomic probabilities

$$m(A)=\sup\{\int A(x,w)\,d\hat{\nu}(x,w)\,|\,\hat{\nu}\in\mathcal{H}\}$$

We also denote by $\mu_{\infty,A}=\mu_\infty$ any holonomic probability which attains the supremum value and call it a maximizing probability for A.

Remark 5: If we consider a potential A such that has supremum exactly on the support of a holonomic probability, then this probability will be maximizing.

We refer the reader to [13] for the similar problem when one considers potentials A defined on the Bernouilly space $\{1,2,,,d\}^{\mathbb{Z}}$.

As an example consider a positive function $\phi(x,w)=\phi(x)$ and $A=\log\phi$. For ϕ fixed (and so A fixed) one can consider a real parameter β and the problem

$$p(\phi^\beta)=\sup_{\hat{\nu}\in\mathcal{H}}\{h(\hat{\nu})+\beta\int\ln(\phi)d\hat{\nu}\}=\sup_{\hat{\nu}\in\mathcal{H}}\{h(\hat{\nu})+\beta\,Ad\hat{\nu}\}.$$

For each value β, denote by $\hat{\nu}_\beta$ a solution (therefore, normalized) of the above variational problem. Consider the limit problem when $\beta\to\infty$. Any subsequence (weak limit) $\hat{\nu}_{\beta_n}\to\nu$, $\beta_n\to\infty$ will determine a maximizing holonomic probability ν (in the sense, of maximizing $\sup_{\hat{\nu}\in\mathcal{H}}\{\int\ln(\phi)d\hat{\nu}\,\}$) because the entropy of any holonomic probability is bounded by $\ln d$. We refer the reader to [13] for properties on maximizing holonomic probabilities and we point out that these results apply also for the iterated setting as described above in the first two sections.

Maximizing probabilities correspond to the limit of β-Gibbs states, $\beta \to \infty$, and they are also called of Gibbs states at temperature zero. Large Deviations results for the limit zero temperature case of σ-invariant (not holonomic) Gibbs states are described in [2].

In the setting of C^*-algebras it corresponds to the ground state [11].

Now we return to consider the case of $A(x, w)$ depending on both variables x and w. We will describe bellow the dual problem and in this setting will be clear that is quite natural to consider the concept of holonomic probability.

Theorem 9.2. *Given a potential $A \in C^0(\hat{\Sigma})$, then*

$$m(A) = \inf_{f\in[0,1]} \{ \max_{(x,w)\in\hat{\Sigma}} [A(x, w) + f(x) - f(\tau_{X_1(w)}(x))] \}.$$

Proof:

We will outline the proof which is similar to the one in [13].

Remark 6: Note that the right side of the equality above does not mention any holonomic concept. This expression correspond in classical mechanics to what is called the effective hamiltonian (see (5.1) and (5.7) in [10]). Here we consider discrete time difference equations instead od derivative as in classical mechanics. The holonomic condition corresponds to (1.3) in [10] and to (1.2) in [14]. The holonomic maximizing probability corresponds to the semiclassical measure in [10] and to the Aubry-Mather measure in [14].

First consider the convex function $F : C^0(\hat{\Sigma}) \to \mathbb{R}$ taking $F(g) = \max(A + g)$. Consider also the subset

$$\mathcal{C} = \{g \in C^0(\hat{\Sigma}) : g(x, w) = f(x) - f(\tau_{X_1(w)}(x)), \text{ for some } f \in C^0(\Sigma)\}.$$

We also define a concave function $G : C^0(\hat{\Sigma}) \to \mathbb{R} \cup \{-\infty\}$ taking $G(g) = 0$, if $g \in \bar{\mathcal{C}}$, and taking $G(g) = -\infty$ otherwise.

Let $\mathcal{M}$ be the convex set of probabilities over the sigma-algebra of Borel in $\hat{\Sigma}$ (with the usual metric).

Let $\mathcal{S}$ be the set of the signed measures on the Borel sigma-algebra of $\hat{\Sigma}$.

The corresponding Fenchel transforms, $F^* : \mathcal{S} \to \mathbb{R} \cup \{+\infty\}$ and $G^* : \mathcal{S} \to \mathbb{R} \cup \{-\infty\}$, are given by

$$F^*(\hat{\mu}) = \sup_{g\in C^0(\hat{\Sigma})} \left[\int_{\hat{\Sigma}} g(x, w)\, d\hat{\mu}(x, w) - F(g) \right] \text{ and}$$

$$G^*(\hat{\mu}) = \inf_{g\in C^0(\hat{\Sigma})} \left[\int_{\hat{\Sigma}} g(x, w)\, d\hat{\mu}(x, w) - G(g) \right].$$

Denote

$$\mathcal{S}_0 = \left\{ \hat{\mu} \in \mathcal{S} : \int_{\hat{\Sigma}} f(\tau_{X_1(w)}(x))\, d\hat{\mu}(x, w) = \int_{\hat{\Sigma}} f(x)\, d\hat{\mu}(x, w) \,,\ \forall\, f \in C^0([0, 1]) \right\}.$$

Proof. First we need a lemma (its proof is quite similar to the one in [13]).

Lemma 9.3. *: Given F and G as above*

$$F^*(\hat{\mu}) = \begin{cases} -\int_{\hat{\Sigma}} A(x,w)\, d\hat{\mu}(x,w) & , \textit{if } \hat{\mu} \in \mathcal{M}, \\ +\infty & , \textit{in the other case,} \end{cases}$$

$$G^*(\hat{\mu}) = \begin{cases} 0 & , \textit{ if } \hat{\mu} \in \mathcal{S}_0, \\ -\infty & , \textit{ in the other case.} \end{cases}$$

The claim of the theorem is a consequence of the Fenchel-Rockafellar theorem:

$$\sup_{g \in C^0(\hat{\Sigma})} [G(g) - F(g)] = \inf_{\hat{\mu} \in \mathcal{S}} [F^*(\hat{\mu}) - G^*(\hat{\mu})] \,.$$

In this way, by lemma 2,

$$\sup_{g \in \mathcal{C}} \left[- \max_{(x,w) \in \hat{\Sigma}} (A+g)(x,w) \right] = \inf_{\hat{\mu} \in \mathcal{H}} \left[- \int_{\hat{\Sigma}} A(x,w)\, d\hat{\mu}(x,w) \right].$$

Finally, using the definition of the set $\mathcal{C}$ we finish the proof of the theorem.

□

Remark 7: A function f_* which attains such minimum value exists and it is special. It satisfies the sub-cohomological equation: $\forall (x,w) \in [0,1] \times \Sigma$

$$m(A) \geq A(x,w) + f_*(x) - f_*(\tau_{X_1(w)}(x)).$$

A function f_* which satisfies this equation is called sub-action.

For such f_*, in the support of any maximizing holonomic probability μ_∞ (see proposition 8 in [13]), one gets the equality

$$m(A) = A(x,w) + f_*(x) - f_*(\tau_{X_1(w)}(x)).$$

Sub-actions are important because, among other things, they bring information which help us to locate the support of the maximizing probability.

This expression corresponds, in discrete time IFS dynamics, to the well known Hamilton-Jacobi equation (5.3) in [10] and Theorem 5.4 in [14] (the equality is true in the support of the holonomic probability).

The holonomic condition appeared initially on the setting of Classical Mechanics and in the search for a duality expression as the one described by the last theorem.

Therefore, if (eventually) the maximum of a potential A is exactly on the support of a holonomic probability (see remark 5), then the properties described in our paper are general enough to describe (via subactions) this set. In other words, in this case, to consider the alternative problem of maximization (of the integral of A) on σ-invariant probabilities would not result in so precise information.

The bottom line is: we described above a pressure problem for holonomic probabilities (which takes care of the finite temperature case) and also the limit case (temperature zero). If the holonomic maximizing probability is unique, then similar results, like in [2], can be easily extended to the present situation.

Acknowledgements. We would like to thanks the referee for their valuable comments and suggestions.

REFERENCES

[1] A. L. Ambrosio, N. Gigli and G. Savare *"Gradient Flows in Metric Spaces and in the Space of Probability Measures"*, ETH Birkhauser Verlag, Basel, 2005.

[2] A. Baraviera, A. O. Lopes and P. Thieullen, *A large deviation principle for equilibrium states of Holder potentials: the zero temperature case*, Stochastics and Dynamics, **6** (2006), 77–96.

[3] M. F. Barnsley, S. Demko, J. Elton and J. Geronimo, *Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities*, Ann. Inst. H. Poincaré - Probab. Statist., 24 **3**. (1988), 367–394.

[4] G. G. de Castro and A. O. Lopes, *KMS States, Entropy and a Variational Principle for Pressure*, preprint (2008)

[5] J.-P Conze and A. Raugi, *Fonctions harmoniques pour un operateur de transition e applications*, Bull. Soc. Math. France, 118 **4** (1990), 273–310.

[6] C. Dellacherie, *"Probabilities and potential"*, Amsterdam: North-Holland, 1978. p.189.

[7] D. Dutkay and P. Jorgensen, *Disintegration of projective measures.*, Proc. Amer. Math. Soc., 135 **1** (2007), 169–179.

[8] D. Dutkay and P. Jorgensen, *Iterated function systems, Ruelle operators, and invariant projective measures*, Math. Comp., 75 **256** (2006), 1931–1970

[9] J. Elton, *An ergodic theorem for iterated maps*, Ergodic Theory and Dyn. Systems 7 **4** (1987), 481–488.

[10] L. C. Evans, *Towards a quantum analog of weak KAM theory*, Comm. in Math. Physics, **244** (2004), 311–334

[11] R. Exel and A. O. Lopes, *C^*-Algebras, approximately proper equivalence relations, and Thermodynamic Formalism*, Ergodic Theory and Dyn. Systems, 24 **4**, (2004) pp 1051–1082

[12] A. H. Fan, K-S. Lau, *Iterated function system and Ruelle operator*, J. Math. Anal. Appl., 231 **2** (1999), 319–344.

[13] E. Garibaldi and A. O. Lopes, *On the Aubry-Mather theory for symbolic dynamics*, Ergodic Theory and Dynamical Systems, Vol 28 **3** (2008), 791–815

[14] D. Gomes, *Viscosity solution method and the discrete Aubry-Mather problem*, Discrete and Continuous Dynamical Systems - A, **13** (2005), 103–116.

[15] J. Jaroszewska, *Iterated function systems with continuous place dependent probabilities*, Univ. Iag. Acta Math, **XL**, (2002) 137–146

[16] J. Jaroszewska and M. Rams, *On the Hausdorff dimension of invariant measures of weakly contracting on average measurable IFS*, preprint (2007)

[17] P. Jorgensen, *"Analysis and Probability: wavelets, signals, fractals"*, Springer Verlag, 2006

[18] M. Keane, *Strongly mixing g-measures*, Invent. Math., **16** (1972), 309–324.

[19] A. O. Lopes, *An analogy of the charge distribution on Julia sets with the Brownian motion*, J. Math. Phys., 30 **9** (1989), 2120–2124.

[20] M. Mauldin and M. Urbanski, *"Graph Directed Markov Systems: Geometry and Dynamics of Limit Sets"*, 2003, Cambridge Press

[21] E. R. Oliveira, *Propriedades genéricas de lagrangianos e princípios variacionais para probabilidades holonômicas em IFS*, PhD thesis, Universidade Federal do Rio Grande do Sul, 2007.

[22] L. Olsen, *"Random Geometrically Graph Directed Self-similar Multifractals"*, 1994, Longman Group

[23] W. Parry and M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérsque, **187-188** (1990), 1–268.

[24] O. Stenflo, *Uniqueness of invariant measures for place-dependent random iterations of functions*, Fractals in multimedia (Minneapolis, MN, 2001), 13–32, IMA Vol. Math. Appl., 132, Springer, New York, 2002.

[25] L. Schwartz, *"Radon measures in arbitrary topological spaces and cylindrical measures"*, Oxford University Press, Londres, 1973.

[26] O. Stenflo, *A note on a theorem of Karlin*, Statist. Probab. Lett., 54 **2** (2001), 183–187.

[27] M. Urbanski, *Hausdorff measures versus equilibrium states of conformal infinite iterated function systems*, International Conference on Dimension and Dynamics (Miskolc, 1998), Period. Math. Hungar., 37 **1-3**. (1998), 153–205.

E-mail address: arturoscar.lopes@gmail.com

E-mail address: oliveira.elismar@gmail.com