ON THE ERGODIC THEORY OF DISCRETE DYNAMICAL SYSTEMS

A. CANDEL

Abstract. In this paper we study a class of measures, called harmonic measures, that one can associate to a dynamical system consisting og a space X and a finitely generated group of transformations. If the group of transformations is infinite cyclic, then these measures are invariant. We show how the theory of classical dynamical systems with invariant measure can be extended to the case of harmonic measure. Other properties of harmonic measures are best understood in the light of the harmonic analysis on a free group.

Contents

1. Introduction 1
2. Discrete potential theory 2
3. Discrete dynamical systems 6
4. Markov operators and random walks 7
5. The ergodic theorem 11
6. The Jacobian cocycle and the tautological action 17
7. Ergodicity and mixing 25
8. Topological entropy 31
9. On the variational principle 37
References 40

1. Introduction

Ergodic theory deals with measurable actions of groups of transformations. When the action is generated by a single measure preserving transformation then the basic theory is well developed and understood. In this paper we explore the situation of dynamical systems with more than one generator which do not necessarily admit an invariant measure. The basic idea is to consider the concept of harmonic measure, that is, a measure whose averages are invariant, and study its basic properties by comparison with the ergodic theory of a single transformation.

The concept of harmonic measure studied here is the discrete version of the harmonic measures for foliations introduced by L. Garnett [11]. In fact, one of the motivations of this paper was to understand how to use harmonic measures for foliations defined by actions of semisimple Lie groups.

We start by providing an existence theorem for harmonic measures on dynamical systems constituted by a compact metric space and a finitely generated group of transformations. This group can always be taken to be an abstract free group, and indeed a free group acting on its natural boundary provides the main model for an action with ergodic harmonic measure.

The main tool for understanding a harmonic measure is the construction of the path space of the action, similar in spirit to Wiener space of a manifold. The points of this space are paths on the orbits of the original state space and carries with it a shift transformation. This path space sits over the state space and the harmonic measures lift to shift invariant measures on it, so that in a sense there is a classical system overlying a non–classical one. This fact allows us to interpret properties of our system in terms of those of the path space.

For example, ergodicity of a harmonic measure is equivalent to ergodicity of the induced measure in path space. We study the related concept of mixing, which is defined with respect to the invariant measures in the path space and then translated back to the harmonic measures. It turns out that a harmonic measure which is not invariant has stronger ergodicity properties than the classical invariant measures. All these properties have also interpretations in terms of the unitary representation associated to the action and the harmonic measure.

Another topic that we discuss is the ergodic decomposition of harmonic measures, following the original work of Krilloff and Bogoliugoff [17]. This was a theory called for in [22] and also [9]. Our solution involves certain probabilistic considerations, and in particular shows that harmonic measures are the natural replacement for invariant ones.

After this we discuss some properties of the jacobian cocycle of a harmonic measure, and some relations between ergodicity, mixing, and the unitary representation to a group action.

The last part of the paper is devoted to entropy. The definition of topological entropy that we consider is the natural extension of the entropy of a map, and it is then a particular case of the notion of entropy of pseudogroups of homeomorphisms studied in [13]. There is another definition of entropy that one can put together by going to the path space. We will show that this entropies are unrelated.

What is done here admits, like any other thing, generalizations. Perhaps a more proper context would have been pseudogroups of homeomorphisms or, even better, countable equivalence relations. On the other hand one has more concrete applications in the case of group actions. In any case, even when I confine the discussion to this category, I have tried to present it in a way that extends with the least number of changes to the other categories mentioned above. A reference for countable equivalence relations is [7]. In particular, note that their first theorem shows that this category is not more general that that of actions of countable groups (but of course, what is done here depends in certain sense on the structure of the orbits).

2. Discrete potential theory

In this section we recall a few concepts pertaining to discrete potential theory which will be used in the sequel.

Let K be a graph, that is, a one–dimensional connected complex. Let K_0 and K_1 denote the set of vertices and the set of edges of K, respectively. We say that two vertices x and y of K are neighbors, and write $x \sim y$, if there is an edge in K_1 connecting them. We write xy for the directed edge connecting them.

Assume that each directed edge xy has a number $q(xy) > 0$ associated to it, in such a way that $\sum_{y\sim x} q(xy) = 1$ for each vertex x. This system of weights $\{q\}$ in turn produces a diffusion (or averaging) operator D which acts on functions $f : K_0 \to \mathbb{C}$ by

$$Df(x) = \sum_{y\sim x} q(xy)f(y).$$

There is a second operator, called the Laplacian and denoted by Δ, which is defined as $\Delta f = Df - f$.

The standard Laplacian on a graph K is associated with the system of weights given by $q(xy) = \text{Card}\{y \mid y \sim x\}$.

A (real valued) function f on the set of vertices K_0 of K is called harmonic if $\Delta f = 0$ or, equivalently, if $Df = f$. The version of the weak maximum principle for this operator is as follows:

Proposition 2.1. *Let $f : K_0 \to \mathbb{R}$, where K_0 is the set of edges of a connected locally finite graph K. If x is a point of K_0 where f reaches a local maximum (respectively minimum), then $\Delta f(x) \leq 0$ (respectively ≥ 0).*

Another useful fact is the discrete version of Harnack's inequality:

Proposition 2.2. *Let f be a positive function on K_0 which satisfies $Df \leq f$. Then there is a constant C, depending only on the system of weights, such that*

$$C^{-n}f(x) \leq f(y) \leq Cf(x)$$

for any two points at distance n in the path metric of K.

The graphs that we will consider in this paper are those associated with a group G endowed with a finite symmetric system of generators G_1 in the following way. If $G_1 = \{s_1^{\pm 1}, \cdots, s_n^{\pm 1}\}$, $(s_i \neq 1)$, then a graph is constructed by taking as set of vertices the elements of G and putting and edge from s to $s_i^{\pm 1}s$. The function theory pertaining to this example has been certainly well studied, specially in recent years [4], [26].

The systems of weights that we will consider are given by probability measures q on G whose support is a symmetric generating set for G. What we do can certainly be extended to more general situations, but from the point of view of group actions this framework is general enough to understand the basic features of the theory.

Thus the weight q can be thought of as a function which assigns a positive number $q(s) > 0$ to each element s of G_1 and such that $\sum_{G_1} q(s) = 1$. In the graph of G relative to G_1 this gives a system of weights by setting $q(xy) = q(s)$ if $y = sx$, $s \in G_1$. Other pieces of nomenclature are the following. If q is a system of weights for the pair (G, G_1), then we will call the dual system to q to be that which assigns the weight $q(s^{-1})$ to s. The system of weights q is said to be symmetric if $q(s) = q(s^{-1})$, and it is equidistributed if $q(s) = 1/\text{Card } G_1$.

It will also be clear that, for actions, one needs only consider the free group generated by the symbols of G_1, a fact that simplifies all the discussions and which we take from now on. Some possible amplifications are described in the introduction.

We will utilize two compact topological spaces which are naturally associated to G. One of them is the path space of G, denoted by $\Omega(G)$, and consisting of all infinite sequences of elements of G_1, that is, $\Omega(G)$ is the infinite product G_1^∞. It comes equipped with a measure P, which is the product measure determined by the system of weights q, and with a shift transformation θ which takes an infinite path $(\cdots, w(2), w(1))$ to $(\cdots, w(2))$ and preserves the measure P. (Note that we write the paths from right to left).

Also associated to G there is a compact topological space called the boundary of G, and denoted by ∂G. It can be described as the subset of the infinite product space G_1^∞ consisting of all sequences $b = \cdots s_1 s_0$ of elements of G_1 with the proviso that $s_{n+1} \neq s_n^{-1}$. There is a natural action of G on this space: an element $s \in G_1$ acts on $b \in \partial G$ according to the rule

$$sb = s(\cdots s_1 s_0) = \begin{cases} \cdots s_1 s_0 s^{-1} & \text{if } s_0 \neq s, \\ \cdots s_2 s_1 & \text{otherwise.} \end{cases}$$

There is also a natural measure on ∂G, which will be denoted by λ. Its construction is as follows. Each element s of G determines a set in ∂G comprised by those words whose initial string is s. These sets form a base for a topology of ∂G, and the measure of the set determined by s is $q(s) = q(s_n) \cdots q(s_1)$, if $s = s_n \cdots s_1$ is the shortest expression of s in the elements of G_1.

The relation between these two spaces associated to G is summarized in the following:

Proposition 2.3. *Suppose that G_1 has more than two elements. Then almost every path w in $\Omega(G)$ hits a point* $\lim w$ *in ∂G. The hitting distribution is given by the measure λ.*

In other words, there is a map $\lim : \Omega(G) \to \partial G$ defined almost everywhere, and which takes the measure on $\Omega(G)$ to that of ∂G.

This measure λ is not invariant by G, but its transforms by G are all mutually absolutely continuous. The Radon–Nikodym derivative of this transformation is denoted by $K(s, b)$, that is

$$\int_{\partial G} \varphi(b) K(s, b) \lambda(b) = \int_{\partial G} \varphi(b) \lambda(sb).$$

Being the Radon-Nikodym derivative of a group action, this kernel K satisfies the following cocycle identity:

$$K(sr, b) = K(s, rb) K(r, b).$$

The function K is analogous to the Poisson kernel for the action of $\mathrm{PSl}(2, \mathbb{R})$ on the boundary of the unit disc in the complex plane. In the rest of this section we collect some facts concerning harmonic functions on graphs and their boundary behavior. Proofs can be found in [4]. For the graphs of a free group that we consider here, the proofs can be carried out exactly like in the case of the unit disc.

Proposition 2.4. *Let h be a positive harmonic function on G. Then there is a positive measure m on ∂G of total mass $h(e)$ such that*

$$h(s) = \int_{\partial G} K(s, b) m(b).$$

For instance, the measure λ represents the constant function 1, and the point mass measure at b represents the function $K(\cdot, b)$.

The space of positive harmonic functions on G will be denoted by $\mathcal{H}$, and endowed with the topology of pointwise convergence. As such it can be seen to be a convex cone in the topological vector space $\mathcal{H} - \mathcal{H}$. This cone has a compact base, namely the positive harmonic functions h on G such that $h(e) = 1$. This compact base is denoted by $\mathcal{H}_1$, and it is a compact and metrizable convex set. Its boundary is formed by the positive constant multiples of the minimal harmonic functions on G. These ones are precisely the functions $K(\cdot, b)$; hence the boundary of $\mathcal{H}_1$ is the same as the boundary ∂G of G, and thus $\mathcal{H}_1$ can also be interpreted as the space of probability measures on ∂G.

The group G acts on $\mathcal{H}$ by right translation:

$$(R_s h)(r) = h(rs).$$

This action preserves the rays of $\mathcal{H}$, so it induces an action on the quotient space $\mathcal{H}/\sim$. This quotient space is canonically identified with $\mathcal{H}_1$ where the action of G now reads:

$$(s, h)(r) = h(rs)/h(s).$$

This action of G on $\mathcal{H}_1$ has a fixed point, namely the constant function 1 and, because of the cocycle property of the kernels K, leaves invariant the boundary ∂G of G.

There is a similar integral representation theorem for bounded (not necessarily positive) harmonic functions that reads as follows:

Proposition 2.5. *Let h be a bounded harmonic function on G. Then there is a function φ in $L^\infty(\partial G, \lambda)$ such that*

$$h(s) = \int_{\partial G} K(s, b)\varphi(b)\lambda(b).$$

The function φ is non–positive if and only if h is non–positive.

Other fact that we need to know about positive harmonic functions is the following:

Proposition 2.6. *Let h be a positive harmonic function on G. Then the non tangential limits exist almost everywhere on ∂G with respect to the measure λ.*

The same type of result is available for bounded harmonic functions.

Proposition 2.7. *Let h be a bounded harmonic function with integral representation $h = \int K\varphi\lambda$. Then h converges to φ non–tangentially almost everywhere.*

If m is a probability measure on ∂G, then Lebesgue decomposition theorem guarantees the existence of a non–negative function φ in $L^1(\partial G, \lambda)$ and a measure σ singular with respect to λ, such that

$$m = \varphi\lambda + \sigma.$$

In the theorem concerning non–tangential limits of a function h in $\mathcal{H}_1$ (corresponding to a probability measure m on ∂G) one can furthermore prove that the boundary values of h agree with the function φ almost everywhere. This brings us another piece of notation: we say that the harmonic function h is singular if its corresponding boundary measure m is singular with respect to λ, i.e., its boundary values are zero λ–almost everywhere. Analogously, we say that h is absolutely continuous if its boundary measure has no singular part.

3. Discrete dynamical systems

The dynamical systems that we will consider consist of a space X (usually a compact metric space) and a finitely generated group G of transformations of X. More important than the group G itself is a generating system for it. Therefore we will denote by G_1 a finite generating system for the group G. (The emphasis on the generating system instead of on the group is already present in the classical theory of dynamical systems as well.) The group G is the free group generated by G_1.

We will consider G acting on the left on X, so we have a map

$$G \times X \to X$$

taking (s, x) to sx and satisfying $(rs)x = r(sx)$ for any $r, s \in G$. The qualities of the action would be imposed by those of X, namely, it will be assumed continuous if X is a topological space, measurable if X is a measure space, etc.

Assume that G has a system of weights q of the type described in the previous section. Associated with it there are the diffusion and Laplacian operators acting on continuous functions on X, called D and Δ, respectively, which are defined as follows: if f is in $C(X)$ then

$$Df(x) = \sum_{s \in G_1} q(s) f(sx)$$

and

$$\Delta f = Df - f.$$

The action of G on X induces an action of G on $C(X)$ on the right, namely $R_s f(x) = f(sx)$. By duality, this in turn induces a left action of G on measures on X, and we will denote by $s\mu$ the measure $s\mu(A) = \mu(s^{-1}(A))$. Thus, if δ_x is the point mass measure concentrated at x, then $s\delta_x = \delta_{sx}$. Also, by duality, the diffusion operator D and Laplacian Δ act on measures on X: if μ is a measure, then $D\mu$ is defined by

$$\int_X f(x) D\mu(x) = \int_X Df(x)\mu(x).$$

A measure μ on X is called harmonic if $D\mu = \mu$, equivalently, if

$$\int_X \Delta f(x)\mu(dx) = 0$$

for every continuous function f on X.

The existence of harmonic measures for a dynamical system is an easy consequence of the maximum principle and the Hahn–Banach theorem:

Theorem 3.1. *Let (X, G_1) be a dynamical system, where X is a compact topological space. Then there is a measure μ on X which is harmonic.*

Proof. If f is a continuous function on X, so is Δf. Suppose that x_0 is a point of X where f reaches a maximum, $f(x_0) \geq f(x)$ for all $x \in X$. Then $\Delta f(x_0) = Df(x_0) - f(x_0) \leq 0$.

It follows that the subspace $\mathrm{Im}(\Delta)$ of $C(X)$ does not contain the constant function 1. The linear functional σ defined on $\mathbb{R} \cdot 1 + \mathrm{Im}(\Delta)$ by $\Lambda(t1 + \Delta f) = t$ has norm $\|\Lambda\| = 1$, vanishes on $\mathrm{Im}(\Delta)$ and $\Lambda(1) = 1$. It therefore extends to a linear functional on $C(X)$ with the same properties, by the Hahn–Banach theorem. The

Riesz representation theorem implies the existence of a probability measure μ on X such that

$$\int_X \Delta f(x)\,\mu(dx) = 0$$

for all f in $C(X)$. □

It is clear that this method of proof applies to more general types of actions: all one needs is that the operator D takes continuous functions to continuous ones, and that it satisfies some kind of weak maximum principle. Of course, the standard proof using Cesaro averages of sequences $D^n\delta_x$ also works.

It is curious that this simple argument gives a new proof of the famous theorem of Kriloff and Bogoliougoff [17] to the effect that a homeomorphism of a compact metric space has an invariant measure. For if $T : X \to X$ is the endomorphism, then we consider the operator $\Delta : f \to f - f \circ T$. At a maximum (respectively, minimum) of f we have $\Delta f \geq 0$ (respectively ≤ 0), and the Hahn–Banach theorem implies the existence of a measure μ which is T–invariant. I have not been able to locate such proof in the literature; but E. Ghys says it is known to him.

Proposition 3.2. *Let (X, G_1) be a dynamical system with harmonic measure μ. Then μ is quasi–invariant. In particular, the support of μ is a union of orbits of G.*

Proof. That μ is harmonic means that

$$\mu = D\mu = \sum_{s \in G_1} q(s) s\mu.$$

If A is a Borel subset of X such that $\mu(A) = 0$, then this equality says that $s\mu(A) = 0$ for all $s \in G_1$. By iteration, $s\mu$ is absolutely continuous with respect to μ for all $s \in G$. Moreover, the relation $s\mu \prec \mu$ implies $\mu = (s^{-1})s\mu \prec s\mu$. The argument is finished by noticing that if $\mu(A) = 0$, then all iterates of A also have measure 0, hence so does $\cup_{s\in G} sA$. □

4. Markov operators and random walks

In this section we discuss in more detail the diffusion operator D and construct the path space associated with the dynamical system. The operator D has a probabilistic interpretation on this space, and we will analyze some of it aspects in order to deal with the ergodic decomposition of harmonic measures.

The Laplacian Δ is the infinitesimal generator of the semi group of operators $D^n : C(X) \to C(X)$ constructed by iteration of D. We have $D^n 1 = 1$, and $\|Df\| \leq \|f\|$ with respect to the supremum norm in $C(X)$. Since D is also a positive operator ($Df \geq 0$ if $f \geq 0$), the map $f \in C(X) \mapsto Df(x)$ corresponds to a probability measure on X. We denote this probability measure by $p(1; x, \cdot)$, so that we have

$$Df(x) = \int_X f(y) p(1; x, dy).$$

Clearly, if δ_x denotes the Dirac measure at the point x, then $p(1; x, \cdot) = D\delta_x$. We set $p(0; x, \cdot) = \delta_x$ to maintain homogeneity in notation. If $p(n; x, dy) = D^n\delta_x$ is associated to D^n, then there is a convolution formula, which is a reflection of the Markov property, that says:

$$p(n+m;x,B) = \int_X p(n;x,dy)p(m;y,B).$$

The measure p is easily described as follows

Proposition 4.1. *Let $\phi : s \in G \mapsto sx$ be the map of G onto the orbit of the point $x \in X$. Let $q(n;s,\cdot)$ be the probability distribution associated to the random walk in G determined by the system of weights q. Then $\phi q(n;s,\cdot) = p(n;sx,\cdot)$.*

Proof. By definition, if f is a function on X,

$$Df(x) = \sum_{s\in G_1} q(s)f(sx) = \int_X f(y)p(1;x,dy).$$

But $Df(x) = DF(e)$, where $F = \phi \circ f$, and

$$DF(s) = \int_G F(r)p(1;s,dr)$$

and the claim follows. □

A path in (X, G_1) is a map $w : \mathbb{N} \to X$ such that $w(n+1) \in \{sw(n) \mid s \in G_1\}$. The path w is said to start at x if $w(0) = x$. The path space of (X, G_1) is the set of all such paths, and it will be denoted by $\Lambda(X)$. It is topologized via the inclusion in the infinite product space $X^\infty = \mathrm{Maps}(\mathbb{N}, X)$. In what follows we describe a few topological and measure theoretic aspects of $\Lambda(X)$ that will be used in the sequel.

Proposition 4.2. *The set $\Lambda(X)$ is closed in X^∞, and it is non–empty if X and G_1 are non–empty.*

Proof. The set Ω_1 consisting of all sequences $(x_0, x_1, \cdots)$ of points of X with the requirement that $x_1 = sx_0$ for some s in G_1 is easily seen to be a finite union of closed sets, hence closed. Using a similar scheme, one constructs a nested sequence $\Omega_1 \supset \Omega_2 \supset \Omega_3 \cdots$ of closed sets such that $\Lambda(X) = \cap_{n=1}^\infty \Omega_n$. □

The path space $\Lambda(X)$ comes equipped with a family of projections $\pi_n : \Lambda(X) \to X$ defined as the evaluation maps $\pi_n(w) = w(n)$, and with a semi group of transformations generated by the shift map $\theta : \Lambda(X) \to \Lambda(X)$, which acts via $(\theta w)(n) = w(n+1)$. A given distance function d_X on X induces a distance d on $\Lambda(X)$ by setting

$$d(w,w') = \sum_{n=0}^{\infty} \frac{1}{2^n} d_X(\pi_n w, \pi_n w').$$

The shift map θ is actually Lipschitz, indeed:

$$d(\theta w, \theta w') = 2d(w,w') - 2d_X(w(0), w'(0)) \leq 2d(w,w').$$

To see that this theory is a true generalization of the classical one concerning a map $T : X \to X$, note that in this case the path space $\Lambda(X)$ is X itself, as one point determines the full forward orbit.

For each $x \in X$ we want to construct a probability measure P_x in the fiber $\pi_0^{-1}(x)$, similar to Wiener measure. One can define a probability measure on X^∞ by viewing it as the inverse limit of the finite products X^F, $F \subset N$ a finite subset, and using Kolmogorov's extension technique. Specifically, for a finite set $F = \{n_1, \cdots, n_k\}$ (ordered), and Borel set $B = B_1 \times \cdots \times B_k$ in X^F, we set

$$P_x^F(B) = \int_{B_1} p(n_1; x, dx_1) \int_{B_2} p(n_2 - n_1; x_1, dx_2) \cdots \int_{B_k} p(n_k - n_{k-1}; x_{k-1}, dx_k).$$

This is a consistent family of measures on the finite products X^F, so there is a unique probability measure P_x on X^∞ whose marginals are the measures P_x^F. Clearly the measure P_x gives total mass to the fiber $\pi_0^{-1}(x)$.

Our path space $\Lambda(X)$ is a proper subset of X^∞ which is where the measures P_x are defined, and therefore we would like to show that it has total probability with respect to P_x.

Proposition 4.3. *The space $\Lambda(X)$ has total probability with respect to P_x.*

Proof. If $w \notin \Lambda(X)$, then there is an integer $n \geq 0$ such that $sw(n) \neq w(n+1)$ for all $s \in G_1$. Therefore there are open neighborhoods V of $w(n)$ and W of $w(n+1)$ in X such that W is disjoint from sV for all $s \in G_1$. Let U be the open set $\pi_n^{-1}V \cap \pi_{n+1}^{-1}W$. This set U is clearly disjoint from $\Lambda(X)$ and its probability under P_x is

$$\begin{aligned} P_x(U) &= P_x[w(n+1) \in W, w(n) \in V] \\ &= \int_V p(n; x, dy) \int_W p(1; y, dz) \\ &= 0, \end{aligned}$$

because the support of the measure $p(1; y, \cdot)$ is the set $\{sy \mid s \in G_1\}$, which in turn is disjoint from W whenever y belongs to V. □

For future use we record the following:

Proposition 4.4. *The map*

$$P : X \times \mathfrak{F} \to \mathbb{R}$$

with the properties:

(1) *For each fixed measurable subset A of $\Lambda(X)$, the map*

$$x \to P_x[A]$$

is measurable.

(2) *For each fixed x in X, P_x is a probability measure in $\Lambda(X)$.*

We also record the following relation between diffusion in X and expectation in path space: if f is a Borel function on X, then

$$D^n f(x) = E_x[f \circ \pi_n] = \int_{\Lambda(X)} f \circ \pi_n(w) P_x(w).$$

A consequence of this expression is that if μ is a harmonic measure for the system (X, G_1) then the measure $\widetilde{\mu}$ defined in $\Lambda(X)$ by

$$\widetilde{\mu}(A) = \int P_x[A] \mu(x)$$

is invariant under the shift θ, as one easily verifies.

Let $\mathcal{B}$ denote the Borel algebra of X. Let $\mathcal{F}$ denote that of X^∞ generated by cylinder sets. For $n \geq 0$, let $\mathcal{F}_n$ denote the sub–algebra of $\mathcal{F}$ generated by cylinders sets in X^∞ of the form

$$\{w \mid w(n_1) \in B_1, \cdots, w(n_k) \in B_k\},$$

where $n_i \leq n$ $(i = 1, \cdots, k)$ and $B_i \in \mathcal{B}$. We have a filtration

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \mathcal{F}.$$

Analogously, let $\mathcal{F}^n$ denote the borel algebra of X^∞ generated by cylinder sets of the form

$$\{w \mid w(n_1) \in B_1, \cdots, w(n_k) \in B_k\},$$

where the steps $n_i \geq n$. Finally, let $\mathcal{F}^\infty$ denote the algebra $\mathcal{F}^\infty = \cap_{n=0}^\infty \mathcal{F}^n$, also known as the algebra of tail events.

Note that if C is a cylinder set in $\mathcal{F}^n$, then $\theta^{-1}C$ is a cylinder set that belongs to $\mathcal{F}^{n+1}$, that is, $\theta^{-1}\mathcal{F}^n = \mathcal{F}^{n+1}$. Therefore the map

$$\theta^k : (\Lambda(X), \mathcal{F}^{n+k}) \to (\Lambda(X), \mathcal{F}^n)$$

is measurable for $n, k \geq 0$.

The result that follows is called the Zero-One law of Kolmogorov [15], [5], and is a ergodicity fact.

Proposition 4.5. *If $A \in \mathcal{F}^\infty$, then $P_x[A] = 0$ or 1. In particular, this holds if A is a θ–invariant set.*

Proof. The reason is that a tail event is independent of itself. I include the proof of [5] for convenience. Let $A \in \mathcal{F}^\infty$ be a tail event. Since $\mathcal{F}_n$ and $\mathcal{F}^\infty$ are independent fields, A is independent of $\mathcal{F}_n$ for each n, that is $P_x(A \cap B) = P_x(A)P_x(B)$ for each $B \in \cup_{n=1}^\infty \mathcal{F}_n$. The probability

$$P_A[B] = \frac{P_x[A \cap B]}{P_x[A]}$$

agrees with P_x on $\cup_{n=0}^\infty \mathcal{F}_n$, hence also in $\mathcal{F}$. Taking $B = A$ we see that $P_x[A] = P_x[A]^2$.

□

We need the stronger fact concerning shift invariant subsets of $\Lambda(X)$.

Proposition 4.6. *If A is a shift invariant subset of $\Lambda(X)$, then $P_x[A] = 0$ or 1, and the function $x \to P_x[A]$ is constant along the G–orbit of x.*

Proof. If A is a shift invariant subset of $\Lambda(X)$, then $A \in \mathcal{F}^0 = \mathcal{F}$ and $A = \theta^{-n}A \in \mathcal{F}^n$ for all $n \geq 0$, that is, $A \in \mathcal{F}^\infty$ and moreover $\chi_A = \chi_A \circ \theta$. Therefore, utilizing the Markov property we obtain

$$\begin{aligned} P_x[A] = E_x[\chi_A] &= E_x[\chi_A \circ \theta] \\ &= E_x[E_x[\chi_A \circ \theta | \mathcal{F}_1]] \\ &= \int p(1; x, dy) P_y[A] \\ &= \sum_{s \in G_1} q(s) P_{sx}[A], \end{aligned}$$

which says that if $P_x[A] = 0$ (respectively 1), then also $P_{sx}[A] = 0$ (respectively 1). □

As a consequence we obtain the following corollary.

Proposition 4.7. *Let μ be a harmonic measure on X and let $\widetilde{\mu}$ be the induced measure on $\Lambda(X)$. Then $\widetilde{\mu}$ is ergodic for the shift θ if and only if μ is ergodic for the system G_1.*

Proof. If A is G–invariant, then $\pi_0^{-1}A$ is shift invariant. Moreover,

$$\mu(A) = \widetilde{\mu}(\pi_0^{-1}A),$$

and so μ is ergodic if $\widetilde{\mu}$ is. Conversely, if A is a shift invariant set and μ is ergodic, then the G-invariant function $x \mapsto P_x[A]$ must be essentially constant, so we must have $P_x[A] = 0$ or $P_x[A] = 1$ for almost all x. Hence either A or its complement in $\Lambda(X)$ is a null set with respect to $\widetilde{\mu}$. □

5. The ergodic theorem

In this section we will recall the statement of the ergodic theorem for Markov operators and then use it to study our dynamical system. After some preparatory facts, we will develop the Krillof–Bogoliougoff theory.

Let (X, G_1) be a dynamical system and let

$$D : C(X) \to C(X)$$

be its associated diffusion operator. Given a harmonic probability measure μ on X, the operator D extends to $L^1(\mu)$. Indeed, if f is continuous in X, then $|Df| \leq D|f|$, and so

$$\|Df\|_1 = \int_X |Df|(x)\mu(x) \leq \int_X |f|(x)\mu(x) = \|f\|_1$$

by the D–invariance of μ. Therefore D is bounded in the dense subspace of $L^1(\mu)$ constituted by the continuous functions, and thus it extends to $L^1(\mu)$. The same applies to the other L^p spaces ($p \neq \infty$); it also extends to $L^\infty(\mu)$, the space of essentially bounded functions, because μ is a probability measure, and satisfies $\|D\|_\infty \leq 1$.

The ergodic theorem relates the time averages of a function to the space averages. The proof of the following operator version of the ergodic theorem can be found in [6].

Theorem 5.1. *Let $D : L^1(\mu) \to L^1(\mu)$ be a linear operator with $\|D\|_1 \leq 1$ and $\|D\|_\infty \leq 1$. If f is a μ–integrable function, then there exists a μ–integrable D–invariant function f^* such that the sequence*

$$\frac{1}{n}\sum_{k=0}^{n-1} D^k f$$

converges to f^* almost everywhere.

It is at this point where the importance of harmonic measures reveals itself, being able to pass from D–invariant functions to G–invariant ones. This is the reflection of the well known geometrical fact that a harmonic function on a compact manifold must be constant. For the continuous version see [11].

Proposition 5.2. *If f is a D–invariant and μ–integrable function, then the class of f in $L^1(\mu)$ contains a G–invariant function.*

Proof. For a function f as in the statement and a rational number r, let $f_r = \min\{f, r\}$. We have $Df_r = \min\{Df, r\} \leq \min\{f, r\} = f_r$. Since f is integrable and $|f_r| \leq |f|$, both f_r and Df_r are integrable. Moreover,

$$\int_X Df_r(x)\mu(x) = \int_X f_r(x)\mu(x),$$

because of the invariance of μ. Therefore the set of points x of X where $Df_r(x) < f_r(x)$ has measure 0. If we let A denote this set, then the quasi–invariance of μ implies that $A_r = \cup_{s \in G} sA$ also has μ–measure 0. Hence the G–invariant set $B_r = X \setminus A_r$ has probability one, and in it the equality $Df_r(x) = f_r(x)$ holds.

If x is a point of B_r, then $Df(x) > r$ if $f(sx) > r$ for some s in G_1. On these points, we have $f_r = r$. Since $Df_r = f_r$ on B_r, starting with x we see that $f_r(sx) = r$ for all s in G_1. By iteration, we must have $f_r = r$ on the whole orbit of x. Therefore, the set of points x in B_r where $f(x) > r$ is also invariant. The set $B = \cap_{r>0} B_r$ has full measure and is invariant, and f is constant on each orbit contained in B. □

The proof actually gives something slightly more general:

Proposition 5.3. *If f is a positive excessive integrable function, i.e., $0 \leq Df \leq f$, then f is constant almost everywhere.*

We now proceed to develop the theory Krillof and Bogoliugoff [17] of ergodic sets. It is true that the ergodic decomposition of measures could be carried out using the concepts of integral boundary representation introduced by Choquet [23], but the work of [17] provides a deeper analysis of the ergodic components of a harmonic measure. The presentation that follows has been influenced by those of [22] and [21].

Let S denote the set of points w in $\Lambda(X)$ such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (f \circ \pi_0 \circ \theta^k)(w)$$

exits for every continuous function f on X. Since the subspace of $C(\Lambda(X))$ generated by the functions $f \circ \pi_n$, where $f \in C(X)$ and $n \in \mathbb{N}$ is dense in $C(\Lambda(X))$, the set S coincides with the set of points of $\Lambda(X)$ where the limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (F \circ \theta^k)(w)$$

exists for each continuous function F on $\Lambda(X)$.

By its very definition, this set S is a shift invariant subset of $\Lambda(X)$. Therefore, by our previous discussion, the function $x \in X \mapsto P_x(S)$ takes on the values 0 and 1 only, and it is constant along the orbits of G. Hence the subset Q of X consisting of those points x in X for which $P_x(S) = 1$ is invariant under G, and it is called the set of quasi–regular points of the system (X, G_1).

Moreover, by the ergodic theorem applied to the classical dynamical system $(\Lambda(X), \theta)$, the set S is of invariant measure 1, i.e., $\nu(S) = 1$ for every shift invariant probability measure ν on $\Lambda(X)$. We say that a subset B of X has harmonic measure 1 if $\mu(B) = 1$ for every harmonic probability measure for (X, G_1).

Proposition 5.4. *The set Q of quasi–regular points is measurable and of harmonic measure one.*

Proof. As X is a compact metrizable space, so is $\Lambda(X)$. It then has a countable base, equivalently, $C(\Lambda(X))$ has a countably dense subset. Thus S can be described as a countable intersection of sets, each determined as the set where a sequence of functions has limit. Therefore S is measurable.

That Q is measurable follows from the fact that the family of measures $P = \{P_x\}$ in path space has the property that for fixed measurable set $S \subset \Lambda(X)$ the map

$$x \in X \to P_x[S]$$

is measurable.

Finally, if μ is a harmonic probability measure on X, then $\widetilde{\mu} = P_x(w)\mu(x)$ is shift invariant, and we have

$$1 = \widetilde{\mu}(S) = \int_{\Lambda(X)} P_x[S]\mu(x) = \mu(Q).$$

□

Proposition 5.5. *For each x in Q and each continuous function f on X, the limit*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} D^k f(x)$$

exists.

Proof. Recall that the diffusion of f at x can be computed as

$$Df(x) = E_x[f \circ \pi_1].$$

The functions of paths

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ \pi_k(w)$$

are integrable and, for $w \in S$, they converge pointwise to an integrable function. Hence, if x is such that $P_x[S] = 1$, the dominated convergence theorem ensures the existence of the limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} D^k f(x) = E_x \left[\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (f \circ \pi_0 \circ \theta^k)(w) \right],$$

the interchange of limits being justified because of Fubini's theorem, as f is bounded and P_x is a probability measure. □

If $x \in Q$ a quasi–regular point, then the assignment

$$f \in C(X) \mapsto \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} D^k f(x)$$

is a continuous linear functional on $C(X)$ of norm 1 taking the constant function 1 to 1. Therefore there is a probability measure μ_x on X such that

$$\int_X f(y)\mu_x(y) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} D^k f(x).$$

Proposition 5.6. *The measure μ_x defined by a quasi–regular point in the above manner is harmonic. Furthermore, all the measures μ_{sx}, $s \in G$, are mutually absolutely continuous.*

Proof. It is plain, from its definition, that μ_x is D–invariant. To prove the last statement, we observe that

$$\begin{aligned}\sum_{s\in G_1} q(s)\left(\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} D^k f(sx)\right) &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\sum_{s\in G_1} q(s)D^k f(sx)\\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n-1} D^k f(x)\\ &= \int_X f(y)\mu_x(y),\end{aligned}$$

from what it follows that

$$\mu_x = \sum_{s\in G_1} q(s)\mu_{sx}.$$

This relation implies the mutual absolute continuity of the measures μ_{sx}. □

To obtain the ergodic decomposition of an arbitrary harmonic measure μ, we use the fact that Q has harmonic measure one, hence we can write

$$\begin{aligned}\int_X f(x)\mu(x) &= \frac{1}{n}\sum_{k=0}^{n-1}\int_X D^k f(x)\mu(x)\\ &= \int_Q \frac{1}{n}\sum_{k=0}^{n-1} D^k f(x)\mu(x)\\ &= \int_Q\int_X f(y)\mu_x(y)\mu(x)\end{aligned}$$

for every continuous function. Hence it holds for all functions because, for a fixed borel set B, the assignment $x \to \mu_x(B)$ is a measurable function on X.

Two classes of quasi–regular points are introduced according to properties of the measure. A point x is said to be a point of density if $\mu_x(U) > 0$ for every neighborhood U of x. It is said to be transitive if the measure μ_x is ergodic. These sets are denoted by the symbols Q_D and Q_T, respectively, and the intersection $R = Q_D \cap Q_T$ is called the set of regular points.

Proposition 5.7. *Let f be a nonnegative integrable function on X, and let f^* denote the function*

$$f^* = \lim_n \frac{1}{n}\sum_{k=0}^{n-1} D^k f(x) > 0.$$

Let μ be a harmonic measure for (X, G_1). Then for almost all points x in X, either $f(x) = 0$ or $f^*(x) > 0$.

Proof. This is a consequence of the ergodic theorem. Let B denote the set of points x of X where $f^*(x) = 0$ or undefined. This is a G–invariant set, up to measure 0. Applying the ergodic theorem to $\chi_B f$ we have that the Cesaro averages of the sequence $D^k\chi_B f$ converge to $\chi_B f^*$, which is equal to 0 almost everywhere. Hence

$$\int_X \chi_B(x)f(x)\mu(x) = 0.$$

Thus $f = 0$ on B and $f^* > 0$ on $X \setminus B$. □

Proposition 5.8. *The set Q_D is a borel measurable G–invariant subset of X of harmonic measure one.*

Proof. The set of points of density is invariant because, as it was proved above, the measures μ_{sx} are mutually absolutely continuous.

Let U_i be a countable base of open sets for the topology of X. For each i, let f_i be a continuous function on X such that $f_i > 0$ on U_i and $f_i = 0$ on $X \setminus U_i$. The set B_i on which $f_i = 0$ or $\widetilde{f}_i > 0$ has harmonic measure one. Moreover,

$$Q_D = Q \cap \bigcap_{i=1}^{\infty} E_i.$$

If x is a quasi–regular point in $\cap E_i$, then for every i, either $f_i(x) = 0$ or $f_i^*(x) > 0$. By definition of μ_x, it must be that $\mu_x(U_i) > 0$ whenever $x \in U_i$.

Conversely, if x is such that $f_i(x) > 0$ and $f_i^*(x) = 0$, then the neighborhood U_i of x has measure $\mu_x(U_i) = 0$.

□

Proposition 5.9. *The set Q_T of transitive points is invariant and of harmonic measure one.*

Proof. This set is invariant because of the mutual absolute continuity of the measures μ_{sx}, $s \in G$.

As in [21], one shows that the set of points $x \in Q$ for which

$$\int_Q (f^*(x) - f^*(y))^2 \mu_x(y) = 0$$

holds for any continuous function is measurable and of harmonic measure one.

By the ergodic theorem, a point $x \in Q$ is transitive if and only if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} D^k f(y) = \int_X f(z)\mu_x(z)$$

for all f in some countable dense subset of $C(X)$ and μ_x–almost all $y \in X$. Together with the previous paragraph, we obtain the desired result.

□

If μ be an ergodic harmonic measure for (X, G_1), then the ergodic theorem implies that there is an invariant measurable subset E of Q with $\mu(E) = 1$ and such that $\mu = \mu_x$ for all $x \in E$. This set E is called the quasi–ergodic set of μ. The ergodic set of μ is the intersection $E \cap R$.

If $T : Y \to Y$ is a map of a space and y and E are a point and a subset of Y respectively, we let

$$\widetilde{\tau}_E(y, n) = \frac{1}{n} \sum_{k=0}^{n-1} \chi_E(T^k y).$$

Let (X, G_1) be a dynamical system. For a point x in X and a borel subset B of X, let

$$\tau_B(x, n) = \frac{1}{n} P_x[\widetilde{\tau}_{\widetilde{B}}(w, n)].$$

Using the correspondence between diffusion on X and the shift on $\Lambda(X)$, it is clear that

$$\tau_B(x,n) = \frac{1}{n}\sum_{k=0}^{n-1} D^k \chi_B(x),$$

and this would be used to close up the discussion of the Krilloff–Bogoliugoff theory. We need one more piece of terminology: say that a closed subset F of X is a center of attraction if

$$\tau_U(x) = \lim_{n\to\infty} \tau_U(x,n) = 1$$

for every $x \in F$ and every neighborhood U of F. Then we have the following result, the proof of which is exactly as in [17] [21].

Proposition 5.10. *The closure F of the set R of regular points of (X, G_1) is a minimal center of attraction.*

Further applications of the theory of Kryloff and Bogoliouboff can be found in [22]. All of them have an adequate translation to the case of harmonic measures considered here, but we leave that to the interested reader.

I will conclude this section with a problem on recurrence motivated by a reading of [9]. Recall that a subset A of $\mathbb{N}$ is said to have positive density if the limit

$$\limsup_{n\to\infty} \frac{1}{n}\mathrm{Card}(A \cap [0, n-1]) > 0.$$

Analogously, we say that a subset A of the group G (or semi group) has positive density if

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} D^k \chi_A(e) > 0,$$

which clearly agrees with the previous definition for $G = \mathbb{N}$.

On the space $X = \{0,1\}^G$ of subsets of G (with the product topology) there is an action of G defined by

$$(s,x)(r) = x(rs).$$

Let $p \in X$ be a point such that the set $S = \{s \mid p(s) = 1\}$ has positive upper density in G. The sequence L_n of linear functionals on $C(X)$, defined by

$$L_n f = \frac{1}{n}\sum_{k=0}^{n-1} D^k f(p)$$

has a subsequence which converges to a harmonic measure μ_p on X. Let A be the subset of X consisting of all points x for which $x(e) = 1$. Its characteristic function χ_A is continuous, $\chi_A(x) = x(e)$, and

$$D^n \chi_A(p) = D^n \chi_S(e)$$

where the left hand side D is the operator in X and the right on G. Hence if S has positive density, then $\mu_p(A) > 0$.

Conversely, by the ergodic theorem,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} D^k \chi_A(x) = \widetilde{\chi}_A(x)$$

exists for almost all $x \in X$, and

$$\mu(A) = \int_X \chi_A(x)\mu(dx) = \int_X \widetilde{\chi}_A(x)\mu(dx).$$

Thus if $\mu(A) > 0$, there is $p \in A$ such that $\widetilde{\chi}_A(p) > 0$; in other words, the set $\{s \mid p(s) = 1\}$ has positive upper density.

These types of construction of measures have been exploited by Furstenberg [10] in his dynamical systems approach to the Szemeradi theorem, which says that a subset of the integers of positive upper density contains arbitrarily large arithmetic progressions. We can formulate a similar type of problem when replacing the group of integers $\mathbb{Z}$ by a finitely generated group as follows. If S is a subset of G of positive upper density, then there is a subgroup H of G of finite index, a positive integer $r > 0$ and $s \in G$ such that S contains the translate $H_r s$ of the set H_r of elements of H of length less that r in the generators of H. Apparently this does not follow by applying Furstenberg's techniques to the shift transformation in the path space. One certainly obtains a finite index subgroup of G with certain recurrence property, but not as strong as the one just mentioned.

6. The Jacobian cocycle and the tautological action

In this section we will discuss some properties of harmonic measures for a dynamical system (X, G_1) suggested by the harmonic analysis on the group G. To this end, we introduce another space associated to the dynamical system (X, G_1) called the greater path space $\Omega(X) = \Omega(G) \times X$. It comes equipped with a shift transformation, usually denoted by θ, defined by $\theta(w, x) = (\theta w, w(1)x)$. This transformation θ inherits all properties of the action of G on X, because the coordinate functions $w(n)$ are locally constant on $\Omega(G)$.

If μ is a harmonic measure for (X, G_1), then the product measure $m = P \times \mu$ is shift invariant; in fact, for a measurable subset $W \times A$ of $\Omega(X)$ we have that $\theta^{-1}(W \times A)$ can be written as the disjoint union

$$\bigcup_{s \in G_1} Ws \times s^{-1}A.$$

Therefore:

$$m(\theta^{-1}(W \times A)) = \sum_{s \in G_1} P(Ws)\mu(s^{-1}A) = P(W)D\mu(A).$$

As it was the case with $\Lambda(X)$ and the measure $\widetilde{\mu}$, if μ is ergodic, then the measure $m = P \times \mu$ is ergodic for the shift θ on $\Omega(X)$.

The first consequence of the harmonic analysis on G consists on applying the the solution of the Dirichlet problem on ∂G to prove the ergodicity of the measure λ.

Proposition 6.1. *Let G be a free group with symmetric generating set G_1. Then the measure λ on its boundary ∂G is harmonic and ergodic.*

Proof. We have already mentioned that the measure λ is harmonic. Actually the action is ergodic when restricted to a finite index subgroup G' of G (but note that λ may not be harmonic for G'). Indeed, let G' be such subgroup and let A be a G'-invariant measurable subset of ∂G with measure $\lambda(A) > 0$. We will show that A has full measure.

If we denote by χ_A the characteristic function of A, then the function

$$h(s) = \int_{\partial G} K(s,b)\chi_A(b)\lambda(b)$$

is harmonic on G. Moreover, it satisfies $0 \le h(s) \le 1$, and it is G'–invariant because A has such property.

Since $\lambda(A) > 0$, there are points x in A for which $h(s) \to 1$ as s approaches b radially, i.e., while remaining at bounded distance from the geodesic ray which starts at s and limits at b.

Since G' is of finite index in G, it has a finite fundamental domain F, and so we can find a sequence $\{s_i\}$ of elements of G and s in F such that $s_i s$ approaches b, so $h(s_i s) \to 1$. By the G'-invariance of h and the finiteness of F, we conclude that $h(s) = 1$. The maximum principle then implies that h is constant equal to 1. Thus

$$1 = h(e) = \int_{B(G)} K(e,b)\chi_A(b)\lambda(b) = \lambda(A).$$

□

One can furthermore prove that λ is the only harmonic measure on ∂G which is ergodic. This can be done by utilizing the theory of integral representation on convex sets [23], or via the ergodic theorem as it is done in [11] for foliations arising from $\mathrm{PSl}(2,\mathbb{R})$. I will briefly sketch how this works because the context of [11] is slightly different.

Proposition 6.2. *The measure λ on ∂G is the only ergodic harmonic measure for the system $(\partial G, G_1)$*

Proof. We may view the path space of the system $(\partial G, G_1)$ as the equivariant image of the product $\Omega(G) \times \partial G$, where the action is $(w,x) \mapsto (\theta w, w(1)x)$. The key is to note that if x and y are two distinct points in ∂G, then almost all paths in G will hit ∂G at some other point distinct from x or y, and if w is one of those paths, then

$$\lim_{n\to\infty} d(w(n)\cdots w(1)x, w(n)\cdots w(1)y) = 0.$$

If f is a continuous function on ∂G, then applying the ergodic theorem to λ we have that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(\theta^k w) = \int_{\partial G} f(b)\lambda(b)$$

for almost all paths starting at x. By the geometric property above, this then holds for almost all paths starting at any other point of ∂G. Therefore the limit is independent of the measure, i.e., the system $(\partial G, G_1)$ is uniquely ergodic. □

Continuing with the discussion of the dynamical system (X, G_1) with harmonic measure μ, we recall that each transformed measure $s\mu$ is absolutely continuous with respect to μ. If the transformation s of G is invertible, then, by the Radon–Nikodym theorem, we can find a positive integrable function $j(s,\cdot)$ in $L^1(X,\mu)$ such that

$$\int_X f(x)s^{-1}\mu(x) = \int_X f(x)j(s,x)\mu(x).$$

When G is the group generated by G_1, this allows us to define the jacobian cocycle

$$j : G \times X \to (0,\infty).$$

It is defined for almost all $x \in X$ and all $s \in G$, and satisfies the chain rule relation

$$j(rs, x) = j(r, sx)j(s, x)$$

for all r, s in G and almost all x in X.

Proposition 6.3. *If $x \in X$ is fixed, then the map*

$$s \in G \to j(s, x) \in (0, \infty)$$

is harmonic on G for the dual system of weights q^{-1}.

Proof. This follows by simply taking the Radon–Nikodym derivative of the identity $\mu = D\mu$. Specifically,

$$\begin{aligned} 1 = \frac{dD\mu}{d\mu}(rx) &= \sum_{s \in G_1} \frac{ds\mu}{d\mu}(rx) \\ &= \sum_{s \in G_1} q(s)j(s^{-1}, rx) \\ &= \sum_{s \in G_1} q(s^{-1})j(sr, x)j(r, x)^{-1}. \end{aligned}$$

Hence

$$j(r, x) = \sum_{s \in G_1} q(s^{-1})j(sr, x),$$

as desired. □

In fact, not only $j(s, \cdot)$ belongs to $L^1(X, \mu)$ for each $s \in G$, but also to $L^\infty(X, \mu)$, because by Harnack's inequality we have

$$C^{-n} \leq j(s, x) \leq C^n,$$

for each x in X, where n is the length of s as a word in G_1.

More generally, the Radon-Nikodym derivative defines a positive harmonic function on the group G which is invariant by the stabilizer of the corresponding point. If we agree to call a subgroup G' of G martinian if the only G'–invariant positive harmonic functions on G are the constants, then we also have:

Proposition 6.4. *If the orbit of almost every point in the support of μ is martinian, then μ is equivalent to an invariant probability measure.*

Let $\mathcal{R}$ denote the quotient of $\mathcal{H}$ by the equivalence relation $u \sim v$ if $u = av$ for some positive real number a. The map $X \to \mathcal{H}$ taking x to the function $j_x(\cdot) = j(\cdot, x)$ is a Borel function on X. Replacing x by another point on its orbit $y = sx$ has the following effect:

$$j(\cdot, sx) = R_s j(\cdot, x)j(s, x)^{-1}.$$

This relation implies that the quotient map

$$J : X \to \mathcal{R}$$

is G–equivariant, that is $J(sx) = sJ(x)$, where the left action of G on $\mathcal{R}$ is induced by that on $\mathcal{H}$ which consists of precomposition with right translation. Since there is a canonical identification of $\mathcal{H}_1$ with $\mathcal{R}$, we will denote by $J : X \to \mathcal{H}_1$ the map just constructed.

Note that the action of G on $\mathcal{H}_1$ has a tautological harmonic cocycle, namely the map

$$\psi : G \times \mathcal{H}_1 \to (0, \infty)$$

defined by $\psi(s, h) = h(s)$ satisfies $\psi(sr, h) = \psi(r, sh)\psi(s, h)$, and evidently it is a harmonic function of s for fixed h.

If the measure μ on X is invariant under the action of G, then the map J is essentially constant equal to 1, and the push forward measure $J\mu$ is the point measure at the constant function 1 on G. In general, there is a kind of centralization about this constant function. This can be seen by observing that the measure $J\mu$ is a probability measure on the compact convex set $\mathcal{H}_1$ and therefore it has a barycenter, that is, there is a point β in $\mathcal{H}_1$ such that

$$F(\beta) = \int_{\mathcal{H}_1} F(h) \cdot J\mu(h)$$

for every linear functional F on $\mathcal{H}$. In particular, for the functional F_s given by evaluation at $s \in G$ we obtain

$$\begin{aligned} \beta(s) &= \int_{\mathcal{H}_1} h(s) J\mu(h) \\ &= \int_X j(s, x) \cdot \mu(x) \\ &= 1 \end{aligned}$$

where the second identity follows by virtue of the change of variables formula, and the last because $j(s, \cdot)\mu = s^{-1}\mu$ is also a probability measure. Summarizing:

Proposition 6.5. *The barycenter of $J\mu$ is the constant function 1. The jacobian cocycle of $J\mu$ agrees with the tautological cocycle for the action of G on $\mathcal{H}_1$, almost everywhere with respect to $J\mu$.*

We may decompose the measure μ with respect to the projected measure $J\mu$. That is, there is a measurable assignment of a probability measure ν_h on the fiber $J^{-1}(h)$ such that

$$\int_X f(x)\mu(x) = \int_{\mathcal{H}_1} \left(\int_{J^{-1}(h)} f(x)\nu_h(x) \right) J\mu(h).$$

From this expression, we see that the jacobian cocyle of $J\mu$ is

$$\frac{dJ\mu(sh)}{d\mu(h)} = \int_{J^{-1}(h)} j(s, x)\nu_h(x),$$

but if $J(x) = h$, then $j(s, x) = h(s)$, so we obtain the claim.

Applying the boundary limit theorems to the functions in the jacobian cocycle j we obtain a measurable map

$$\varphi : \partial G \times X \to (0, \infty)$$

Since it is positive and has integrable marginals, Fubini–Tonelli theorem implies that this function φ belongs to $L^1(\partial G \times X, \lambda \times \mu)$.

If h is a function on $\mathcal{H}_1$, then we can write

$$h(s) = \int_{\partial G} K(s, b) m_h(b)$$

for some probability measure m_h on ∂G. This measure m_h admits a unique Lebesgue decomposition $m_h = \varphi\lambda + \sigma$, where $\varphi \in L^1(\partial G, \lambda)$ and $\sigma \perp \lambda$. Under the action of $s \in G$ on $\mathcal{H}_1$, the function h goes to $h_s(r) = h(rs)/h(s)$, whose corresponding measure is easily computed to be

$$[h(r)K(r^{-1}, \cdot)]^{-1} r m_h.$$

Since any two measures $r\lambda$ are mutually absolutely continuous, the decomposition of the measure of h_s takes the form

$$h(r)^{-1}\varphi(r^{-1}\cdot)\lambda + [h(r)K(r^{-1}, \cdot)]^{-1} r\sigma.$$

We therefore see that G acting on $\mathcal{H}_1$ leaves invariant the subspace of singular harmonic functions. This subspace contains all harmonic functions $K(\cdot, b)$, $b \in B$; it is a subspace which forms the set of extreme points of the convex set $\mathcal{H}_1$, and remains itself invariant (the action there is the action of G on ∂G).

A calculation like the one already done shows that the function φ satisfies the following identity almost everywhere:

$$j(s, x)\varphi_{sx}(sb) = \varphi_x(b).$$

For a measurable subset A of X, let

$$j(s, A) = \int_A j(s, x)\mu(dx).$$

For fixed A, this is a harmonic function on G as one sees differentiating under the integral sign. It satisfies the inequality $0 \le j(s, A) \le 1$, the second one because $j(s, A) = \mu(sA) \le \mu(X) = 1$. Also, as one easily verifies, if r, $s \in G$, then $j(s, rA) = j(sr, A)$. Therefore the harmonic function $j(s, A)$ is bounded, hence absolutely continuous, with boundary measure $\varphi_A\lambda$, where $\varphi_A \in L^\infty(\partial G, \lambda)$, and it satisfies

$$\varphi_{sA}(b) = \varphi_A(sb).$$

It follows that the measure μ can also be defined as

$$\mu(A) = \int_{\partial G} \varphi_A(b)\lambda(b).$$

Define a new measure μ_a by setting

$$\mu_a(A) = \int_{\partial G}\int_X \varphi_x(b)\chi_A(x)\mu(x)\lambda(b).$$

By Fatou's lemma, we have $\varphi_A(b) \ge \int_A \varphi_x(b)\mu(x)$, so this measure μ_a is absolutely continuous with respect to μ. We compute, using the property of φ_x listed above, that

$$\mu_a(sA) = \int_{\partial G}\int_X \chi_A(x)\varphi_x(b)K(s, b)\mu(x)\lambda(b),$$

which means that μ_a is also a harmonic measure for (X, G_1).

Similarly, we can define a measure μ_z, absolutely continuous with respect to μ, by integrating the singular component of the boundary functions. Namely

$$\mu_z(A) = \int_{\partial G}\int_X \chi_A(x)\sigma_x(b)\mu(x).$$

A calculation like the one above, using the transformation formula for the boundary measures of harmonic functions, shows that this measure is harmonic for the system (X, G_1), and absolutely continuous with respect to μ.

Consequently, we have

Proposition 6.6. *A harmonic measure μ can be decomposed as $\mu = \mu_a + \mu_z$, where μ_a is a harmonic measure of continuous type, and μ_z is harmonic of singular type, and $\mu_a \perp \mu_z$.*

The last step of our analysis will consist of making this decomposition more explicit, which will be achieved by analyzing the boundary values of the jacobian cocycle j. Associated to the jacobian cocycle of the harmonic measure μ, there is a measurable function ℓ on $\Omega(X)$ defined by

$$\ell(w, x) = -\log j(w(1), x).$$

Because j is a cocycle, we immediately note that

$$\sum_{k=0}^{n-1} \ell(\theta^k(w, x)) = -\log j(w(n) \cdots w(1), x),$$

and hence the partial sums converge almost everywhere to the value of boundary function $-\log \varphi_x$ at the point $\lim w = \lim_{n\to\infty} w(n) \cdots w(1)$ of ∂G.

To study the asymptotic behavior of the partial sums of ℓ along the orbits of θ we consider the associated transformation

$$T : \Omega(X) \times \mathbb{R} \to \Omega(X) \times \mathbb{R}$$

defined by $T(y, t) = (\theta y, t + \ell(y))$.

This transformation T of $\Omega(X)\times\mathbb{R}$ preserves the measure $m\times dt$, where dt denotes Lebesgue measure on $\mathbb{R}$. As such, the map T is either dissipative or conservative, that is, either it admits a wandering set or it does not admit one. Recall that a wandering set for T is a set F of positive measure such that the sets $F, T^{-1}F, T^{-2}F, \cdots$ are pairwise disjoint.

This dichotomy for the transformation T translates into the following one for an ergodic harmonic measure μ: either μ is invariant or singular. Equivalently, for an ergodic harmonic measure μ we have

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} \ell(\theta^k(w, x)) = 0$$

or

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} \ell(\theta^k(w, x)) = \infty$$

for almost all points (w, x) of $\Omega(X)$.

We first examine the case in which T is conservative. For this we do not need ergodicity.

Proposition 6.7. *The transformation T is conservative if and only if*

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} \ell(\theta^k(w, x)) = 0$$

almost everywhere on $\Omega(X)$, equivalently, if and only if the measure μ is invariant.

Proof. If T is conservative, then for $\epsilon > 0$ we have

$$\sum_{n=1}^{\infty} \chi_{Y_\epsilon} \circ T^n = \infty$$

almost everywhere on $Y_\epsilon = \Omega(X) \times [-\epsilon, \epsilon]$, according to the recurrence theorem of Halmos [14]. Therefore, since $T^n(y,t) = (\theta^n y, t + \sum_{k=1}^n \ell(\theta^k y))$, we have that

$$\sum_{n=1}^{\infty} \chi_{Y_\epsilon}(t + \sum_{k=1}^{n} \ell(\theta^k y)) = \infty$$

for almost all $y = (w,x) \in \Omega(X)$ and $t \in [-\epsilon, \epsilon]$. This implies that

$$|\log j(\lim w, x)| \leq 2\epsilon$$

for almost all $(w,x) \in \Omega(X)$.

Conversely, if

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} \ell(\theta^k(w,x)) = 0$$

then $\varphi_x(b) = 1$ for almost all $b \in \partial G$ and almost all $x \in X$. In view of the relation $j(s,x)\varphi_{sx}(sb) = \varphi_x(b)$ we then have that $j(s,x) = 1$ almost everywhere, hence that μ is invariant. But then the transformation T of $\Omega(X) \times \mathbb{R}$ is essentially $T(y,t) = (\theta(y), t)$, which is obviously conservative.

□

Before considering the case in which T is dissipative, we first we note that the set Z of points $(w,x) \in \Omega(X)$ where

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} \ell(\theta^k(w,x)) = \infty$$

is shift invariant. Indeed, using the cocycle property of j, we have that

$$\sum_{k=0}^{n-1} \ell\theta^k(\theta w, w_1 x) + \ell(w_1, x) = \sum_{k=0}^{n} \ell\theta^k(w,x)$$

so the last term converges to ∞ if $\theta(w,x) \in Z$, and vice versa; that is $\theta^{-1}Z = Z$.

Hence, in case that the measure μ is ergodic, which we assume for the rest of this discussion, either Z or its complement has m–measure 0. If it is of full measure then the harmonic measure is of singular type and T is dissipative.

If Z has measure zero, then let

$$L(w,x) = \lim_{n\to\infty} \sum_{k=0}^{n-1} \ell(\theta^k(w,x))$$

This function is finite almost everywhere and equal to $-\log \varphi_x(\lim w)$. Moreover, the action of T on $\Omega(X) \times \mathbb{R}$ has the property that

$$T(w,x,L(w,x)+t) = (\theta(w,x), L(\theta(w,x)) + t)$$

i.e., it leaves the graph of L and of its vertical translates invariant. Therefore, any subset B of $\Omega(X) \times \mathbb{R}$ of the form

$$B = \{(w,x,t) \mid L(w,x) + a \leq t \leq L(w,x) + b\}$$

(with $a < b$ real numbers) is invariant under T and of finite measure. Hence such set B cannot meet any wandering set for T.

To complete the proof we have to show that such wandering set F exists. As T is dissipative, the union W of all wandering sets of T is a positive measure subset of $\Omega(X) \times \mathbb{R}$. Moreover, if F is wandering, then also $T^{-1}F$ is wandering; therefore W is invariant under T. Furthermore, if F_r denotes the set of points (y, s) such that $(y, s + r) \in F$, we have that $T^{-1}F_r = (T^{-1}F)_s$. Thus W is invariant under the translations $(y, s) \mapsto (y, s + r)$, $r \in \mathbb{R}$. If ϕ denotes the transformation defined by $\phi(y, s) = (\theta y, s)$, then $\phi T = T\phi$, from what it follows that W is also invariant under ϕ. Therefore W is invariant under the action of $\mathbb{N} \times \mathbb{R}$ defined by $(n, r)(y, s) = (\theta^n y, s + r)$. The action of $\mathbb{R}$ on itself by translations is ergodic, and since we are assuming that θ is ergodic, Fubini's theorem implies that this action on $\Omega(X) \times \mathbb{R}$ is also ergodic. Therefore, $W = \Omega(X) \times \mathbb{R}$ (up to a null set), and the existence of the appropriate wandering set F is assured. In fact, in this situation, one can express $\Omega(X) \times \mathbb{R}$ as a countable union of wandering sets. Hence, if T is dissipative and $B \subset \Omega(X) \times \mathbb{R}$ has positive measure, then B contains a wandering set.

In conclusion,

Proposition 6.8. *The transformation T is dissipative if and only if the harmonic measure μ is of singular type (hence non–invariant).*

Since the decomposition $\mu = \mu_c + \mu_z$ of a harmonic measure is preserved under its ergodic decomposition (i.e., the ergodic components of a harmonic measure of singular type are almost all of singular type, and analogously for absolutely continuous type), we in fact have the following:

Proposition 6.9. *A harmonic measure μ for the system (X, G_1) can be written uniquely as $\mu = \mu_a + \mu_z$, where μ_a is invariant, and μ_z is harmonic of singular type. Moreover, μ_a and μ_z are mutually singular.*

The property of the jacobian cocycle described here should be contrasted with those properties of the jacobian cocycle of a quasi–invariant measure of an action of the integers, as studied in [19], [24] and [16].

A corollary to this discussion is that a system (X, G_1) with singular ergodic harmonic measure is of type III. More precisely:

Proposition 6.10. *Let (X, G_1) be a dynamical system and let μ be an ergodic probability measure which is not invariant. Then μ is not equivalent to a σ–finite invariant measure.*

Proof. If such σ–finite measure ν existed, then we can write $\mu = f\nu$, for some positive measurable function f on X, integrable with respect to ν. This function f satisfies the functional equation

$$f(sx) = j(s, x)f(x)$$

almost everywhere on X.

Let A be a subset of X with $0 < \nu(A) < \infty$. We may assume that A is such that $1/C \leq f \leq C$ for some constant $C > 0$. By recurrence with respect to the measure μ, and a fortiori with respect to the measure ν, almost all paths starting at points in A return to A. Since the measure μ is of singular type, we also have that $j(s, x)$ approaches 0 along almost all paths staring in A.

In view of the functional equation, these two facts are contradictory. Therefore no such function f could exist.

□

In terms of the tautological action, this property says that the support of an ergodic harmonic measure on $\mathcal{H}_1$ for the action of G is either the constant function 1 or else almost every function in its support is singular. I do not know whether a more precise description of the minimal center of attraction of the dynamical system $(\mathcal{H}_1, G_1)$ can be obtained. Certainly the topological structure is of no much help because the singular harmonic functions are dense in $\mathcal{H}_1$; on the other hand, there is some regularity implicit in those harmonic functions arising from harmonic measures, i.e., the cocycle property.

Note also that it is not difficult to construct ergodic systems whose jacobian cocycle does not have image contained in ∂G. A simple example consists on taking the action of G on ∂G generated by G_1 and then adding some new elements to G_1 which act trivially on ∂G. The measure λ is also harmonic for this enlarged system, but now there are points in the boundary of the new group where the harmonic functions in the cocycle take on the value one.

From a different point of view, this property of the jacobian of a non invariant harmonic measure is in some sense a weak form of the vanishing of coefficients for unitary representation of semisimple Lie groups. Such vanishing of coefficients is of course false for unitary representations of a free group.

7. Ergodicity and mixing

We have seen that the concept of ergodicity for a harmonic measure on a dynamical system can be successfully interpreted in terms of the corresponding measure in path space. In this section we explore this and other ergodic properties of harmonic measures, specially in connection with the induced action of G and of D in the hilbert space of square integrable functions on the underlying space.

Associated to the action of G on X having quasi–invariant measure μ there is a unitary representation of G on $L^2(X,\mu)$, called the Koopman representation, defined by

$$\kappa(s^{-1})f(x) = f(sx)j(s,x)^{1/2},$$

where j is the jacobian cocycle.

By analogy with representations of $\mathrm{Sl}(2,\mathbb{R})$, one has a family of unitary representations obtained by replacing the multiplicative factor $j(s,x)^{1/2}$ by $j(s,x)^{1/2+it}$, where t is a real number. In general, these representations are not equivalent.

In the ergodic theory of a single measure preserving transformation of a probability space there is an interpretation of the ergodicity of the measure as being equivalent to the constants being the only invariant square integrable functions. In the case of singular harmonic measures (i.e., non invariant) the situation is essentially more complicated.

We need some preliminary facts. The first one is from classical dynamical systems (a very elegant proof, due to M. Sion, can be found in [23]), and reads as follows:

Proposition 7.1. *Let μ be an invariant measure for a dynamical system (Y,T) and let $\nu = f\mu$ be absolutely continuous with respect to μ. Then ν is invariant if and only if $f \circ T = f$.*

The second is essentially the harmonic measure version of the first one:

Proposition 7.2. *Suppose that μ is a harmonic measure for (X, G_1) and ν is absolutely continuous with respect to μ, with Radon-Nikodym derivative $f = d\nu/d\mu$. Then ν is harmonic if and only if f is G–invariant.*

Proof. If f is G–invariant, then not only $Df = f$, but also $D(fg) = fDg$ for every function g on X. Therefore we have

$$\int_X gD\nu = \int_X D(fg)\mu = \int_X gf\mu.$$

Suppose that ν is harmonic. Both measures $\widetilde{\mu}$ and $\widetilde{\nu}$ are shift invariant, and from the definition of $\widetilde{\nu}$ it follows that it is absolutely continuous with respect to $\widetilde{\mu}$, with corresponding Radon–Nikodym derivative $f \circ \pi_0$. Indeed,

$$\int_{\Lambda(X)} F(w)\widetilde{\nu}(dw) = \int_X f(x) \int F(w)P_x(dw)\mu(dx)$$

and since $f \circ \pi_0$ is constant equal to $f(x)$ with respect to P_x, this last integral equals

$$\int_X \int_{\Lambda(X)} f \circ \pi_0(w)F(w)P_x(dw)\mu(dx) = \int_{\Lambda(X)} F(w)f \circ \pi_0(w)\widetilde{\mu}(dw).$$

In view of the previous statement, this implies that the composite function $f \circ \pi_0$ is shift invariant, which in turn means that f is invariant under the action of G on X.

□

With these preliminary facts out of the way, we state the first consequence of ergodicity for harmonic measures.

Proposition 7.3. *Let (X, G_1) be a dynamical system with harmonic singular measure μ. Then the representation κ has no finite dimensional invariant subspace.*

Proof. Arguing for a contradiction, suppose that V is an invariant subspace of finite, but positive, dimension. Let $f_1, \cdots, f_n$ be an orthonormal basis of V, and let f denote the function in $L^1(X, \mu)$ defined by

$$f(x) = \sum_{i=1}^{n} |f_i(x)|^2.$$

To each $s \in G$ there corresponds a unitary matrix $(a_{ij}(s))$ determined by

$$f_i(sx)j(s,x)^{1/2} = \sum_{j=1}^{n} a_{ij}(s)f_j(x).$$

This implies that the function g satisfies the equation

$$f(sx)j(s,x) = f(x).$$

Since by construction f is also positive, the measure $f\mu$ is invariant and absolutely continuous with respect to μ. By the above, it then follows that f is constant along the orbits of G. But then the functional equation $f(sx)j(s,x) = f(x)$ says that the jacobian cocycle is essentially constant equal to 1 as a function of $s \in G$, which we know to be false for singular harmonic measures. □

This property of the Koopman representation is reminiscent of the weakly mixing property. The weakly mixing property for unitary representations of semisimple Lie groups is related to the vanishing of the coefficients at infinity. A coefficient of κ is a function φ_{uv} on G which one associates to a pair of elements u, v of $\mathcal{H}$ by setting

$$\varphi_{uv}(s) = \langle \kappa(s)u, v\rangle.$$

The coefficients of κ are said to vanish at infinity if $\varphi_{uv}(s) \to 0$ as $s \to \infty$ in G, for all $u, v \in \mathcal{H}$. For the relevance of the vanishing of coefficients property to ergodic theory, see [27].

We will briefly discuss this property for the Koopman representation attached to a dynamical system (X, G_1) with ergodic non invariant measure μ.

A coefficient of κ is a function φ on G of the form

$$\varphi(s) = \int_X f(s^{-1}x)\overline{g(x)}j(s^{-1}, x)^{1/2}\mu(x),$$

where $f, g \in L^2(X, \mu)$. In order to show that φ vanishes at infinity on G, it is enough to show it for the case $f = g = \chi_X$, for then it would follow for simple functions in $L^2(X, \mu)$, hence for all square integrable functions.

We then have only to consider the function

$$\varphi(s) = \int_X j(s, x)^{1/2}\mu(x).$$

If we assume that no element $s \neq e$ of G preserves μ, then we have, by Schwarz's inequality

$$\varphi(s) < \int_X j(s, x)\mu(x) = 1.$$

Moreover, using the concavity of the square root function,

$$j(s, x)^{1/2} = \left(\sum_{r\in G_1} q(r^{-1})j(rs, x)\right)^{1/2} \geq \sum_{r\in G_1} q(r^{-1})j(rs, x)^{1/2},$$

for almost all $x \in X$. Upon integration,

$$0 \leq D\varphi(s) = \sum_{r\in G_1} q(r^{-1})\varphi(rs) \leq \varphi(s),$$

that is, φ is excessive (superharmonic and non negative) on G. By iteration of this inequality we have $D^{n+1}\varphi \leq D^n\varphi$, hence the limit $h = \lim_{n\to\infty} D^n\varphi$ exists and it is a nonnegative harmonic function on G. We have then a Riesz decomposition

$$\varphi = h + k$$

where k is a potential. The boundary limit of a potential such as k is zero almost everywhere on $(\partial G, \lambda)$ [4]; thus φ and h have the same boundary values. However, the vanishing of coefficients is apparently a stronger condition that the vanishing of the jacobian cocycle; one needs some regular contraction of the measure μ under the action of G. For instance, under the assumption that for no element $s \neq e$ of G the measure μ has no nontrivial subordinate s–invariant measure, one can prove that φ vanishes on a dense subset of ∂G. But, in contrast with semisimple lie groups, the action of G on its boundary is not transitive.

Next we turn to the action of the diffusion operator on $L^2(X, \mu)$. Ergodicity can be characterized in terms of the operator D as follows:

Proposition 7.4. *The system* (X, G_1) *is ergodic if and only if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int (D^k \chi_A)\chi_B \mu = \mu(A)\mu(B)$$

for all measurable subsets A, B *of* X.

Proof. If μ is ergodic so is $\widetilde{\mu}$. Therefore the ergodic theorem applied to the shift θ implies that for any measurable subset A of X

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (\chi_A \circ \pi_0 \circ \theta^k)(w) = \widetilde{\mu}(\pi_0^{-1} A) = \mu(A).$$

But

$$\int_{\Lambda(X)} (\chi_A \circ \pi_0 \circ \theta^k)(w)\widetilde{\mu}(dw) = \int_X D^k \chi_A.$$

Therefore, after multiplication by χ_B and integration, the dominated convergence theorem gives the first statement. The rest is proved in a similar fashion. □

As for the eigenvalues of the operator D acting on $L^2(X, \mu)$ we have the following:

Proposition 7.5. *Let* (X, G_1) *be a dynamical system with harmonic measure* μ *which is ergodic. If* f *is an eigenfunction of* D *in* $L^2(X, \mu)$ *then either* f^2 *is constant (and the eigenvalue is* ± 1*) or else* $D^n f$ *converges to* 0 *as* $n \to \infty$.

Proof. If f is an eigenfunction of D, i.e., $Df = \lambda f$ and $\lambda \neq 1$, then from the inequality $|Df| \leq D|f|$ and the invariance of μ we obtain $\int f\mu = 0$ and $|\lambda| \leq 1$.

In case $|\lambda| < 1$, we have $|D^n f| = |\lambda|^n |f|$, so $D^n f$ converges to 0 in $L^2(X, \mu)$. On the other hand, if $|\lambda| = 1$ then $|f| \leq D|f|$. If μ is ergodic, being D–invariant in turn implies that the functions $|f|$ and $D|f|$ must be equal almost everywhere, hence that $|f|$ is essentially constant. Therefore we may assume that f takes values in the unit circle in the complex plane. But then the identity

$$Df(x) = \sum_{s\in G_1} q(s) f(sx) = \lambda f(x)$$

says that the circle number $\lambda f(x)$ is in the convex hull of the circle numbers $f(sx)$, $s \in G_1$. This can only happen if $f(sx) = \lambda f(x)$ for all $s \in G_1$. Iteration of this identity shows that $\lambda = \pm 1$. This in turn implies that such function f satisfies $D(f^2) = f^2$, and so, by the ergodicity of μ, f^2 is constant almost everywhere. □

This property of the diffusion operator suggests it has stronger spectral properties than those of the unitary operator associated to a measure preserving transformation of a probability space. In fact, in the case of invariant measure, the operator D satisfies a strong form of weak mixing as we will show next.

We need to introduce one more piece of notation. For a function f on X, let ∂f denote the function on X with values on $\mathbb{C}^{G_1}$ given by

$$\partial f(x)(s) = q(s)(f(sx) - f(x)).$$

Then we have the following identities, reminiscent of familiar ones in differential geometry:

$$\Delta f = \mathrm{tr}\, \partial f,$$

(where tr operates on a vector $v = (v_1, \cdots, v_n)$ by $\mathrm{tr}\,(v) = \sum v_i$), and

$$\Delta(fg) = f\Delta g + g\Delta f + (\partial f, \partial g)$$

for functions f and g on X, where the product $(u(s), v(s)) = \sum_s q(s)u(s)v(s)$.

Let $(\ ,\ )$ denote the inner product in $\mathbb{C}^{G_1}$ and let $\langle\ |\ \rangle$ denote the inner product in $L^2(X, \mu)$. For every function f in $L^2(X, \mu)$ we have:

$$\begin{aligned}\|D^{n+1}f\|^2 - \|D^n f\|^2 &= \langle D^{n+1}f \mid D^{n+1}f\rangle - \langle D^n f \mid D^n f\rangle \\ &= \langle \Delta D^n f \mid \Delta D^n f\rangle + \langle D^n f \mid \Delta D^n f\rangle + \langle \Delta D^n f \mid D^n f\rangle \\ &= \|\Delta D^n f\|^2 - \int_X (\partial D^n f, \partial D^n f)\mu\end{aligned}$$

where we have used the fact that μ is harmonic. Since we always have

$$(\mathrm{tr}\ (\partial D^n f))^2 \leq (\partial D^n f, \partial D^n f),$$

it follows that the sequence $\|D^n f\|^2$ is non–increasing, hence that it has a limit as $n \to \infty$.

Note that in general the diffusion operator D is not unitary in $L^2(X, \mu)$, its adjoint D^* being given by

$$D^* f(x) = \sum_{s \in G_1} q(s^{-1}) f(sx) j(s, x).$$

Therefore, D is a symmetric operator if the measure μ is invariant and the system of weights is symmetric. Hence, assuming that μ is invariant and $q(s) = q(s^{-1})$, we have:

$$\|D^{2n} f - D^{2n+2m} f\|^2 = \|D^{2n} f\|^2 + \|D^{2m} f\|^2 - 2\|D^{m+2n} f\|^2,$$

and similarly for the odd indexed sequence. Therefore both sequences $\{D^{2n} f\}$ and $\{D^{2n+1} f\}$ are Cauchy in $L^2(X, \mu)$, and so they converge there to functions f_e and f_o, respectively.

In general, both limits could be distinct: for instance, let $X = \{+1, -1\}$ and let G_1 consist of the map $s(x) = -x$ and its inverse. The function $f(x) = x$ provides an ergodic example where the even and odd limits are distinct. In any case, we have the following:

Proposition 7.6. *Let (X, G_1) be a dynamical system with invariant measure μ and suppose that $q(s)$ is symmetric. If D acting on $L^2(X, \mu)$ does not have -1 as an eigenvalue, then $D^n f$ converges to a D–invariant function in $L^2(X, \mu)$. In particular, if μ is ergodic, $D^n f$ converges to $\int_X f\mu$.*

Proof. With the notation introduced above, it is clear that $Df_e = f_o$ and $Df_o = f_e$. Thus if they are not equal, $f_e - f_o$ is an eigenfunction of D corresponding to the eigenvalue -1. In the ergodic case, the identification of the limit is obtained via the ergodic theorem. □

All the concepts used in the study of a single transformation can be extended to multidimensional systems by formulating the corresponding concept in path space $(\Lambda(X), \theta)$. We have done that for the ergodicity property, and next look at the weakly mixing property.

A dynamical system $T : X \to X$ with invariant measure ν is said to be weakly mixing if the system $T : X \times X \to X \times X$ is ergodic with respect to the product measure $\nu \times \nu$. When trying to extend this concept to a dynamical system (X, G_1) with harmonic measure μ we face some difficulties. First of all, and even in the case of a symmetric equidistributed system of weights on G, the product $\mu \times \mu$ is not harmonic, unless μ is invariant. Second, the path space of $(X \times X, G_1)$ is not the product of the path space of (X, G_1) with itself. These difficulties are eliminated once we define correctly what the product action should be.

Given a dynamical system (X, G_1), we consider the dynamical system on $X \times X$ generated by the set of maps $G_1 \times G_1$. It acts on $X \times X$ in the obvious way, namely by $(s_1, s_2)(x, y) = (s_1 x, s_2 y)$. We denote this product system by $(X \times X, G_1^2)$, and note that the path spaces are related by

$$(\Omega(X \times X, G_1^2), \theta) \sim (\Lambda(X) \times \Lambda(X), \theta).$$

If q is a system of weights on G_1, then the product G_1^2 has the system of weights $q^2(sr) = q(s)q(r)$. With this notion of product action the concept of product of harmonic measures is well behaved, that is:

Proposition 7.7. *If μ is a harmonic measure for (X, G_1) then the product measure $\mu \times \mu$ is harmonic for $(X \times X, G_1^2)$.*

Proof. The operator D on $X \times X$ for the system generated by G_1^2 is defined by

$$Df(x, y) = \sum_{r,s \in G_1} q(s)q(r)f(sx, ry).$$

In particular, for a function f on $X \times X$ of the form $f(x, y) = f_1(x)f_2(y)$ we have that $Df(x, y) = Df_1(x)Df_2(y)$. The statement follows from this fact. □

We say the the dynamical system (X, G_1) is weakly mixing if the associated path space system $(\Lambda(X), \theta)$ is weakly mixing. Since this one is a classical system, and the product $(\Lambda(X) \times \Lambda(X), \theta \times \theta)$ is the path space of the system $(X \times X, G_1^2)$, weak-mixing means that this last system is ergodic with respect to the product measure.

There is a characterization of the weak mixing property in terms of the operator D analogous to the ergodicity property. It reads

Proposition 7.8. *The system (X, G_1) is weakly mixing if and only if*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left| \int_X (D^k \chi_A)\chi_B \mu - \mu(A)\mu(B) \right| = 0$$

for all measurable subset A, B of X.

Proof. If $(\Lambda(X), \theta)$ is weakly mixing, then for every pair A, B of subsets of X,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left| \widetilde{\mu}(\theta^{-k} A \cap B) - \mu(A)\mu(B) \right| = 0.$$

Using the Markov property, i.e., the definition of the measures P_x and $\widetilde{\mu}$, one checks that

$$P_x[\theta^{-k} A \cap B] = P_x[w(k) \in A; w(0) \in B] = D^k \chi_A(x)\chi_B(x),$$

and the rest follows. □

To conclude this section, note that if (X, G_1) has harmonic measure μ, while (Y, G_1) has invariant measure ν, then the product measure $\mu \times \nu$ is harmonic for the diagonal system $(X \times Y, G_1)$. It would be interesting to know whether this system is ergodic when μ is weakly mixing and ν is ergodic.

8. Topological entropy

In these section, we collect some observation on the topological entropy of a dynamical system (X, G_1). They were motivated as complementing some aspects of [13], and as a search for a variational principle for the entropy.

In order to define the topological entropy of a dynamical system (X, G_1) using open covers, we recall a few standard facts about open covers of a space. The collection of open covers of a space X is partially ordered by refinement. If $\mathcal{U}$ and $\mathcal{V}$ are open covers, we write $\mathcal{U} \prec \mathcal{V}$ if every element of $\mathcal{U}$ is contained in some element of $\mathcal{V}$. Every pair of covers has a greatest lower bound which is denoted by $\mathcal{U} \vee \mathcal{V}$, and consists of the sets of the form $A \cap B$ with $A \in \mathcal{U}$ and $B \in \mathcal{V}$. Analogously, if $\mathcal{U}_1, \cdots, \mathcal{U}_n$ is a finite collection of open covers of X, the greatest lower bound is denoted by $\vee_{i=1}^n \mathcal{U}_i$, its elements being the sets of the form $A_1 \cap \cdots \cap A_n$ with $A_j \in \mathcal{U}_j$.

The entropy of a finite open cover $\mathcal{U}$ of X is the number $\log N(\mathcal{U})$, where $N(\mathcal{U})$ denotes the minimum cardinality of a finite sub-cover of $\mathcal{U}$.

Let G_n denote the set of elements of G of length at most n with respect to the path distance defined via G_1. If $\mathcal{U}$ is a finite cover of X, then the entropy of (X, G_1) with respect to $\mathcal{U}$ is

$$h(X, G_1, \mathcal{U}) = \lim_{n\to\infty} \frac{1}{n} \log N(\vee_{s\in G_{n-1}} s\mathcal{U}),$$

and the topological entropy of the system (X, G_1) is

$$h(X, G_1) = \sup_{\mathcal{U}} h(X, G_1, \alpha).$$

As in the case of single transformations, the entropy can also be defined in terms of a distance function, which gives the version of entropy of pseudogroups utilized in [13]. To do so, denote by d_X a distance function on the compact metric space X, and declare two points x and y in X to be (n, ϵ)–separated if there is s in G_{n-1} such that

$$d(sx, sy) > \epsilon.$$

Let $S(n, \epsilon)$ denote the maximum cardinality of an (n, ϵ)–separated subset of X. Then we set

$$h(X, G_1, \epsilon) = \limsup_{n\to\infty} S(n, \epsilon)$$

and the entropy of the dynamical system is defined as

$$h(X, G_1) = \lim_{\epsilon\to 0} h(X, G_1, \epsilon).$$

Analogously, say that a set F of X is (n, ϵ)–spanning if for any $y \in X$ there exists $x \in F$ such that

$$d(sx, sy) < \epsilon$$

for all s in G_{n-1}. Lets $R(n, \epsilon)$ denote the minimum cardinality of the (n, ϵ)–spanning sets. Proceeding as in the previous paragraph, one obtains yet another possible definition of entropy. As in [25], one can prove that the entropy of (X, G_1) can be computed with respect to open covers, separated sets or spanning ones, giving the same result, which is denoted by $h(X, G_1)$.

We note another useful fact concerning the entropy of a system (X, G_1) on a compact metric space.

Proposition 8.1. *The entropy of a system (X, G_1) on a compact metric space depends only on the uniform structure of X, that is, the number $h(X, G_1)$ is independent of the distance on X used in the sequence of calculations leading to its definition.*

For instance, using the definition of entropy with coverings, we can compute the topological entropy of $(\partial G, G_1)$. We can cover ∂G with $\mathrm{Card}(G_1)$ mutually disjoint discs, namely those determined by the initial symbol of the corresponding element of ∂G. If $\mathcal{U}$ denotes this open cover of ∂G, then the elements of the cover $\vee_{G_n} s\mathcal{U}$ are the discs determined by the first $n-1$ symbols, a total of $\mathrm{Card}(G_1)(\mathrm{Card}(G_1)-1)^{n-2}$. Hence

$$h(\partial G, G_1) = \log(\mathrm{Card}(G_1) - 1)$$

(One has equality by the standard properties of generators.) Curiously, the hausdorff dimension of ∂G, with respect to the obvious non–archimedian distance, is one.

As noted in [25], the concept of entropy is closely related to the notion of Kolmogorov dimension of a metric space. Let (Y, d) be a complete metric space. A subset E of Y is said to be ϵ–separated if $d(x, y) > \epsilon$ for every pair of distinct points x and y of Y. The Kolmogorov dimension of (Y, d) is

$$\limsup_{\epsilon \to 0} \frac{\log S(Y, d, \epsilon)}{\log(1/\epsilon)}$$

where $S(Y, d, \epsilon)$ denotes the maximum cardinality of an ϵ–separated subset of Y.

Unlike entropy, the Kolmogorov dimension is not an invariant of the uniform structure of the metric space. It is mentioned here because of the following:

Proposition 8.2. *Let (X, G_1) be a dynamical system on compact metric space X of finite Kolmogorov dimension. If all elements of G_1 are Lipschitz with constant $K > 1$, then*

$$h(G, G_1) \leq C \log K.$$

A property of the dynamical system (X, G_1) guaranteeing that the entropy is positive involves the notion of resilient point. This is noted in [13] in the case of codimension one foliations. Here we simply generalize the definition.

A point x of X is resilient if there are elements s and r in the group G with the following properties:

(1) x is a contracting fixed point of s; that is, $s(x) = x$ and there is a neighborhood U of x such that the closure of sU is contained in U and $\cap_{n=0}^{\infty} s^n U = \{x\}$.
(2) There is a neighborhood V of x such that $rV \subset U \setminus \{x\}$.

The proof of the following statement is exactly the same as that presented in [13] for pseudogroups arising from codimension one foliations. Note that their proof actually gives an explicit lower bound on the entropy.

Proposition 8.3. *If the system (X, G_1) has a resilient point then the entropy $h(X) \geq \log 2$.*

We have mentioned the concept of resilient point because many of the familiar examples of dynamical systems have them. On one hand, it is clear from the definition that an action of the group of the integers $\mathbb{Z}$ cannot have a resilient point.

Also note that the orbit of a resilient point has exponential growth; therefore many other groups cannot have actions with this kind of points.

On the other hand, a cocompact Fuchsian group acting on the boundary of the unit disc, and more generally a negatively curved group acting on its boundary, has resilient points. Abstracting such situation we have the concept of quasi–conformal action of a group G on a metric space X. We recall the definition and refer to [12] for a detailed discussion of this concept.

Let $\phi : Y \to Y$ be a homeomorphism of a metric space Y with distance d. The conformal distortion of ϕ at a point y of Y is defined by

$$\delta(\phi, y) = \limsup_{\epsilon \to 0} \frac{\sup\{d(\phi(y), \phi(y')) \mid d(y, y') = \epsilon\}}{\inf\{d(\phi(y), \phi(y')) \mid d(y, y') = \epsilon\}}.$$

Thus $1 \leq \delta(\phi, y) \leq \infty$. If $\delta(\phi, y)$ is bounded on Y, then ϕ is said to be quasi-conformal.

We say that G is a group of quasi–conformal homeomorphisms of X if there is a constant C such that $\delta(s, x) \leq C$ for every element s of G and every x in X.

Proposition 8.4. *Let (X, G_1) be a dynamical system on a metric space X and suppose that G is a group of quasi conformal maps of X. If X has a harmonic measure which is not invariant, then there is a resilient point.*

Proof. Using the ergodic decomposition, we may assume that there is a harmonic measure μ which is not invariant and which is defined by a regular point x. We then know that there is a set of positive measure such that the boundary limits of the jacobian maps $j(s, y)$ are zero almost everywhere for almost all points y in such set. (Ergodicity and the cocycle property then imply that this is true for almost every point of X with respect to such measure.)

Recalling that the Radon-Nikodym derivative can be computed as

$$j(s, y) = \lim_{D \to y} \frac{\mu s D}{\mu D},$$

that $j(s, y)$ is a singular harmonic function, together with the fact that y is a point of density, means that there is an open disc D around y such that the measure $\mu(sD)$ tends to 0 along almost all paths in G, and that almost every path starting at y passes near y at some later time. Quasi–conformality of the action together with density of the measure implies that there is an element s in G such that the closure of sD is contained in D. Since $\mu(s^n D) \to 0$, s must have a contracting fixed point z in D. The other element r needed to fit the requirement that z be resilient is easily found. Indeed, since the measure is non invariant, G is a free group in more that one generator. Moreover, using once more the density and non–invariance of the measure μ, there must be other points in the orbit of z which pass arbitrarily close to it and are different from it. □

The next proposition is a classical one. I mention it because it would be interesting to know whether it also holds true for distal actions.

Proposition 8.5. *If G is an equicontinuous group of transformations of X then the entropy $h(X, G_1) = 0$.*

Proof. It is evident from the definition that if G is a group of isometries of X, then the entropy $h(X, G_1) = 0$. Conversely, as in [1], let d_X be a distance function on

X and define

$$d(x,y)=\sup_G d_X(sx,sy).$$

Then d is also a distance function on X compatible with its topology and invariant under the action of G. □

The concept of entropy can be generalized to deal with potentials which affect the system under consideration. One instance of this is the notion of pressure that is used in classical dynamics [25]. We will consider a slightly more general version of it, because it has two interesting applications in our context.

Let $\varphi : X\times\mathbb{N} \to (0,\infty)$ be a map, and consider the following quantities associated to it and the system (X, G_1)

$$S(n,\epsilon,\varphi)=\sup\left\{\sum_{x\in E}\varphi(x,n)\mid E \text{ is } (n,\epsilon)\text{–separated}\right\}.$$

Note that since the function φ is positive, it suffices to take the supremum over maximal (n,ϵ)–separated sets. Let

$$S(\epsilon,\phi)=\limsup_{n\to\infty}\frac{1}{n}\log S(n,\epsilon,\varphi),$$

and observe that $S(\epsilon,\phi)$ is a non–increasing function of ϵ, so that the following limit exists:

$$h_S(X,\phi)=\lim_{\epsilon\to 0}S(\epsilon,\phi).$$

There is also a related definition involving spanning sets. Let

$$R(n,\epsilon,\varphi)=\inf\left\{\sum_{x\in E}\varphi(x,n)\mid E \text{ is } (n,\epsilon)\text{–spanning}\right\},$$

and $R(\epsilon,\varphi)=\limsup_{n\to\infty}(1/n)\log R(n,\epsilon,\varphi)$, and then define

$$h_R(X,\varphi)=\lim_{\epsilon\to 0}R(\epsilon,\varphi).$$

It always happens that $h_R(X,\varphi)\le h_S(X,\varphi)$ because a maximal separated set is also spanning of the same order (n,ϵ), hence $R(n,\epsilon,\varphi)\le S(n,\epsilon,\phi)$. We will not be concerned about whether these two quantities are equal.

The first application of the concept of entropy of a potential concerns the constructions of measures on X with given jacobian cocycle.

Let α be a continuous cocycle for the action of G on X, that is, a map $\alpha : G\times X\to(0,\infty)$ which satisfies

$$\alpha(rs,x)=\alpha(r,sx)\alpha(s,x)$$

for all r,s in G. The potential associated to the cocycle α is the function φ defined by

$$\varphi(\alpha,n,x)=\sup_{s\in G_n}\alpha(s,x)^{-1}.$$

We set

$$h(\alpha,\epsilon)=\limsup_{n\to\infty}\frac{1}{n}\log S(\alpha,n,\epsilon)$$

and

$$h(\alpha)=\lim_{\epsilon\to 0}h(\alpha,\epsilon).$$

This number $h(X, \alpha)$ is called the entropy of the cocycle α with respect to the generating set G_1. Clearly the value of the entropy depends on the generating set G_1, but if $h(\alpha, G_1) = 0$ for one choice of generating set, it is also 0 for any other choice.

The entropy of the potential associated to a cocycle α over (X, G_1) is finite if that of (X, G_1) is finite. Indeed, since α is continuous, there exists a constant $C = C(G_1, \alpha)$ such that if s is an element of G_n, then $\alpha(s, x) \leq C^n$. Hence $h(X, G_1) \leq h(\alpha) \leq h(X, G_1) + C$.

As we mentioned, the first application of entropy is to the construction of measures with given jacobian cocycle. The technique of construction of the measure follows the pattern of constructions of invariant measures utilized in [13].

Theorem 8.6. *Let α be a cocycle for the dynamical system (X, G_1) and suppose that the entropy $h(X, \alpha) = 0$. Then there is a measure μ on X whose jacobian cocycle under the action of G agrees with α (almost everywhere with respect to μ).*

For the proof, one constructs a family of operators $\Lambda_{n,\epsilon}$ on $C(X)$ which are defined as follows: if f is a continuous and non–negative function on X,

$$\Lambda_{n,\epsilon}(f) = \frac{1}{S(\alpha, n, \epsilon)} \sup_E \sum_{x \in E} f(x)\varphi(\alpha, n, x),$$

where the supremum is taken over all (n, ϵ)–separated sets E. The proof of the following fact is as in [13].

Proposition 8.7. *The operators $\Lambda_{n,\epsilon}$ satisfy the following properties:*

(1) *if $a \geq 0$, then $\Lambda_{n,\epsilon}(af) = a\Lambda_{n,\epsilon}(f)$,*
(2) *if $f_1 \leq f_2$, then $\Lambda_{n,\epsilon}(f_1) \leq \Lambda_{n,\epsilon}(f_2)$,*
(3) *$\Lambda_{n,\epsilon}(1) = 1$ and $\Lambda_{n,\epsilon}(f) \leq \sup_{x \in X} f(x)$,*
(4) *$\Lambda_{n,\epsilon}(f_1 + f_2) \leq \Lambda_{n,\epsilon}(f_1) + \Lambda_{n,\epsilon}(f_2)$, with equality if the distance between the supports of f_1 and f_2 is greater than ϵ.*

As in [13], the fact that the entropy of α is zero allows us to find a sequence $n(k, p)$ such that $\lim_{k \to \infty} n(k, p) = \infty$ and for which the limit

$$r_p = \lim_{k \to \infty} \frac{S(n(k, p) + 1, 1/p)}{S(\alpha, n(k, p), 1/p)}$$

exists and the iterated limit satisfies $\lim_{p \to \infty} r_p = 1$.

Moreover, the sequence $n(k, p)$ can be chosen so that the limits

$$\Lambda_{0,1/p}(f) = \lim_{k \to \infty} \Lambda_{n(k,p),1/p}(f)$$

and

$$\Lambda_{1,1/p}(f) = \lim_{k \to \infty} \Lambda_{n(k,p)+1,1/p}(f)$$

both exist, and a subsequence $p(l)$ can also be found so that

$$\Lambda_0(f) = \lim_{l \to \infty} \Lambda_{0,p(l)}(f)$$

$$\Lambda_1(f) = \lim_{l \to \infty} \Lambda_{1,p(l)}(f)$$

both exist.

The reasoning of [13] shows that these operators Λ_0 and Λ_1 define measures μ_0 and μ_1 on X, and that the choices of sequences made above imply that $\mu_0 = \mu_1$.

We call this common measure μ, and next verify that its jacobian cocycle agrees with the initial cocycle α, almost everywhere with respect to μ.

To begin with, we note that if $s \in G_1$, then

$$\begin{aligned}\varphi(\alpha, n, sx) &= \sup_{r \in G_n} \alpha(r, sx)^{-1} \\ &= \sup_{r \in G_n} \alpha(rs, x)^{-1}\alpha(s, x) \\ &\leq \alpha(s, x)\varphi(\alpha, n+1, x).\end{aligned}$$

It follows that

$$\Lambda_{n,\epsilon}(f \circ s) \leq \frac{S(\alpha, n+1, \epsilon)}{S(\alpha, n, \epsilon)} \Lambda_{n,\epsilon}(f\alpha(s^{-1}, \cdot)),$$

because if E is an (n, ϵ)–separated set, then sE is $(n+1, \epsilon)$–separated. Taking appropriate limits we have that

$$\int_X f(sx)\mu(x) \leq \int_X f(x)\alpha(s^{-1}, x)\mu(x).$$

Applying the same reasoning to the function $x \mapsto f \circ s(x)\alpha(s, x)^{-1}$ and the element s^{-1}, and noticing that $\alpha(s, s^{-1}x) = \alpha(s^{-1}, x)^{-1}$, we obtain the reverse inequality:

$$\int_X f(x)\alpha(s^{-1}, x)\mu(x) \leq \int f(sx)\mu(x).$$

Consequently the jacobian cocycle of the measure μ is

$$\frac{d(s^{-1})^*\mu}{d\mu}(x) = \alpha(s, x).$$

Our second application of this concept just introduced will be to the study of the relation between the topological entropy of (X, G_1) and that of $(\Lambda(X), \theta)$. To this end, we consider the following potential on X. For a point x in X and a non–negative integer n, let $\varphi(x, n)$ denote the number of paths of length $2n$ which start at x, where by a path of length $2n$ which starts at x we mean a map $c : \{0, 1, \cdots, 2n\} \to X$ such that $c(0) = x$ and $c(i) \in \{sc(i-1) \mid s \in G_1\}$ for $i = 1, \cdots, 2n$.

Before continuing, we state a useful trick:

Proposition 8.8. *Let (X, G_1) be a system on a compact metric space X with distance d_X. Then there exists a uniformly equivalent distance d on X under which all elements of G_1 are lipschitz maps of constant not larger than 2.*

Proof. Actually, what we will use is the explicit form of the distance function d. It is defined by

$$d(x, y) = \sum_{n=0}^{\infty} \frac{1}{2^n} \max_{s \in G_k} d_X(sx, sy).$$

This new distance on X is easily seen to be equivalent to the original one d_X, because the compactness of X implies that a finite collection of maps from X to itself is equicontinuous.

Incidentally, this observation can be used to exhibit Lipschitz maps of compact metric spaces with infinite entropy. □

Proposition 8.9. *The entropy of* $(\Lambda(X), \theta)$ *is the entropy of the potential* φ, *that is*

$$h(\Lambda(X), \theta) \leq h(X, G_1, \varphi).$$

Moreover, the entropies of X *and of* $\Lambda(X)$ *are simultaneously finite, in fact*

$$h(\Lambda(X), \theta) \leq h(X, G_1) + \log \operatorname{Card}(G_1).$$

Proof. Using the fact that the entropy of a system depends only on the uniform structure of the metrizable space X, and not on the particular distance used in the sequence of calculations leading to the final result (see [25] for a proof), we start with a distance d_X in X, which we assume to be of diameter one, and define the distance $d_{\Lambda(X)}$ in $\Lambda(X)$ and d on X as above.

The calculation of the entropy involves fixing ϵ and computing the limit superior of a sequence of numbers. Therefore, we may look at those terms $S(n, \epsilon)$ of the sequence with $1/2^n < \epsilon/2$. Also, for fixed ϵ, the quantities $R(n, \epsilon)$ and $R(n+1/\epsilon, \epsilon)$ behave asymptotically the same as $n \to \infty$ in the pertinent calculation.

Suppose that F is an $(n, \epsilon/2)$–spanning set for (X, G_1). For each x in F we can choose $\varphi(x, n)$ paths which start at x and whose initial $2n$–strings represent each of the $\varphi(x, n)$ paths of length $2n$ which start at x. We do this for each point of F and denote the resulting set of paths by $\widetilde{F}$. We claim that $\widetilde{F}$ is an (n, ϵ)–spanning set for $(\Lambda(X), \theta)$. Indeed, if w is a path of the form $w = (z, s_1 z, \cdots)$, then there is x in F such that $d(sx, sz) < \epsilon$ for every $s \in G_n$.

Now we write, for $0 \leq j \leq n$,

$$\begin{aligned} d_{\Lambda(X)}(\theta^j w, \theta^j w_x) &= \sum_{k=0} \frac{1}{2^k} d_X(w(k+j), w_x(k+j)) \\ &\leq \sum_{k=0}^{n} \frac{1}{2^k} d_X(w(k+j), w_x(k+j)) + \epsilon/2. \end{aligned}$$

By construction, we know that if $k \leq 2n$ and $w(k) = sz$, then also $w_x(k) = sx$. This implies, because the $(n, \epsilon/2)$ separation of F, that for $j \leq n$ and some $s \in G_n$, we have

$$\sum_{k=0}^{n} \frac{1}{2^k} d_X(w(k+j), w_x(k+j)) \leq \sum_{k=0}^{\infty} \frac{1}{2^k} \max_{G_k} d_X(rsz, rsz) \leq d(sx, sz) \leq \epsilon/2.$$

This shows that

$$\widetilde{R}(n, \epsilon) \leq \inf \left\{ \sum_{x \in F} \varphi(x, n) \mid F \text{ is } (n, \epsilon/2)\text{–spanning for } (X, G_1) \right\},$$

and the conclusion follows from this inequality. □

9. On the variational principle

In this section we make a few observation on the variational principle for the entropy of a system (X, G_1). If one attempts to build a category including the classical dynamical systems as well as the systems with harmonic measure that we have been studying, then there should be some variational principle which links the entropy of a system (X, G_1) with some of the other systems associated to it. We have seen such partial relation between the topological entropy of (X, G_1) and that of its associated path space $(\Lambda(X), \theta)$ in the previous section. In order to have a better understanding of this, it is best to place our discussion in terms of the greater path space $\Omega(X)$ and its shift transformation θ. Each harmonic measure μ for (X, G_1) induces the shift invariant measure $m = P \times \mu$ on $\Omega(X)$.

We fix a distance d_X on X, the obvious one in $\Omega(G)$, and consider the product distance on $\Omega(X)$. Then the projection $\phi : \Omega(X) \to \Omega(G)$ satisfies $\theta\phi = \phi\widetilde{\theta}$, and a theorem of Bowen [3] gives the following string of inequalities:

Proposition 9.1. *The following relation between topological entropies holds:*

$$h(X, G_1) \leq h(\Omega(X), \widetilde{\theta}) \leq h(\Omega(G), \theta) + \sup_{w \in \Omega(G)} h(w) \leq h(\Omega(G), \theta) + h(X, G_1).$$

Note that $h(\Omega(G), \theta) = \log \mathrm{Card}(G_1)$. The number $h(w)$ associated to the path $w \in \Omega(G)$ is the topological entropy of the map $\widetilde{\theta}$ relative to the fiber $\phi^{-1}(w)$. We make a few remarks on its calculation, as they will be needed later in the discussion.

If w is a path in G and $\mathcal{U}$ an open covering of X, let $N(w, \mathcal{U}, n)$ denote the minimum cardinality of a subcovering of $\mathcal{U} \vee w(1)^{-1}\mathcal{U} \vee \cdots (w(n-1)^{-1} \cdots w(1)^{-1})\mathcal{U}$, and let

$$h(w, \mathcal{U}, n) = \frac{1}{n} \log N(w, \mathcal{U}, n).$$

We note that $N(w, \mathcal{U}, n+m) \leq N(w, \mathcal{U}, n) + N(w, \mathcal{U}, m)$, for reasons similar to those used in the calculation of the entropy of a map. Therefore the limit

$$h(w, \mathcal{U}) = \lim_{n \to \infty} h(w, \mathcal{U}, n)$$

exists for each $w \in \Omega(G)$. Finally, let

$$h(w) = \sup_{\mathcal{U}} h(w, \mathcal{U}).$$

One can equally define the entropy along a path in terms of spanning sets, or in terms of separated sets, and prove that the three functions constructed are the same on $\Omega(G)$.

The facts concerning this function $h(w)$ that we need are as follows. First, for n and $\mathcal{U}$ fixed, $h(w, \mathcal{U}, n)$ depends only on the initial n–symbols of w. That is, $h(w, \mathcal{U}, n)$ is a locally constant function on $\Omega(G)$, hence continuous. Moreover, the convexity property of $\log N(w, \mathcal{U}, n)$ as a function of n implies that $h(w, \mathcal{U})$ is actually equal to $\inf_n h(w, \mathcal{U}, n)$. Therefore, the function $h(w, \mathcal{U})$ is upper semicontinuous on $\Omega(G)$ (and hence measurable).

The second observation concerns the relation between $h(w, \mathcal{U})$ and $h(\theta w, \mathcal{U})$. Directly from the definitions we have that

$$\begin{aligned} \log N(w, \mathcal{U}, n) &\leq \log N(U) + \log N(\theta w, w(1)^{-1}\mathcal{U}, n-1) \\ &= \log N(U) + \log N(\theta w, \mathcal{U}, n-1), \end{aligned}$$

where the last equality holds because $w(1)$ is a homeomorphism. It follows that $h(w, \mathcal{U}) \leq h(\theta w, \mathcal{U})$ is a shift super invariant function on $\Omega(G)$. It is therefore constant almost everywhere with respect to P (any P induced by a system of weights on G_1) because this measure is ergodic for the shift. Since a set of total measure with respect to some P is dense in $\Omega(G)$, the upper semicontinuity of $h(w, \mathcal{U})$ implies that this constant is independent of P. Summarizing

Proposition 9.2. *The function $h(w, \mathcal{U})$ is upper semicontinuous and constant almost everywhere on $\Omega(G)$ (with respect to P). This constant is in fact independent of the measure P on $\Omega(G)$. Moreover, $h(w) \leq h(X)$.*

Next we will find a formula for the entropy of the measure $m = P \times \mu$ on $\Omega(X)$. A path of length n in G is a sequence $\alpha = \alpha(n-1)\cdots\alpha(1)$ with $\alpha(k) \in G_1$. Let $\mathcal{P}_n$ denote the paths of length n in G. For a finite measurable partition ξ of X and a path α of length n, let $\alpha\xi$ denote the partition $\xi \vee \alpha(1)^{-1}\xi \vee \cdots \vee \alpha^{-1}\xi$ of X. Let $\widetilde{\xi}$ denote the partition of $\Omega(X)$ with atoms $\Omega(G)s \times A$, where $s \in G_1$ and $A \in \xi$. To compute the entropy of $(\Omega(X), \theta)$ it is enough to consider partitions of $\Omega(X)$ of the form $\widetilde{\xi}$. Moreover, an element B of $\widetilde{\xi} \vee \theta^{-1}\widetilde{\xi} \vee \cdots \vee \theta^{-n}\widetilde{\xi}$ has the form

$$B = \Omega(G)\alpha \times A$$

where $A \in \alpha\xi$. Therefore, for a partition ξ we let

$$h(X, G_1, \xi, n) = \sum_{\alpha \in \mathcal{P}_n} q(\alpha) \int_X \log \mu(\alpha\xi(x))\mu(x)$$

where $q(\alpha)$ is the weight of α and $\alpha\xi(x)$ denotes the atom of the partition $\alpha\xi$ which contains the point x.

Then define

$$H_\mu(X, G_1, \xi) = \lim_{n\to\infty} \frac{1}{n} H_\mu(X, G_1, n),$$

and

$$H_\mu(X, G_1) = \sup_\xi H_\mu(X, G_1, \xi).$$

Therefore, we have

$$H_\mu(X, G_1) = H_m(\Omega(X), \theta) - \sum_{s \in G_1} q(s) \log q(s).$$

One may expect that it could be possible to obtain a variational principle for the entropy using this calculation. The number $H_\mu(X, G_1)$ is the P–average of the metric entropy along paths in G.

As I realized when this paper was being written, there is a relative variational principle which when applied to our situation gives the following:

Proposition 9.3. *Let (X, G_1) be a dynamical system and P a measure on $\Omega(G)$ associated to a system of weights. Then*

$$\sup_{\nu; \phi\nu = P} H_\nu(\Omega(X), \widetilde{\theta}) = \sum_{s \in G_1} q(s) \log q(s) + \int_{\Omega(G)} h(w)P(w).$$

This is proven in [18] (see also [20]). Note that in [18] there is the implicit assumption of invertibility of the transformations involved. This assumption is removed in [2].

Since not every invariant measure for $(\Omega(X), \theta)$ is of the form $P \times \mu$ with μ harmonic for (X, G_1), we only obtain the inequality

$$H_\mu(X, G_1) \leq \int h(w)P(w).$$

I do not know whether this is actually an equality. Perhaps it is in case that the shift map is expansive, but I will not attempt to discuss it now. My contribution here would be to show that this notion of metric entropy is not related to the topological entropy discussed in the previous section, apart from the obvious inequality. That is, we exhibit an example of a dynamical system (X, G_1) which has positive topological entropy, but the entropy along every path in G is zero.

Let X be a compact metric space and G be a group of homeomorphisms of X with the property that the topological entropy of each element of G is zero. For example, X could be the unit circle, or it could be the boundary ∂G of G with the canonical action.

In such a situation, it is easy to see that the entropy along every path in G is zero. Indeed, that the entropy of every element of G is zero implies that the entropy along every periodic path in G is zero. Therefore $h(w, \mathcal{U}) = 0$ for every covering $\mathcal{U}$ of X and every periodic path w. If $h(w)$ were positive on a subset of $\Omega(G)$ of total P–measure, then we could find a covering $\mathcal{U}$ for which $h(w, \mathcal{U})$ on a set of positive measure, hence almost everywhere. But using the previously proven fact that this function $h(w, \mathcal{U})$ is upper semicontinuous on $\Omega(G)$, it would be positive everywhere. This contradicts that it is zero on the periodic paths.

Finally, if we take X to be the circle, and G_1 to consist of two elements s and r. One of then, say r is a rotation. The other is a hyperbolic element of $\mathrm{PS}(2, \mathbb{R})$ acting on the circle. This system has a resilient point, hence its topological entropy is positive. Similar considerations apply to the system $(\partial G, G_1)$.

References

[1] R. Adler, A. Konheim and M. McAndrew, *Topological entropy*, Trans. Amer. Math. Soc. **114** (1965), 309–319.

[2] T. Bogenschütz and H. Crauel, *The Abramov–Rokhlin formula*, Ergodic theory and related topics, III. Lecture Notes in Math. 1514, Springer–Verlag, Berlin, 1992, 32–35.

[3] R. Bowen, *Entropy for group endomorphisms and homogeneous spaces*, Trans. Amer. Math. Soc. **153** (1971), 401–414.

[4] P. Cartier, *Fonctions harmoniques sur un arbre*, Symp. Math. **9** (1972), 203–270.

[5] K. L. Chung, *A course in probability theory*, Academic Press, San Diego, 1974.

[6] N. Dunford and J. Schwartz, *Linear Operators. Part I*, Interscience, New York, 1958.

[7] J. Feldman and C. Moore, *Ergodic equivalence relations, cohomology, and von Neumann algebras*, Trans. Amer. Math. Soc. **234** (1977), 289–324.

[8] A. Figà–Talamanca and M. Picardello, *Harmonic Analysis on Free Groups*, Lecture Notes in Pure and Applied Mathematics, vol. 87, Marcel Dekker, New York, 1983.

[9] S. V. Fomin, *On measures invariant under a group of transformations*, Izv. Akad. Nauk SSSR, Ser. Mat. **14** (1950), 261–274 (Russian), English transl. in Amer. Math. Soc. Transl. (2) **51** (1966), 317–332.

[10] H. Furstenberg, *Poincaré recurrence and number theory*, Bull. Amer. Math. Soc. **5** (1981), 211–234.

[11] L. Garnett, *Foliations, the ergodic theorem and brownian motion*, J. Funct. Anal. **51** (1983), 285–311.

[12] E. Ghys and P. de la Harpe, *Sur les Groups Hyperboliques d'après Gromov*, Progress in Mathematics **83**, Birkhäuser, Boston, 1990.

[13] E. Ghys, R. Langevin and P. Walczak, *Entropie des feuilletages*, Acta Math. **160** (1988), 107–142.

[14] P. Halmos, *Lectures on Ergodic Theory*, Chelsea, New York, 1953.

[15] A. Kolmogorov, *Foundations of the Theory of Probability*, Chelsea, New York, 1956.

[16] W. Krieger, *On ergodic flows and the isomorphism of factors*, Math. Ann. **223** (1976), 19–70.

[17] N. Kriloff and N. Bogoliuboff, *La théorie générale de la measure dans son application à l'étude des systèmes dynamiques de la mécanique non linéare*, Annals of Math. **38** (1937), 65–113.

[18] F. Ledrappier and P. Walters, *A relativised variational principle for continuous transformations*, J. London Math. Soc. **16** (1977), 568–576.

[19] D. Maharam, *Incompressible transformations*, Fund. Math. **56** (1964), 35–50.

[20] M. Nakamura, *Invariant measures and entropies for random dynamical systems and the variatonal principle for random Bernouilli shifts*, Hiroshima Math. J. **21**, 187–216.

[21] V. Nemytskii and V. Stepanov, *Qualitative Theory of Differential Equations*, Princeton University Press, Princeton, 1960.
[22] J. Oxtoby, *Ergodic sets*, Bull. Amer. Math. Soc. **58** (1952), 116–136.
[23] R. Phelps, *Lectures on Choquet's Theorem*, Van Nostrand, New York, 1966.
[24] K. Schmidt, *On recurrence*, Z. Wahrscheinlichkeitctheorie verw. Gebiete **68** (1984), 75–95.
[25] P. Walters, *An Introduction to Ergodic Theory*, Springer–Verlag, New York, 1982.
[26] W. Woess, *Random walks on infinite graphs and groups*, Bull. London Math. Soc. **26** (1994), 1–60.
[27] Zimmer, R. *Ergodic theory and semisimple groups*, Monographs in Mathematics 81, Birkhäuser, Boston, 1984.

Department of Mathematics, 253–37. California Institute of Technology, Pasadena, CA 91125.