FUNDAMENTA
MATHEMATICAE
206 (2009)

Predictability, entropy and information of infinite transformations

by

Jon Aaronson (Tel Aviv) and **Kyewon Koh Park** (Suwon)

Abstract. We show that a certain type of quasifinite, conservative, ergodic, measure preserving transformation always has a maximal zero entropy factor, generated by predictable sets. We also construct a conservative, ergodic, measure preserving transformation which is not quasifinite; and consider distribution asymptotics of information showing that e.g. for Boole's transformation, information is asymptotically mod-normal with normalization $\propto \sqrt{n}$. Lastly, we show that certain ergodic, probability preserving transformations with zero entropy have analogous properties and consequently entropy dimension of at most 1/2.

0. Introduction. Let $(X, \mathcal{B}, m, T)$ be a conservative, ergodic, measure preserving transformation and let $\mathcal{F} := \{F \in \mathcal{B} : m(F) < \infty\}$. Call a set $A \in \mathcal{F}$ *T-predictable* if it is measurable with respect to its own past in the sense that $A \in \sigma(\{T^{-n}A : n \geq 1\})$ (the σ-algebra generated by $\{T^{-n}A : n \geq 1\}$) and let $\mathcal{P} = \mathcal{P}_T := \{T\text{-predictable sets}\}$.

If $m(X) < \infty$, Pinsker's theorem ([Pi]) says that

- $\mathcal{P}_T$ is the *maximal*, *zero-entropy factor algebra*,

i.e. $\mathcal{P} \subset \mathcal{B}$ is a factor algebra (T-invariant sub-σ-algebra), $h(T, \mathcal{P}) = 0$ (see §1) and if $\mathcal{C} \subset \mathcal{B}$ is a factor algebra with $h(T, \mathcal{C}) = 0$, then $\mathcal{C} \subseteq \mathcal{P}$. $\mathcal{P}$ is also known as the *Pinsker algebra* of $(X, \mathcal{B}, m, T)$.

When $(X, \mathcal{B}, m, T)$ is a conservative, ergodic, measure preserving transformation with $m(X) = \infty$, the above statement fails and indeed $\sigma(\mathcal{P}) = \mathcal{B}$: Krengel has shown ([K2]) that

- For all $A \in \mathcal{F}$ and $\epsilon > 0$, there exists $B \in \mathcal{F}$ with $m(A \triangle B) < \epsilon$ which is a *strong generator* in the sense that $\sigma(\{T^{-n}B : n \geq 1\}) = \mathcal{B}$, whence $\sigma(\mathcal{P}_T) = \mathcal{B}$.

2010 *Mathematics Subject Classification*: 37A40, 60F05.
Key words and phrases: measure preserving transformation, conservative, ergodic, entropy, quasifinite, predictable set, entropy dimension.

DOI: 10.4064/fm206-1-1

It is not known if there is always a maximal, zero-entropy factor algebra (in case there is some zero-entropy factor algebra).

We recall the basic properties of entropy in §1 and define the class of *log lower bounded* conservative, ergodic, measure preserving transformations in §2.

These are quasifinite in the sense of [K1] and are discussed in §2 in this context where also examples are constructed, including a conservative, ergodic, measure preserving transformation which is not quasifinite.

A log lower bounded conservative, ergodic, measure preserving transformation with some zero-entropy factor algebra has a maximal, zero-entropy factor algebra generated by a specified hereditary subring of predictable sets (see §5).

We obtain information convergence (in §4) for quasifinite transformations (cf. [KS]).

For quasifinite, pointwise dual ergodic transformations with regularly varying return sequences, we obtain (in §6) distributional convergence of information. Lastly, we construct a probability preserving transformation with zero entropy with analogous distributional properties and estimate its entropy dimension in the sense of [FP]. This example is unusual in that it has a generator with information function asymptotic to a nondegenerate random variable (the range of Brownian motion).

1. Entropy. We recall the basic entropy theory of a probability preserving transformation $(\Omega, \mathcal{A}, P, S)$. Let $\alpha \subset \mathcal{A}$ be a countable partition.

The *entropy* of α is

$$H(\alpha) := \sum_{a\in\alpha} P(a)\log\frac{1}{P(a)}.$$

The *S-join of α from k to l* (for $k < l$) is

$$\alpha_k^l(S) := \Big\{\bigcap_{j=k}^{l} S^{-j}a_j : a_k, a_{k+1}, \ldots, a_l \in \alpha\Big\}.$$

By subadditivity, the limit

$$\lim_{n\to\infty}\frac{1}{n}H(\alpha_0^{n-1}(S)) =: h(S,\alpha)$$

exists (the *entropy* ([1]) *of S with respect to α*). The *entropy of S with respect to the factor algebra* (S-invariant σ-algebra) $\mathcal{C} \subset \mathcal{A}$ is

$$h(S,\mathcal{C}) := \sup_{\alpha\subset\mathcal{C}} h(S,\alpha).$$

([1]) Mean entropy rate.

By the generator theorem, if α is a partition, then $h(S,\alpha) = h(S,\sigma(\{S^n\alpha : n \in \mathbb{Z}\}))$.

The *information* of the countable partition $\alpha \subset \mathcal{A}$ is the function $I(\alpha) : \Omega \to \mathbb{R}$ defined by

$$I(\alpha)(x) := \log \frac{1}{P(\alpha(x))}$$

where $\alpha(x) \in \alpha$ is defined by $x \in \alpha(x) \in \alpha$. Evidently

$$H(\alpha) = \int_{\Omega} I(\alpha)\, dP.$$

Convergence of information is given by the celebrated Shannon–McMillan–Breiman theorem (see [S], [M], [Br]), the statement ($\mathfrak{I}$) here being due to Chung [C] (see also [IT]).

Let $(\Omega, \mathcal{A}, P, S)$ be an ergodic probability preserving transformation and let α be a partition with $H(\alpha) < \infty$. Then

$$(\mathfrak{I}) \qquad \frac{1}{n} I(\alpha_1^N(S)) \to h(S,\alpha) \quad \text{a.s. as } n \to \infty;$$

equivalently $P(\alpha_1^N(S)(x)) = e^{-nh(S,\alpha)(1+o(1))}$ for a.e. $x \in \Omega$ as $n \to \infty$ where $x \in \alpha_1^N(S)(x) \in \alpha_1^N(s)$.

We will need *Abramov's formula* for the entropy of an induced transformation of an ergodic probability preserving transformation $(\Omega, \mathcal{A}, P, S)$:

$$h(S_A) = \frac{1}{P(A)} h(S) \quad \forall A \in \mathcal{A}$$

where $S_A : A \to A$ is the *induced transformation* on A defined by

$$S_A x := S^{\varphi_A(x)} x, \quad \varphi_A(x) := \min\{n \ge 1 : S^n x \in A\} \quad (x \in A).$$

Abramov's formula can be proved using convergence of information (see [Ab] and §4 here).

Krengel entropy. Suppose that $(X, \mathcal{B}, m, T)$ is a conservative, ergodic, measure preserving transformation. Then using Abramov's formula (as shown in [K1]) we obtain

$$m(A)h(T_A) = m(B)h(T_B) \quad \forall A, B \in \mathcal{F} := \{F \in \mathcal{B} : 0 < m(F) < \infty\}.$$

Then $\underline{h}(T) := m(A)h(T_A)$ (for any $A \in \mathcal{B}$ with $0 < m(A) < \infty$) is the *Krengel entropy of T*. More generally, the *Krengel entropy of T with respect to the factor* (i.e. σ-finite, T-invariant sub-σ-algebra) $\mathcal{C} \subset \mathcal{B}$ is

$$\underline{h}(T, \mathcal{C}) := m(A)h(T_A, \mathcal{C} \cap A) \quad (A \in \mathcal{C},\ 0 < m(A) < \infty).$$

Another definition of entropy is given in [Pa].

It is shown in [Pa] that for *quasifinite* (see §2 below) conservative, ergodic, measure preserving transformations, the two entropies coincide.

2. Quasifiniteness and log lower boundedness

Quasifiniteness. Let $(X, \mathcal{B}, m, T)$ be a conservative, ergodic, measure preserving transformation. Recall from [K1] that a set $A \in \mathcal{F}$ is called *quasifinite* (qf) if $H_A(\rho_A) < \infty$ where $\rho_A := \{A \cap T^{-n}A \setminus \bigcup_{j=1}^{n-1} T^{-j}A : n \geq 1\}$, and that T is so called if there exists such a set. As shown in Proposition 7.1 in [K1],

- for $A \in \mathcal{F}$ quasifinite, $A \in \mathcal{P}_T \Leftrightarrow h(T_A, \rho_A) = 0$.

There are conservative, ergodic, measure preserving transformations which are not quasifinite. An unpublished example by Ornstein is mentioned in [K2, p. 82].

Here we construct a conservative, ergodic, measure preserving transformation with no quasifinite extension. To do this we first establish a saturation property for the collection of quasifinite sets:

Proposition 2.0. *Suppose that $(X, \mathcal{B}, m, T)$ is a conservative, ergodic, quasifinite, measure preserving transformation. Then for every $F \in \mathcal{F}$ there exists $A \in \mathcal{B} \cap F$ such that $m(A) > 0$ and each $B \in \mathcal{B} \cap A$ is quasifinite.*

Proof. We first show that

¶1 *if $F \in \mathcal{F}$ is quasifinite, then for every $\epsilon > 0$ there exists $A \in \mathcal{B} \cap F$ such that $m(F \setminus A) < \epsilon$ and each $B \in \mathcal{B} \cap A$ is quasifinite.*

By $(\mathfrak{I})$, $n^{-1}I(\rho_F)_0^{n-1}(T_F)) \to h(T_F, \rho_F)$ a.e. as $n \to \infty$. By Egorov's theorem, there exists $A \in \mathcal{B} \cap F$ such that $m(F \setminus A) < \epsilon$ and the convergence is uniform on A.

For $B \in \mathcal{B} \cap A$, let $N_{n,B} := \#\{a \in (\rho_F)_0^{n-1}(T_F) : m(a \cap B) > 0\}$ (where $\#F$ means the number of elements in the set F). Then $N_{n,B} = e^{nh(T_F,\rho_F)(1+o(1))}$ as $n \to \infty$.

Define $\psi : B \to \mathbb{N}$ by $\psi(x) := \min\{n \geq 1 : T_F^n x \in B\}$. Then

$$\int_B \psi \, dm = \sum_{n=1}^{\infty} nm([\psi = n]) = m(F) < \infty \quad \text{(by Kac's formula)},$$

$$\varphi_B(x) = \sum_{j=0}^{\psi(x)-1} \varphi_F(T_F^j x),$$

whence

$$\rho_B \prec \gamma_B := \bigcup_{n=1}^{\infty} \{[\psi = n] \cap a : a \in (\rho_F)_0^{n-1}(T_F)\}.$$

Thus

$$H_{m_B}(\rho_B) \le H_{m_B}(\gamma_B) = \sum_{n=1}^{\infty} m_B([\psi = n]) H_{m_{[\psi=n]}}((\rho_F)_0^{n-1}(T_F))$$
$$\le \sum_{n=1}^{\infty} m_B([\psi = n]) \log N_{n,B} < \infty$$

because $\log N_{n,B} \sim nh(T_F, \rho_F)$, proving ¶1.

To complete the proof, let $F \in \mathcal{F}$. Suppose that $Q \in \mathcal{F}$ is quasifinite. Then evidently so is $T^{-n}Q$ for all $n \ge 1$. By ergodicity, there exists $n \ge 1$ such that $m(F \cap T^{-n}Q) > 0$. By ¶1, there exists $G \in \mathcal{B} \cap T^{-n}Q$ such that $m(T^{-n}Q \setminus G) < \epsilon := m(F \cap T^{-n}Q)/9$ and each $B \in \mathcal{B} \cap G$ is quasifinite. The set $A = G \cap F$ is as required. ∎

EXAMPLE 2.1. Let $(X_0, \mathcal{B}_0, m_0, T_0)$ be the conservative, ergodic, measure preserving transformation defined as in [Fr] by the cutting and stacking construction

$$B_0 = 1, \qquad B_n = \bigoplus_{k=1}^{N_n} B_{n-1} 0^{L_{n,k}}$$

where $N_n,\ L_{n,k},\ 1 \le k \le N_n$, satisfy

$$N_{n+1} \ge e^{nN_1 \dots N_n}, \qquad L_{n,k+1} > \sum_{j=1}^{k} L_{n,j} + kh_{n-1},$$

where $h_n := |B_n|$.

PROPOSITION 2.1. *No extension T of the conservative, ergodic, measure preserving transformation T_0 defined in Example 2.1 is quasifinite.*

Proof. Suppose (without loss of generality) that $(X, \mathcal{B}, m, T)$ is a conservative, ergodic extension of T_0 and that $F \in \mathcal{F}$ is quasifinite. Then evidently so is $T^n F$ for all $n \ge 1$. By Proposition 2.0 there exists $A \in \mathcal{B}$ quasifinite with $A \subset B_0$. We will contradict this (and therefore the assumption that a quasifinite $F \in \mathcal{F}$ exists).

¶1 Write $B_n = \bigcup_{j=0}^{h_n-1} T^j b_n$ where $b_n \subset B_0,\ m(b_n) = 1/N_1 \dots N_n$ and

$$B_n = \biguplus_{k=1}^{N_{n+1}} B_n^{(k)} = \biguplus_{k=1}^{N_{n+1}} T^{\kappa(n+1,k)} B_n^{(1)}$$

where $\kappa(n+1,1) = 0$ and $\kappa(n+1,k) = (k-1)|B_n| + \sum_{j=1}^{k-1} L_{n+1,j}$ (i.e. the $B_n^{(k)}$ $(1 \le k \le N_{n+1})$ are the subcolumns of B_n appearing in B_{n+1}).

¶2 For $n \ge 1$, let $\mathfrak{k}_n := \{0 \le j \le h_n - 1 : T^j b_n \subset B_0\}$. Then $B_0 = \biguplus_{j \in \mathfrak{k}_n} T^j b_n$, $|\mathfrak{k}_n| = N_1 \dots N_n$ and $\{T_{B_0}^k x\}_{k=0}^{N_1 \dots N_n - 1} = \{T^j x : j \in \mathfrak{k}_n\}$ for $x \in b_n$.

¶3 Fix $0 < \epsilon < 1/3$ and let

$$b_{n,\epsilon} := \left\{x \in b_{n+1} : \left|\frac{1}{|\mathfrak{k}_{n+1}|}\sum_{k\in\mathfrak{k}_{n+1}} 1_A(T^k x) - m(A)\right| < \epsilon m(A)\right\}.$$

By ¶2 above, for $x \in b_{n+1}$,

$$\frac{1}{|\mathfrak{k}_{n+1}|}\sum_{k\in\mathfrak{k}_{n+1}} 1_A(T^k x) = \frac{1}{N_1 \dots N_{n+1}} \sum_{k=0}^{N_1 N_2 \dots N_n - 1} 1_A(T_{B_0}^k x)$$

and a standard argument using the ergodic theorem for T_{B_0} shows that there exists M so that $m(b_{n,\epsilon}) > (1-\epsilon)m(b_{n+1})$ for all $n \ge M$.

¶4 Fix $n \ge M$ and $x \in b_{n+1}$, and let $\mathfrak{k}_{A,n,x} := \{k \in \mathfrak{k}_{n+1} : T^k x \in A\}$ and $A_{n,x} := \{T^j x\}_{j\in\mathfrak{k}_{A,n,x}}$. Then for $x \in b_{n,\epsilon}$,

$$\#\{1 \le k \le N_{n+1} : A_{n,x} \cap B_n^{(k)} \ne \emptyset\} \ge (1-e)m(A)\frac{|\mathfrak{k}_{n+1}|}{h_n} = (1-\epsilon)m(A)N_{n+1}.$$

For $n \ge M$ and $x \in b_{n,\epsilon}$, write

$$\{1 \le k \le N_{n+1} : A_{n,x} \cap B_n^{(k)} \ne \emptyset\} =: \{\kappa_i(x) : 1 \le i \le \nu\}$$

where $\nu - 1 > (1-\epsilon)m(A)N_{n+1}$ and $\kappa_i(x) < \kappa_{i+1}(x)$ for all i.

For $1 \le i \le \nu$, let $\mathfrak{k}_{A,n,x}^{(i)} := \{k \in \mathfrak{k}_{n+1} : T^k b_{n+1} \subset A_{n,x} \cap B_n^{(\kappa_i)}\}$ and let $\underline{m}_i := \min \mathfrak{k}_{A,n,x}^{(i)}$, $\overline{m}_i := \max \mathfrak{k}_{A,n,x}^{(i)}$; $y_i := \overline{m}_{i+1} - \underline{m}_i$ $(1 \le i \le \nu - 1)$. Note that

$$y_i \le \sum_{j=1}^{\kappa_i} L_{n+1,j} + \kappa_i h_n < L(n+1, \kappa_i + 1) \le L(n+1, \kappa_{i+1}) \le y_{i+1}.$$

¶5 For $K \subset \mathfrak{k}_{n+1}$, let $a_K := \{x \in b_{n+1} : \mathfrak{k}_{A,n,x} = K\}$ and

$$\beta_n := \{a_K : K \subset \mathfrak{k}_{n+1}\}, \qquad \alpha_n := \Big\{\widehat{a} := \bigcup_{j\in\mathfrak{k}_n} T^j a : a \in \beta_n\Big\}.$$

For $a \in \beta_n,\ a \subset b_{n,\epsilon}$, and $1 \le i \le \nu - 1$, we have

$$m(a \cap [\varphi_A = y_i(a)]) = \frac{m(a)}{N_1 \dots N_{n+1}}.$$

Thus

$$H(\rho_A) \ge H(\rho_A \| \alpha_n)$$

$$\ge \sum_{a\in\beta_n,\, a\subset b_{n,\epsilon}} m(a) \sum_{i=1}^{\nu-1} m([\varphi_A = y_i(a)]|a) \log \frac{1}{m([\varphi_A = y_i(a)]|a)}$$

$$\geq m(\widehat{b}_{n,\epsilon})\,\frac{(\nu-1)\log(N_{n+1})}{N_1\dots N_{n+1}}$$
$$\geq (1-\epsilon)^2 m(A)\frac{\log N_{n+1}}{N_1N_2\dots N_n}$$
$$> (1-\epsilon)^2 m(A) n \uparrow \infty. \blacksquare$$

Log lower boundedness. For $(X,\mathcal{B},m,T)$ a conservative, ergodic, measure preserving transformation, set

$$\mathcal{F}_{\log,T} := \Big\{A\in\mathcal{B} : 0<m(A)<\infty,\ \int_A \log\varphi_A\,dm<\infty\Big\}.$$

Note that $\mathcal{F}_{\log,T}\subset\{\text{quasifinite sets}\}$, because

$$(\bigstar)\qquad p_n\geq 0,\ \sum_{n=1} p_n\log n<\infty\ \Rightarrow\ \sum_{n=1} p_n\log\frac{1}{p_n}<\infty.$$

Call T *log lower bounded* (LLB) if $\mathcal{F}_{\log,T}\neq\emptyset$.

PROPOSITION 2.2.

(i) *T is LLB iff*

$$\frac{1}{\log n}\sum_{k=0}^{n-1} f\circ T^n\to\infty \qquad \text{a.e. as } n\to\infty$$

for some and hence all $f\in L^1(m)_+ := \{f\in L^1 : f\geq 0,\ \int_X f\,dm>0\}$.

(ii) *T is not LLB iff*

$$\liminf_{n\to\infty}\frac{1}{\log n}\sum_{k=0}^{n-1} f\circ T^n=0 \qquad \text{a.e. for some and hence all } f\in L^1_+.$$

(iii) *If $(X,\mathcal{B},m,T)$ is LLB and $\mathcal{C}\subset\mathcal{B}$ is a factor, then $\mathcal{C}\cap\mathcal{F}_{\log,T}\neq\emptyset$.*

(iv) *$\mathcal{F}_{\log,T}$ is a hereditary ring.*

Proof. Statements (i) and (ii) follow from Theorem 2.4.1 in [A], and (iii) follows from these. We prove (iv).

Suppose that $A\in\mathcal{F}_{\log,T}$, $B\in\mathcal{B}$, and $B\subset A$. Then

$$\varphi_B(x)=\sum_{k=0}^{\psi(x)-1}\varphi_A(T_A^k x)\qquad (x\in B)$$

where $\psi : B\to\mathbb{N}$, $\psi(x) := \min\{n\geq 1 : T_A^n x\in B\}$.

By the Kac formula,

$$\int_B\sum_{k=0}^{\psi-1} f\circ T_A^k\,dm=\int_A f\,dm\qquad \forall f\in L^1(m).$$

To see that $B \in \mathcal{F}_{\log,T}$, we use this and $\log(k+l) \le \log(k) + \log(l)$:

$$\int_B \log \varphi_B \, dm = \int_B \log\Big(\sum_{k=0}^{\psi-1} \varphi_A \circ T_A^k\Big)\, dm$$
$$\le \int_B \sum_{k=0}^{\psi-1} \log(\varphi_A \circ T_A^k)\, dm = \int_A \log \varphi_A \, dm < \infty.$$

Suppose that $A, B \in \mathcal{F}_{\log,T}$. Then $\varphi_{A\cup B} \le 1_A \varphi_A + 1_B \varphi_B$, whence

$$\int_{A\cup B} \log(\varphi_{A\cup B})\, dm = \int_A \log(\varphi_{A\cup B})\, dm + \int_B \log(\varphi_{A\cup B})\, dm$$
$$\le \int_A \log(\varphi_A)\, dm + \int_B \log(\varphi_B)\, dm < \infty. \ \blacksquare$$

3. Examples of LLB transformations

Pointwise dual ergodic transformations. A conservative, ergodic, measure preserving transformation $(X, \mathcal{B}, m, T)$ is called *pointwise dual ergodic* if there is a sequence of constants $(a_n(T))_{n\ge 1}$ (called the *return sequence* of T) so that

$$\frac{1}{a_n(T)} \sum_{k=0}^{n-1} \widehat{T}^k f \to \int_X f\, dm \quad \text{a.e. for some (and hence all) } f \in L^1(m)_+$$

where $\widehat{T} : L^1(m) \to L^1(m)$ is the *transfer operator* defined by

$$\int_A \widehat{T} f\, dm = \int_{T^{-1}A} f dm \quad (f \in L^1(m),\ A \in \mathcal{B}).$$

See [A, 3.8].

Proposition 3.1. *Let $(X, \mathcal{B}, m, T)$ be a pointwise dual ergodic, conservative, ergodic, measure preserving transformation. Then*

$$T \text{ is } LLB \ \Leftrightarrow\ \sum_{n=1}^{\infty} \frac{1}{n a_n(T)} < \infty.$$

Proof. Let $A \in \mathcal{F}$ be a *uniform set* in the sense that for some $f \in L^1(m)_+$,

$$\frac{1}{a_n(T)} \sum_{k=0}^{n-1} \widehat{T}^k f \to \int_X f\, dm \quad \text{uniformly on } A.$$

By Lemma 3.8.5 in [A],

$$\int_A (\varphi_A \wedge n)\, dm = m\Big(\bigcup_{k=0}^{n} T^{-k} A\Big) \asymp \frac{n}{a_n(T)},$$

whence

$$A \in \mathcal{F}_{\log} \Leftrightarrow \sum_{n=1}^{\infty} \frac{m(\bigcup_{k=0}^{n} T^{-k}A)}{n} < \infty \Leftrightarrow \sum_{n=1}^{\infty} \frac{1}{n a_n(T)} < \infty. \blacksquare$$

REMARKS. 1) For example, the simple random walk on $\mathbb{Z}$ is LLB (because $a_n(T) \propto \sqrt{n}$), whereas the simple random walk on $\mathbb{Z}^2$ is not LLB (because $a_n(T) \propto \log n$).

2) It is not known whether the simple random walk on $\mathbb{Z}^2$ is quasifinite, or even has a factor with finite entropy.

EXAMPLE 3.2. *There is a quasifinite, conservative, ergodic, Markov shift* $(X, \mathcal{B}, m, T)$ *with* $a_n(T) \asymp \sqrt{\log n}$.

Note that by Proposition 3.1, this T is not LLB.

Proof. Let $f_{4^{4^n}} := 1/2^n$ for $n \geq 1$ and $f_k := 0$ for $k \in \mathbb{N} \setminus 4^{4^{\mathbb{N}}}$. Then $f \in \mathcal{P}(\mathbb{N})$.

Let $\Omega := \mathbb{N}^{\mathbb{Z}}$ and let $P = f^{\mathbb{Z}} \in \mathcal{P}(\Omega, \mathcal{B}(\Omega))$ be the product measure. Then $(\Omega, \mathcal{B}(\Omega), P, S)$ is an ergodic, probability preserving transformation where $S : \Omega \to \Omega$ is the shift.

Define $\varphi : \Omega \to \mathbb{N}$ by $\varphi(\omega) := \omega_0$ and let $(X, \mathcal{B}, m, T)$ be the tower over $(\Omega, \mathcal{B}(\Omega), P, S)$ with height function φ. It follows that $(X, \mathcal{B}, m, T)$ is a conservative, ergodic, Markov shift with $a_n(T) \asymp \sum_{k=0}^{n} u_k$ where u is defined by the renewal equation: $u_0 = 1,\ u_n = \sum_{k=1}^{n} f_k u_{n-k}$.

To see that $(X, \mathcal{B}, m, T)$ is quasifinite, we check that Ω is quasifinite. Indeed,

$$H_{\Omega}(\rho_{\Omega}) = \sum_{k\geq 1,\, f_k>0}^{\infty} f_k \log \frac{1}{f_k} = \sum_{n=1}^{\infty} \frac{n \log 2}{2^n} < \infty.$$

To estimate $a_n(T)$, recall that by Lemma 3.8.5 in [A], $a_n(T) \asymp n/L(n)$ where

$$L(n) := m\Big(\bigcup_{k=0}^{n} T^{-k}\Omega\Big) = \sum_{k=0}^{n} \sum_{l=k+1}^{\infty} f_l.$$

Now,

$$\sum_{l=k+1}^{\infty} f_l = \sum_{n>\log_4 \log_4 k} \frac{1}{2^n} \asymp \frac{1}{2^{\log_4 \log_4 k}} = \frac{1}{\sqrt{\log_4 k}}.$$

Thus $L(n) \asymp n/\sqrt{\log n}$ and $a_n(T) \asymp \sqrt{\log n}$. $\blacksquare$

The Hajian–Ito–Kakutani transformations. Let $\Omega = \{0, 1\}^{\mathbb{N}}$, $\ell(\omega) := \min\{n \geq 1 : \omega_n = 0\}$ and let $\tau : \Omega \to \Omega$ be the *adding machine* defined by

$$\tau(1, \ldots, 1, 0, \omega_{\ell(\omega)+1}, \ldots) := (0, \ldots, 0, 1, \omega_{\ell(\omega)+1}, \ldots).$$

For $p \in (0,1)$, define $\mu_p \in \mathcal{P}(\Omega)$ by $\mu_p([a_1,\dots,a_n]) := p_{a_1}\dots p_{a_n}$ where $p_0 := 1-p,\ p_1 := p$. It follows that $(\Omega,\mathcal{A},\mu_p,\tau)$ is an ergodic, nonsingular transformation with $\frac{d\mu_p\circ\tau}{d\,\mu_p} = \left(\frac{1-p}{p}\right)^{\phi}$ where $\phi := \ell - 2$.

Now let $X := \Omega\times\mathbb{Z}$ and define $T : X \to X$ by $T(x,n) = (\tau x, n+\phi(x))$. For $p\in(0,1)$, define $m_p \in \mathfrak{M}(X)$ by

$$m_p(A\times\{n\}) := \mu_p(A)\left(\frac{1-p}{p}\right)^{-n}.$$

As shown in [HIK] (see also [A]), $T_p = (X,\mathcal{B},m_p,T)$ is a conservative, ergodic, measure preserving transformation (known as the *Hajian–Ito–Kakutani* transformation). The entropy is given by $\underline{h}(T_p) = h((T_p)_{\Omega\times\{0\}}) = 0$ by [MP] since $(T_p)_{\Omega\times\{0\}}$ is the Pascal adic transformation.

PROPOSITION 3.3. *$(X,\mathcal{B},m_p,T)$ is LLB for all $0<p<1$.*

Proof. As in the proof of Proposition 5.1 in [A1],

$$\sum_{k=0}^{2^n-1} 1_{\Omega\times\{0\}}\circ T^k(x,0) = \#\Big\{0\le k\le 2^n-1 : \sum_{j=0}^{k-1}\phi(\tau^j x) = 0\Big\}$$
$$\ge \#\Big\{0\le K\le n-1 : \sum_{j=0}^{2^K-1}\phi(\tau^j x) = 0\Big\}.$$

Now $\sum_{j=0}^{2^K-1}\phi(\tau^j x) = \phi(S^K x)$ where $S:\Omega\to\Omega$ is the shift, and so

$$\sum_{k=0}^{2^n-1} 1_{\Omega\times\{0\}}\circ T^k(x,0) \ge \#\{0\le K\le n-1 : \phi(S^K x) = 0\} \sim (1-p)n$$

for μ_p-a.e. $x\in\Omega$ by Birkhoff's theorem for the ergodic, probability preserving transformation $(\Omega,\mathcal{B}(\Omega),\mu_p,S)$. The LLB property now follows from Proposition 2.2. ∎

Let $\mathfrak{G}$ be the Polish group of measure preserving transformations of $(\mathbb{R},\mathcal{B}(\mathbb{R}),m_{\mathbb{R}})$ equipped with the weak topology.

PROPOSITION 3.4. *The collection of LLB measure preserving transformations is meagre in $\mathfrak{G}$.*

Proof. Let

$$\pounds := \Big\{T\in\mathfrak{G} : \exists n_k\to\infty,\ \frac{S_{n_k}(f)}{\log n_k}\to 0 \text{ a.e. } \ \forall f\in L^1\Big\}$$

where $S_n(f) = S_n^T(f) := \sum_{j=0}^{n-1} f\circ T^j$. By Proposition 2.2, it suffices to show that $\pounds$ is a dense G_δ set in $\mathfrak{G}$.

By Example 3.2, there exists a conservative, ergodic, measure preserving transformation $T \in \pounds$. The set $\pounds$ is conjugacy invariant, and so dense in $\mathfrak{G}$ by the conjugacy lemma (e.g. 3.5.2 in [A]).

To see that $\pounds$ is a G_δ set, let $P \sim m$ be a probability, fix $\{A_n : n \in \mathbb{N}\} \subset \mathcal{F} := \{A \in \mathcal{B} : m(A) < \infty\}$ so that $\sigma(\{A_n : n \in \mathbb{N}\}) = \mathcal{B}$ and let

$$\pounds' := \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \bigcap_{\nu=1}^{k} \Big\{T \in \mathfrak{G} : P\Big(\Big[S_n(1_{A_\nu}) > \frac{1}{k}\log n\Big]\Big) < \frac{1}{2^k}\Big\}.$$

Then $\pounds'$ is a G_δ. We claim $\pounds' = \pounds$.

Evidently,

$$\pounds' = \Big\{T \in \mathfrak{G} : \exists n_k \to \infty,\ \frac{S_{n_k}(1_{A_\nu})}{\log n_k} \to 0 \text{ a.e. } \forall \nu \geq 1\Big\},$$

whence $\pounds' \supset \pounds$.

Now suppose that $T \in \pounds'$ and $S_{n_k}(1_{A_\nu})/\log n_k \to 0$ a.e. for all $\nu \geq 1$, and let $f \in L^1$. Evidently $S_n(f)/\log n \to 0$ a.e. on $\mathfrak{D}$, the dissipative part of T. The conservative part of T is

$$\mathfrak{C} = \bigcup_{\nu=1}^{\infty} \hat{A}_\nu \quad \text{where} \quad \hat{A}_\nu := \Big[\sum_{n=1}^{\infty} 1_{A_\nu} \circ T^n = \infty\Big].$$

By Hopf's theorem, $S_n(f)(x)/S_n(1_{A_\nu})(x) \to h_\nu(f)$ a.e. on A_ν for all $\nu \geq 1$ where $h_\nu(f) \circ T = h_\nu(f)$ and $\int_{A_\nu} h_\nu(f)\,dm = \int_X f\,dm$, whence, a.e. on $\hat{A}_\nu$,

$$\frac{S_{n_k}(f)}{\log n_k} = \frac{S_{n_k}(f)}{S_{n_k}(1_{A_\nu})} \cdot \frac{S_{n_k}(1_{A_\nu})}{\log n_k} \to 0. \ \blacksquare$$

4. Information convergence. Let $(X, \mathcal{B}, m, T)$ be a conservative, ergodic, measure preserving transformation. A countable partition $\xi \subset \mathcal{B}$ is called *cofinite* if there exists $A = A_\xi \in \mathcal{F}$ with $A^c \in \xi$. We call A^c the *cofinite atom* of ξ and A the (*finite*) *core* of ξ.

If $\xi \subset \mathcal{B}$ is cofinite, then $\xi_k^l(T)$ is also cofinite, with core $A_{\xi_k^l(T)} = \bigcup_{j=k}^{l} T^{-j}A$.

The T-process generated by a cofinite partition ξ restricted to its core A is given by *Krengel's formula* [K1]:

$$(\mathfrak{K}) \qquad \xi_1^{\varphi_n(x)}(T)(x) \;=\; (\rho_A \vee ((\xi \cap A) \vee \rho_A)_1^n(T_A))(x) \qquad \text{for a.e. } x \in A$$

where for $x \in X$ and α a partition of X, $\alpha(x)$ is defined by $x \in \alpha(x) \in \alpha$, and

$$\varphi_n(x) := \sum_{k=0}^{n-1} \varphi_A(T_A^k x), \qquad \rho_A := \Big\{A \cap T^{-n}A \setminus \bigcup_{k=1}^{n-1} T^{-k}A : n \in \mathbb{N}\Big\}.$$

A cofinite partition $\xi \subset \mathcal{B}$ is called *quasifinite* (qf) if $A = A_\xi$ is quasifinite and $H_A(\xi) < \infty$. Note that ξ quasifinite $\Rightarrow H_A(\xi \vee \rho_A \vee T_A\,\rho_A) < \infty$.

Convergence of information for quasifinite partitions

PROPOSITION 4.1 (cf. [KS]). *Let $(X, \mathcal{B}, m, T)$ be a conservative, ergodic, measure preserving transformation, let $\xi \subset \mathcal{B}$ be a quasifinite partition and let $p \in L^1(m)$, $p > 0$, $\int_X p\,dm = 1$. Then for a.e. $x \in X$,*

$$\frac{1}{S_n(p)(x)}\, I(\xi_1^n(T))(x) \to \underline{h}(T, \xi)$$

where

$$S_n(p)(x) := \sum_{k=0}^{n-1} p(T^k x) \quad \text{and} \quad I(\xi_1^n(T))(x) := \log \frac{1}{m(\xi_1^n(T)(x))}.$$

Proof. Let A be the core of ξ and set $\varsigma := (\xi \cap A) \vee \rho_A$. Then by $(\mathfrak{K})$,

$$\varsigma_0^{s_n(x)}(T_A)(x) \subseteq \xi_1^n(T)(x) \subseteq \varsigma_1^{s_n(x)-1}(T_A)(x) \quad \text{a.e. } x \in A$$

where $x \in \xi(x) \in \xi$ and $s_n := S_n(1_A)$.

By $(\mathfrak{I})$, for T_A, a.e. on A, $I(\varsigma_1^N(T_A)) \sim N h(T_A, \varsigma)$, whence for a.e. $x \in A$,

$$\log \frac{1}{m(\xi_1^n(T)(x))} \sim \log \frac{1}{m(\varsigma_1^{s_n(x)}(T_A)(x))} \sim s_n(x) h(T_A, \varsigma)$$
$$\sim S_n(p)(x) m(A) h(T_A, \varsigma) = S_n(p)(x) \underline{h}(T, \xi).$$

We obtain convergence a.e. on $\bigcup_{k=0}^N T^{-k} A$ by substituting $\xi_1^N(T)$ for ξ; whence convergence a.e. on X as $\bigcup_{k=0}^N T^{-k} A \uparrow X$. ■

Abramov's formula is proved analogously in case $(X, \mathcal{B}, m, T)$ is an ergodic, probability preserving transformation. As in [Ab],

$$h(T, \xi) \overset{(\mathfrak{I})}{\longleftarrow} \frac{1}{n} \log \frac{1}{m(\xi_1^n(T)(x))} \approx \frac{1}{n} s_n(x) h(T_A, \varsigma) \xrightarrow{\text{Birkhoff's PET}} m(A) h(T_A, \varsigma).$$

5. Pinsker algebra. Let $(X, \mathcal{B}, m, T)$ be a LLB, conservative, ergodic, measure preserving transformation. Define

$$\mathcal{F}_\Pi := \{A \in \mathcal{F}_{\log,T} : A \in \sigma(\{T^{-k} A : k \ge 1\})\} = \mathcal{P} \cap \mathcal{F}_{\log,T}.$$

In this section, we show that (in case $\mathcal{F}_\Pi \neq \emptyset$) $\mathcal{B}_\Pi := \sigma(\mathcal{F}_\Pi)$ is the maximal zero entropy factor of T. To do this, we will need

KRENGEL'S PREDICTABILITY LEMMA ([K1]). *Let $(X, \mathcal{B}, m, T)$ be a quasifinite, conservative, ergodic, measure preserving transformation, let $\xi \subset \mathcal{B}$ be a quasifinite partition with core A, and let $\zeta = \xi \cap A$. Then*

$$\xi \subset \xi_1^\infty(T) \bmod m \;\Leftrightarrow\; h(T_A, \zeta \vee \rho_A) = 0.$$

In particular,

$$A \in \sigma(\{T^{-n}A : n \geq 1\}) \;\Leftrightarrow\; h(T_A, \rho_A) = 0.$$

For $F \in \mathcal{F}$, set

$$\mathcal{P}_F = \mathcal{P}_{T_F} := \{A \in \mathcal{B} \cap F : A \in \sigma(\{T_F^{-k}A : k \geq 1\})\}.$$

By Pinsker's theorem ([Pi]),

- $\mathcal{P}_F$ is a T_F-factor algebra of subsets of F with $h(T_F, \mathcal{P}_F) = 0$,
- if $\mathcal{A} \subset \mathcal{B} \cap F$ is another T_F-factor algebra of subsets of F with $h(T_F, \mathcal{A}) = 0$, then $\mathcal{A} \subset \mathcal{P}_F$.

Theorem 5.1.

(i) *$\mathcal{F}_\Pi$ is a ring and $\mathcal{F}_\Pi \cap F = \mathcal{P}_F$ for all $F \in \mathcal{F}_\Pi$.*
(ii) *If $\mathcal{F}_\Pi \neq \emptyset$, then $\sigma(\mathcal{F}_\Pi)$ is the maximal factor of zero entropy.*

Proof. ¶1 Let $A \in \mathcal{F}_{\log}$. By Krengel's predictability lemma, $F \in \mathcal{F}_\Pi$ iff $h(T_F, \rho_F) = 0$. Thus, $F \in \mathcal{F}_\Pi$ iff there is a factor $\mathcal{B}_0$ with $F \in \mathcal{B}_0$ and $\underline{h}(T, \mathcal{B}_0) = 0$.

¶2 Next, fix $F \in \mathcal{F}_\Pi$. We claim that $\rho_F \subseteq \mathcal{P}_F$. This is because $F \in \mathcal{F}_\Pi \;\Rightarrow\; h(T_F, \rho_F) = 0$.

¶3 We now show that $\mathcal{P}_F \subseteq \mathcal{F}_\Pi \cap F$ for all $F \in \mathcal{F}_\Pi$. Fix $F \in \mathcal{F}_\Pi$ and let $\mathcal{B}_0 := \sigma\{T^n A : n \in \mathbb{Z},\ A \in \mathcal{P}_F\}$. Then $\mathcal{B}_0$ is a factor, $F \in \mathcal{B}_0$ and $\mathcal{B}_0 \cap F = \mathcal{P}_F$. Thus $\underline{h}(T, \mathcal{B}_0) = h(T_F, \mathcal{P}_F) = 0$ and by ¶1, $\mathcal{P}_F \subseteq \mathcal{F}_\Pi \cap F$.

¶4 Now we claim that $A, B \in \mathcal{F}_\Pi \Rightarrow A \cup B \in \mathcal{P}_F$. Set $C := A \cup B$. Then $C \in \mathcal{F}_{\log,T}$. Set $\zeta := \{A \cap B, A \setminus B, B \setminus A\}$ and $\xi := \zeta \cup \{C^c\}$. By ($\mathfrak{K}$),

$$\xi_1^\infty(T) \cap C = \rho_C \vee (\zeta \vee \rho_C)_1^\infty(T_C).$$

By assumption, $\zeta \subset \xi_1^\infty(T) \cap C$, whence also $\rho_C \subset \xi_1^\infty(T) \cap C$. Thus

$$\zeta \vee \rho_C \subset \rho_C \vee (\zeta \vee \rho_C)_1^\infty(T_C), \quad \text{so} \quad \zeta \vee \rho_C \vee T_C\, \rho_C \subset (\zeta \vee \rho_C \vee T \rho_C)_1^\infty(T_C),$$

and (using $H_C(\zeta \vee \rho_C \vee T_C\, \rho_C) < \infty$) we have

$$h(T_C, \rho_C) \leq h(T_C, \zeta \vee \rho_C \vee T_C\, \rho_C) = 0,$$

whence $C \in \sigma(\{T^{-k}C : k \geq 1\})$ and $C \in \mathcal{F}_\Pi$.

¶5 Now we show that $\mathcal{F}_\Pi$ is a ring by proving that $A, B \in \mathcal{F}_\Pi \;\Rightarrow\; \zeta := \{A \cap B, A \setminus B, B \setminus A\} \subset \mathcal{F}_\Pi$. By ¶3, it suffices to show that $\zeta \subset \mathcal{P}_C$ where $C := A \cup B$. To see this, fix $a \in \zeta$. Then

$$h(T_C, \{a, C \setminus a\}) \leq h(T_C, \zeta) \leq h(T_C, \zeta \vee \rho_C \vee T_C\, \rho_C) = 0$$

(as above) and $a \in \mathcal{P}_C$.

¶6 To complete the proof of (i), we show that $\mathcal{F}_\Pi \cap F \subseteq \mathcal{P}_F$ for all $F \in \mathcal{F}_\Pi$. Fix $F \in \mathcal{F}_\Pi$, $A \in \mathcal{F}_\Pi \cap F$. Let $\zeta := \{A, F \setminus A\}$, $\xi := \zeta \cup \{F^c\}$. By the ring property, $A \in \mathcal{F}_\Pi$, whence $\xi \subset \xi_1^\infty(T) \bmod m$. By Proposition 4.1, $h(T_F, \zeta \vee \rho_F) = 0$, whence

$$h(T_F, \zeta) \leq h(T_F, \zeta \vee \rho_F) = 0$$

and $A \in \mathcal{P}_F$.

¶7 To see (ii), fix $F \in \mathcal{F}_\Pi$. Then by (i), $\mathcal{F}_\Pi \cap F = \mathcal{P}_F = \mathcal{F}_\Pi \cap F \cap F$, whence $\underline{h}(T, \sigma(\mathcal{F}_\Pi)) = m(F)h(T_F, \mathcal{P}_F) = 0$ and if $\mathcal{C} \subset \mathcal{B}$ is a factor with $\underline{h}(T, \mathcal{C}) = 0$, then by ¶1, $\mathcal{C} \cap \mathcal{F}_{\log} \subset \mathcal{F}_\Pi$, whence $\mathcal{C} \subset \sigma(\mathcal{F}_\Pi)$. ■

6. Asymptotic distribution of information with infinite invariant measure

Pointwise dual ergodic transformations. Let $(X, \mathcal{B}, m, T)$ be a pointwise dual ergodic, measure preserving transformation and assume that the return sequence $a_n = a_n(T)$ is regularly varying with index α $(\alpha \in [0,1])$. Then by the Darling–Kac theorem (Theorem 3.6.4 in [A]—see also references therein),

$$(\text{☿}) \qquad \frac{1}{a_n} S_n^T(f) \xrightarrow{\mathfrak{d}} \int_X f\, dm \cdot X_\alpha \quad \text{as } n \to \infty\ \forall f \in L^1(m)_+$$

where X_α is a Mittag-Leffler random variable of order α normalised so that $E(X_\alpha) = 1$, and $F_n \xrightarrow{\mathfrak{d}} Y$ means

$$\int_X G(F_n)\, dP \to E(G(Y)) \quad \forall P \in \mathcal{P}(X, \mathcal{B}),\ P \ll m,\ G \in C([0,\infty]).$$

Note that $X_1 \equiv 1$, X_0 has exponential distribution, and for $\alpha \in (0,1)$, $X_\alpha = 1/Y_\alpha^\alpha$ where $E(e^{-tY_\alpha}) = e^{-ct^\alpha}$ (for some $c = c_\alpha > 0$). In particular, $X_{1/2} = |\mathcal{N}|$ where $\mathcal{N}$ is a centred Gaussian random variable on $\mathbb{R}$.

Proposition 6.1. *Suppose that $(X, \mathcal{B}, m, T)$ is a quasifinite, pointwise dual ergodic, measure preserving transformation, and assume that the return sequence $a_n = a_n(T)$ is regularly varying with index $\alpha \in [0,1]$. If $\xi \subset \mathcal{B}$ is quasifinite, then*

$$\frac{1}{a_n(T)} \log \frac{1}{m(\xi_1^n(T)(x))} \xrightarrow{\mathfrak{d}} \underline{h}(T, \xi) X_\alpha \quad \textit{as } n \to \infty.$$

Proof. This follows from Proposition 4.1 and (☿). ■

Example 6.2 (Boole's transformation). Let $(X, \mathcal{B}, m, T)$ be given by $X = \mathbb{R}$, $m =$ Lebesgue measure and $Tx = x - 1/x$. Then T (see [A]) is a pointwise dual ergodic, measure preserving transformation with $a_n(T) \sim \sqrt{2n}/\pi$, so $\mathcal{F}_\Pi \neq \emptyset$ and T is LLB, whence quasifinite.

By Proposition 6.1, if $\xi \subset \mathcal{B}$ is quasifinite, then

$$(\text{♋}) \qquad \frac{1}{a_n(T)} \log \frac{1}{m(\xi_1^n(T)(x))} \xrightarrow{\mathfrak{d}} \underline{h}(T, \xi)|\mathcal{N}| \quad \text{as } n \to \infty.$$

7. Analogous properties of probability preserving transformations. The last section is devoted to the construction of an ergodic, probability preserving transformation having a generating partition with properties analogous to (♋). The related "measure-theoretic invariant" is *entropy dimension* as in [FP].

Let $(\mathbb{T}, \mathcal{T}, m_{\mathbb{T}}, R)$ be an irrational rotation of the circle (equipped with Borel sets and Lebesgue measure). Let $f \in L^2(\mathbb{T})$ satisfy the weak invariance principle, i.e. $B_n(t) \to B(t)$ in distribution on $C([0,1])$ where B is Brownian motion and

$$B_n(t) := f_{[nt]-1} + (nt - [nt]) f \circ T^{[nt]}$$

(where $f_k := \sum_{j=0}^{k-1} f \circ R^j$). Existence of such $f \in L^2(\mathbb{T})$ is shown in [V].

In particular,

$$\frac{L_n}{\sqrt{n}}, \frac{R_n}{\sqrt{n}} \xrightarrow{\mathfrak{d}} |\mathcal{N}|, \qquad \frac{L_n + R_n}{\sqrt{n}} \xrightarrow{\mathfrak{d}} \mathcal{R}$$

where $R_n := \max_{1 \le k \le n} f_k$, $L_n := \max_{1 \le k \le n} (-f_k)$ and $\mathcal{R} := \max_{t \in [0,1]} B(t) - \min_{t \in [0,1]} B(t)$.

The random variable $\mathcal{R}$ is known as the *range of Brownian motion.* Its (non-Gaussian) distribution is calculated in [Fe].

Let $(Y, \mathcal{C}, \mu, S)$ be the 2-shift with generating partition $Q = \{Q_0, Q_1\}$ and symmetric product measure. Let $\rho : Y \to \mathbb{R}$ be defined by $\rho = \alpha_0 1_{Q_0} + \alpha_1 1_{Q_1}$ where $\alpha_0 < \alpha_1$, $\int_Y \rho \, d\mu = 1$ and α_0, α_1 are rationally independent. Then the *special flow* (under ρ) $(Y^\rho, \mathcal{C}^\rho, q, S^\rho)$ is Bernoulli where

$$Y^\rho := \{(y,s) : y \in Y,\ s \in [0, \rho(y))\}, \qquad \mathcal{C}^\rho := \mathcal{C} \times \text{Lebesgue}, \qquad q := \mu \times \lambda,$$

and

$$S_t^\rho(y,s) := (S^n y, s + t - \rho_n(y))$$

where $0 \le s + t - \rho_n(y) < \rho(S^n y)$, $\rho_n := \sum_{j=0}^{n-1} \rho \circ S^j$.

Note that the "vertical" partition $\overline{Q} := \{\overline{Q}_0, \overline{Q}_1\}$ where $\overline{Q}_i := Q_i \times [0, \alpha_i)$ $(i = 0, 1)$ generates $\mathcal{C}$ under S^ρ.

Define the probability preserving transformation $(X, \mathcal{B}, m, T)$ by

(☎)
$$\begin{aligned} &X := \mathbb{T} \times Y^\rho, \qquad m = m_{\mathbb{T}} \times q, \qquad \mathcal{B} := \mathcal{T} \times \mathcal{C}^\rho, \\ &T(x, (y,s)) := (R(x), S^\rho_{f(x)}(y,s)). \end{aligned}$$

For P a finite partition of $\mathbb{T}$ into intervals (which generates $\mathcal{T}$ under R), define the partition $\xi = \xi_P$ of X by

(⛢)
$$\xi(\omega, y, s) := P(\omega) \times \Big(\bigvee_{t \in \iota(0, f(\omega))} S^\rho_{-t} \overline{Q} \Big)(y,s)$$

where for $x, y \in \mathbb{R}, \iota(x,y) := [x \wedge y, x \vee y]$ (the closed interval joining x and y). Next, we show that ξ is measurable and $H(\xi) < \infty$.

Proposition 7.1. *The partition ξ is measurable, generates $\mathcal{B}$ under T, $H(\xi) < \infty$ and*

$$(☿) \qquad \frac{1}{\sqrt{n}} I(\xi_0^{n-1}(T)) \xrightarrow{\mathfrak{d}} h(S^\rho)\mathcal{R} \quad \text{as } n \to \infty,$$

where $\mathcal{R}$ is the range of Brownian motion.

Proof. The proof is in several stages. We first claim that

$$(♪) \quad \xi_0^{n-1}(T)(\omega, y, s) = P_0^{n-1}(R)(\omega) \times \Big(\bigvee_{t \in [-L_n(\omega), R_n(\omega)]} S^\rho_{-t} t\overline{Q}\Big)(y, s).$$

To see this, note that for $n \geq 1$,

$$\begin{aligned}
(T^{-n}\xi)(\omega, y, s) &= \xi(R^n(\omega), S^\rho_{f_n(\omega)}(y, s)) \\
&= P(R^n(\omega)) \times \Big(\bigvee_{t \in \iota(0, f(R^n(\omega)))} S^\rho_{-t} t\overline{Q}\Big)(S^\rho_{f_n(\omega)}(y, s)) \\
&= P(R^n(\omega)) \times \Big(\bigvee_{t \in \iota(f_n(\omega), f_n(\omega) + f(R^n(\omega)))} S^\rho_{-t} t\overline{Q}\Big)(y, s) \\
&= P(R^n(\omega)) \times \Big(\bigvee_{t \in \iota(f_n(\omega), f_{n+1}(\omega))} S^\rho_{-t} t\overline{Q}\Big)(y, s).
\end{aligned}$$

To continue, we need the following (elementary) proposition:

¶ Let $a_n \in \mathbb{R}$ $(n \geq 1)$. Then

$$\bigcup_{k=0}^{n-1} \iota(s_k, s_{k+1}) = [m_n, M_n]$$

where $a_0 := 0$, $s_n := \sum_{k=0}^n a_k$, $m_n := \min_{0 \leq k \leq n} s_k$, $M_n := \max_{0 \leq k \leq n} s_k$.

To finish the proof of (♪), note that

$$\begin{aligned}
\xi_0^{n-1}(T)(\omega, y, s) &= \bigvee_{k=0}^{n-1} T^{-k}\xi(\omega, y, s) \\
&= \bigcap_{k=0}^{n-1} P(R^k(\omega)) \times \Big(\bigvee_{t \in \iota(f_k(\omega), f_{k+1}(\omega))} S^\rho_{-t} t\overline{Q}\Big)(y, s) \\
&= P_0^{n-1}(R)(\omega)) \times \Big(\bigvee_{t \in \bigcup_{k=0}^{n-1} \iota(f_k(\omega), f_{k+1}(\omega))} S^\rho_{-t} t\overline{Q}\Big)(y, s) \\
&\overset{¶}{=} P_0^{n-1}(R)(\omega) \times \Big(\bigvee_{t \in [-L_n(\omega), R_n(\omega)]} S^\rho_{-t} \overline{Q}\Big)(y, s),
\end{aligned}$$

proving (♪).

Now consider $\rho_n : Y \to \mathbb{R}$ defined by

$$\rho_n(y) := \begin{cases} \sum_{k=0}^{n-1} \rho(S^k y), & n > 0, \\ 0, & n > 0, \\ \sum_{k=1}^{|n|} \rho(S^{-k} y), & n < 0. \end{cases}$$

Then $\rho_n(y) < \rho_{n+1}(y)$ and for all $y \in Y$, $\rho_n(y) \to \pm\infty$ as $n \to \pm\infty$.

For $y \in Y$ and $t \in \mathbb{R}$, define $[t]_y \in \mathbb{Z}$ to be so that $\rho_{[t]_y}(y) \le t < \rho_{[t]_y+1}(y)$. It follows that for $t \in \mathbb{R}$,

$$|t|/\alpha_1 - 1 \le |[t]_y| \le |t|/\alpha_0, \qquad S_t^\rho(y,s) = (S^{[s+t]_y} y, s + t - \rho_{[s+t]_y}(y)).$$

Our next claim is that

$$(\text{♫}) \qquad \xi_0^{n-1}(T)(\omega, y, s) = P_0^{n-1}(R)(\omega) \times Q_{[s-L_n(\omega)]_y}^{[s+R_n(\omega)]_y}(S)(y) \times \eta_n(\omega, y)(s)$$

where for each $(\omega, y) \in \Omega \times Y$, $\eta_n(\omega, y)$ is a partition of $[0, \rho(y))$ into at most $(R_n(\omega) + L_n(\omega) + 1)/\alpha_0$ intervals. Indeed, fixing $(\omega, y, s) \in X$ and $n \ge 1$, we have

$$\begin{aligned} \Big(\bigvee_{t \in [-L_n(\omega), R_n(\omega)]} S_{-t}^\rho \overline{Q}\Big)(y,s) &= \bigcap_{t \in [-L_n(\omega), R_n(\omega)]} \overline{Q}(S_t^\rho(y,s)) \\ &= \bigcap_{t \in [-L_n(\omega), R_n(\omega)]} Q(S^{[s+t]_y} y) \times [0, \rho(S^{[s+t]_y} y)) \\ &= \bigcap_{j \in [[s-L_n(\omega)]_y, [s+R_n(\omega)]_y]} S^{-j} Q(y) \times \eta_n(\omega, y, s) \\ &= Q_{[s-L_n(\omega)]_y}^{[s+R_n(\omega)]_y}(S)(y) \times \eta_n(\omega, y)(s) \end{aligned}$$

where for each $(\omega, y) \in \Omega \times Y$, $\eta_n(\omega, y)$ is a partition of $[0, \rho(y))$ into at most $[R_n(\omega)]_y - [-L_n(\omega)]_y \le (R_n(\omega) + L_n(\omega) + 1)/\alpha_0$ intervals. This proves (♫).

Now (♫) with $n = 1$ shows that

$$\xi(\omega, y, s) = P(\omega) \times Q_{-\nu_-(\omega,y,s)}^{\nu_+(\omega,y,s)}(S)(y) \times \eta_1(\omega, y)(s)$$

where

$$\nu_+(w, y, s) = [s + f(\omega) \vee 0]_y, \qquad \nu_-(w, y, s) = [s + f(\omega) \wedge 0]_y.$$

Thus, ξ is measurable. Moreover, writing $\mathcal{Z} := \{[\nu_- = k,\ \nu_+ = l] : k, l \in \mathbb{Z}\}$, we see that

$$\begin{aligned} I(\xi \,|\, \mathcal{Z})(\omega, y, s) &= I(P)(\omega) + I(Q_{[s+f(\omega)\wedge 0]_y}^{[s+f(\omega)\vee 0]_y})(S)(y) + I(\eta_1(\omega, y)(s)) \\ &\le I(P)(\omega) + ([s + f(\omega) \wedge 0]_y + [s + f(\omega) \vee 0]_y) \cdot \log 2 + \log \frac{1 + |f(\omega)|}{\alpha_0} \\ &\le I(P)(\omega) + \frac{|f(\omega)| + 1}{\alpha_0} \cdot \log 2 + \log \frac{1 + |f(\omega)|}{\alpha_0} \end{aligned}$$

and

$$H(\xi\,|\,\mathcal{Z}) \le H(P) + \frac{\log 2}{\alpha_0}(\|f\|_1 + 1) + \int_\Omega \log\frac{1+|f|}{\alpha_0}\,dm < \infty.$$

Now $|\nu_\pm(\omega, y, s)| \le (|f(\omega)| + 1)/\alpha_0$ and

$$(\nu_+(w, y, s), \nu_-(w, y, s)) = \begin{cases} ([s + f(\omega) \vee 0]_y, 0), & f(\omega) \ge 0, \\ (0, [s + f(\omega) \wedge 0]_y), & f(\omega) < 0; \end{cases}$$

whence by (★) (see page 7), $H(\mathcal{Z}) < \infty$ and

$$H(\xi) = H(\xi\,|\,\mathcal{Z}) + H(\mathcal{Z}) < \infty.$$

Since ξ is measurable, (♪) now shows that it generates $\mathcal{B}$ under T.

To establish (♅), we claim that for a.e. (x, y, s), any $\epsilon > 0$, and sufficiently large $n = n(x, y, s)$,

$$(♣)\qquad P_0^{n-1}(R)(x) \times Q_{-L_n(x)(1+\epsilon)}^{R_n(x)(1+\epsilon)}(S)(y) \times \eta_n(x, y)(s) \subseteq \xi_0^{n-1}(T)(x, y, s) \subseteq P_0^{n-1}(R)(x) \times Q_{-L_n(x)(1-\epsilon)}^{R_n(x)(1-\epsilon)}(S)(y) \times \eta_n(x, y)(s)$$

where for each $(\omega, y) \in \Omega \times Y$, $\eta_n(\omega, y)$ is a partition of $[0, \rho(y))$ into at most $(R_n(\omega) + L_n(\omega) + 1)/\alpha_0$ intervals.

To see this note that for a.e. $(x, y, s) \in X$, $R_n(x), L_n(x) \uparrow \infty$ and $\rho_n(y) \sim n$, whence $|[s - L_n(x)]_y| \sim L_n(\omega)$ and $[s + R_n(\omega)]_y \sim R_n(x)$. Now (♣) follows from (♫) using this.

We next claim that for all $(x, y) \in \mathbb{T} \times Y$,

$$(☼)\qquad \frac{1}{\sqrt{n}}\left(I(P_0^{n-1}(R)) + I(\eta_n(x, y))\right) \xrightarrow{m} 0.$$

Indeed, $\#\eta_n(x, y) \le \mathcal{E}_n(x) := (R_n(x) + L_n(x) + 1)/\alpha_0$ and $\#P_0^{n-1}(R) \le Mn$ for some $M > 0$ and all $n \ge 1$, whence

$$m([I(P_0^{n-1}(R)) \ge t\sqrt{n}]) \le \frac{1}{t\sqrt{n}} H(P_0^{n-1}(R)) \lesssim \frac{\log n}{t\sqrt{n}} \to 0 \quad \text{as } n \to \infty$$

and for all (x, y),

$$m([I(\eta_n(x, y)(s))) \ge t\sqrt{n}]) \le \frac{1}{t\sqrt{n}} H(\eta_n(x, y)) \le \frac{\log \mathcal{E}_n(x)}{t\sqrt{n}} \xrightarrow{m} 0 \quad \text{as } n \to \infty,$$

proving (☼).

Using (♣), (☼) and ($\mathfrak{I}$) for S we have, as $n \to \infty$,

$$\frac{1}{\sqrt{n}} I(\xi_0^{n-1}(T))(x,y,s) = \frac{1}{\sqrt{n}} I(Q_{-L_n(x)(1+o(1))}^{R_n(x)(1+o(1))}(S))(y) + O\Big(\frac{\log n}{\sqrt{n}}\Big)$$
$$= \frac{1}{\sqrt{n}} (L_n(x) + R_n(x)) \log 2(1+o(1)) + O\Big(\frac{\log n}{\sqrt{n}}\Big)$$
$$\xrightarrow{\mathfrak{d}} \mathcal{R}\log 2 = \mathcal{R}h(S^\rho). \ ■$$

Estimation of entropy dimension. Let $(Z,\mathcal{D},\nu,R)$ be a probability preserving transformation and let $P \subset \mathcal{D}$ be a countable partition of Z. As in [FP], for $n \ge 1$, $\epsilon > 0$ and $a = \bigcap_{k=0}^{n-1} R^{-k} a_k \in P_0^{n-1}(R)$, let

$$B(n,P,a,\epsilon) := \bigcup_{a' \in P_0^{n-1}(R),\, \bar{d}(a,a')<\epsilon} a$$

where $\bar{d}(a,a') := n^{-1}\#\{0 \le k \le n-1 : a_k \ne a'_k\}$ is the Hamming distance, and let

$$K(P,n,\epsilon) := \min\Big\{\#F : F \subset P_0^{n-1}(R),\ \nu\Big(\bigcup_{a\in F} B(n,P,a,\epsilon)\Big) > 1-\epsilon\Big\}.$$

The ergodic, probability preserving transformation is said to have *upper entropy dimension* $\varDelta \in [0,1]$ if for some countable, measurable generating partition P with finite entropy (and hence—as proved in [FP]—for all such),

$$\overline{\lim_{n\to\infty}} \frac{\log\log K(P,n,\epsilon)}{\log n} \underset{\epsilon\to 0}{\longrightarrow} \varDelta.$$

PROPOSITION 7.2. *Let $(X,\mathcal{B},m,T)$ be as in* (☎). *Then the upper entropy dimension is at most* $1/2$.

Proof. Let $\xi = \xi_P$ be as in (♁) and let $h = h(S^\rho)$. For $n \ge 1$ and $J \subset \mathbb{R}_+$ an interval bounded away from 0 and ∞, define

$$\xi_n(J) := \Big\{a \in \xi_0^{n-1}(T) : \frac{1}{\sqrt{n}} \log \frac{1}{m(a)} \in hJ\Big\}.$$

We claim that

$$\#\xi_n(J) \sim E(1_J(\mathcal{R})e^{h\mathcal{R}\sqrt{n}})e^{o(\sqrt{n})} \qquad \text{as } n \to \infty.$$

To see this, suppose that $J = [r-\delta, r+\delta]$. Then

$$P(\mathcal{R} \in J) \leftarrow m\Big(\Big[\frac{1}{\sqrt{n}} I(\xi_0^{n-1}(T)) \in hJ\Big]\Big)$$
$$= \sum_{a \in \xi_n(J)} m(a) = \#\xi_n(J) e^{-h\sqrt{n}(r\pm\delta)}$$

(because $m(a) = e^{-h\sqrt{n}(r\pm\delta)}$ for all $a \in \xi_n(J)$); whence

$$E(e^{h\sqrt{n}(\mathcal{R}-2\delta)}1_J(\mathcal{R})) \lesssim \#\xi_n(J) \lesssim E(e^{h\sqrt{n}(\mathcal{R}+2\delta)}1_J(\mathcal{R})).$$

Using this on a decomposition of J into a finite union of disjoint short enough intervals yields $\#\xi_n(J) = E(e^{h\sqrt{n}\mathcal{R}}1_J(\mathcal{R}))e^{\pm\epsilon\sqrt{n}}$ for all $\epsilon > 0$, proving the claim.

Evidently $K(\xi, n, \epsilon) \le \#\xi_n([1/M, M])$ for some $M = M_\epsilon > 0$, whence $K(\xi, n, \epsilon) \le e^{c_\epsilon\sqrt{n}(1+o(1))}$ and

$$\overline{\lim_{n\to\infty}} \frac{\log\log K(\xi, n, \epsilon)}{\log n} \le \frac{1}{2} \quad \forall \epsilon > 0. \ \blacksquare$$

Remark on the lower bound. The upper estimate for the entropy dimension follows from the weak invariance principle for the "random walk" f_n. In a similar manner, a lower estimate would follow from an analogous result for the "local time" of the random walk. Such a result is not available for the present example. However, such considerations show that the "relative entropy dimension" over its Bernoulli factor of an aperiodic, centered random walk in random scenery with jumps of finite variance is 1/2.

References

[A] J. Aaronson, *An Introduction to Infinite Ergodic Theory*, Math. Surveys Monogr. 50, Amer. Math. Soc., Providence, RI, 1997.

[A1] —, *The intrinsic normalising constants of transformations preserving infinite measures*, J. Anal. Math. 49 (1987), 239–270.

[AL] J. Aaronson and M. Lemańczyk, *Exactness of Rokhlin endomorphisms and weak mixing of Poisson boundaries*, in: Contemp. Math. 385, Amer. Math. Soc., 2005, 77–87.

[Ab] L. M. Abramov, *Entropy of a derived automorphism*, Dokl. Akad. Nauk SSSR 128 (1959), 647–650 (in Russian); English transl.: Amer. Math. Soc. Transl. Ser. II 49 (1960), 162–176.

[AR] L. M. Abramov and V. A. Rohlin, *Entropy of a skew product of mappings with invariant measure*, Vestnik Leningrad. Univ. 17 (1962), no. 7, 5–13 (in Russian).

[Br] L. Breiman, *The individual ergodic theorem of information theory*, Ann. Math. Statist. 28 (1957), 809–811.

[C] K. L. Chung, *A note on the ergodic theorem of information theory*, ibid. 32 (1961), 612–614.

[Fe] W. Feller, *The asymptotic distribution of the range of sums of independent random variables*, Ann. Math. Statist. 22 (1951), 427–432.

[FP] S. Ferenczi and K. K. Park, *Entropy dimensions and a class of constructive examples*, Discrete Contin. Dynam. Systems 17 (2007), 133–141.

[Fr] N. A. Friedman, *Introduction to Ergodic Theory*, Van Nostrand Reinhold, New York, 1970.

[HIK] A. Hajian, Y. Ito and S. Kakutani, *Invariant measures and orbits of dissipative transformations*, Adv. Math. 9 (1972), 52–65.

[IT] A. Ionescu Tulcea, *Contributions to information theory for abstract alphabets*, Ark. Mat. 4 (1961), 235–247.

[KS] E. M. Klimko and L. Sucheston, *On convergence of information in spaces with infinite invariant measure*, Z. Wahrsch. Verw. Gebiete 10 (1968), 226–235.

[K1] U. Krengel, *Entropy of conservative transformations*, ibid. 7 (1967), 161–181.

[K2] —, *On certain analogous difficulties in the investigation of flows in a probability space and of transformations in an infinite measure space*, in: Functional Analysis (Monterey, CA, 1969), C. O'Wilde (ed.), Academic Press, New York, 1970, 75–91.

[LL] M. Lemańczyk and E. Lesigne, *Ergodicity of Rokhlin cocycles*, J. Anal. Math. 85 (2001), 43–86.

[M] B. McMillan, *The basic theorems of information theory*, Ann. Math. Statist. 24 (1953), 196–219.

[MP] X. Mela and K. Petersen, *Dynamical properties of the Pascal adic transformation*, Ergodic Theory Dynam. Systems 25 (2005), 227–256.

[Pa] W. Parry, *Entropy and Generators in Ergodic Theory*, Benjamin, New York, 1969.

[Pi] M. S. Pinsker, *Dynamical systems with completely positive or zero entropy*, Dokl. Akad. Nauk SSSR 133 (1960), 1025–1026 (in Russian); English transl.: Soviet Math. Dokl. 1 (1960), 937–938.

[S] C. E. Shannon, *A mathematical theory of communication*, Bell System Tech. J. 27 (1948), 379–423, 623–656.

[V] D. Volný, *Invariance principles and Gaussian approximation for strictly stationary processes*, Trans. Amer. Math. Soc. 351 (1999), 3351–3371.

Jon Aaronson
School of Math. Sciences
Tel Aviv University
69978 Tel Aviv, Israel
E-mail: aaro@tau.ac.il

Kyewon Koh Park
Department of Mathematics
Ajou University
Suwon 442-729, South Korea
E-mail: kkpark@ajou.ac.kr

Received 16 January 2009