Spatial Recurrence for Ergodic Fractal Measures

Nadav Dym
Advisor: Dr. Michael Hochman

August 17, 2014

Abstract

We discuss an invertible version of Furstenberg's 'Ergodic CP Shift Systems'. We show that the regularity of these dynamical systems with respect to magnification of measures, implies certain regularity with respect to translation of measures.

1 Introduction

Ergodic CP Shift Systems (ECPS) were introduced by Furstenberg in [4] as a tool in his proof of dimension conservation for homogenous fractals, and were also used by him in a slightly different form in earlier work [2, 3], and by Hochman and Shmerkin in [5]. These are 'dynamical systems whose states represent measures on $\mathbb{R}$, in which progression in time corresponds to progressively increasing magnification of the measures[1]'. A general discussion of ECPS and other 'fractal dynamical systems' can be found in [6].

We discuss the implications or the regularity of an invertible version of ECPS, which we call Extended ECPS, with respect to 'zooming in' into measures, on the behavior of typical measures when translated and normalized. We show that this translation and normalization action is conservative whenever the Extended ECPS is non-deterministic (to be defined), which is equivalent to saying that typical measures are not Dirac measures. This result, and its proof, resembles Host's proof [7] that a 'non deterministic' measure on $[0, 1)$ invariant and ergodic under the $\times p$ map ('zooming in') is conservative with respect to the action of translation by numbers whose base p representation is finite.

We use this conservativity to show that 'ergodic averages' of the form

$$\frac{1}{\nu[0, N)} \sum_{n=0}^{N-1} f(t_n^* \nu) \nu[n, n+1)$$

[1] This is a quote from the abstract of [4], with slight changes

and

$$\frac{1}{\nu[0,T)}\int_0^T f(t_x^*\nu)d\nu(x)$$

(where $t_x^*\nu$ is the measure ν translated by x and then normalized) converge for typical ν.

This result can be compared to [9], where Medynets and Solomyak prove a second-order ergodic theorem for the action of translation by $\mathbb{R}^d$ on a different class of 'fractal dynamical systems' (self-similar tiling systems).

Finally, we discuss a 'pointwise analogue' of the question of conservativity described above, and show that if an Extended ECPS in not bilaterally deterministic (to be defined), then the translation action on $\mathbb{R}$ is conservative with respect to typical measures (which are points in the Extended ECPS).

We now begin Subsection 1.1 in which we present some terminology we use throughout the paper. We then define Extended ECPS in Subsection 1.2, which will enable us to give a complete presentation of our main results in Subsection 1.3. At the end of our discussion of the main results we will give the outline of the rest of the paper.

Acknowledgments I would like to thank Mike for providing me with interesting questions to work on, for many interesting discussions, and in general, for his substantial help in making the work on this thesis a pleasant experience.

I am grateful to my wife Efrat for her companionship in general, and specifically for her understanding and assistance in my (at most partially successful) attempts to find the correct balance between managing a household and getting some work done.

List of Notation

$\mathbb{M}(\mathbb{R})$	Radon positive measures on $\mathbb{R}$.
N	Map normalizing elements of $\mathbb{M}(\mathbb{R})$.
$N\mathbb{M}(\mathbb{R})$	Space of normalized measures.
$\mathbb{P}_0[0,1)$	Space of probability measures on $[0,1)$, and zero.
LS	Space of legal sequences $(\mu_n, i_n)_{n\in\mathbb{Z}}$.
$X, \tilde{X}$	Spaces with points of the form $(\mu_n, i_n)_{n\leq 0}$, $(\nu, (i_n)_{n\leq 0})$.
M_p^{ext}, M_p^{CP}	zooming in maps on Extended ECPS and ECPS.
S_a, T_k	translation maps on X and $\tilde{X}$.
Φ, θ	Isomorphisms $\Phi : LS \to N\mathbb{M}(\mathbb{R}) \times \{0,1,\dots,p-1\}^{\mathbb{Z}}$ and $\theta : X \to \tilde{X}$.
$\hat{\nu}$,$\tilde{\pi}_{\mathbb{M}}$, $\hat{\mu}_n$	Projections to the measure coordinate.
$\hat{i}_n$	Projection to the symbolic coordinate.
$\hat{x}_n$	The n-th digit in the standard representation of x.
t_x,$t_x^{\mathbb{T}}$	Translation by x on $\mathbb{R}$, $\mathbb{T}$.
t_x^*	Translation and normalization map on $N\mathbb{M}(\mathbb{R})$.
$\boldsymbol{\sigma}$,$\boldsymbol{\sigma}_-$	Left shift, right shift.
$\overrightarrow{j}$	Sequence equal to j in all coordinates.
E_j, E_j^-	Sequences $(i_n)_{n\in\mathbb{Z}}, (i_n)_{n\leq 0}$ with $i_n = j$ for all negative enough n.
E_j^+	Sequences $(i_n)_{n\in\mathbb{N}}$ with $i_n = j$ for all large enough n.
$\mathbb{D}_{p^n}$	Partition of $[0,1)$ into p^n equal intervals.
D_{p^n}, D	Points of the form kp^{-n}, and the union of D_{p^n}.
G_n, G	Points of the form $(\dots,0,0,i_n,\dots,i_{-1},i_0)$, and the union of G_n.
$[i_1,\dots,i_n)_{p^n}$	the interval $[\sum_{j=1}^n i_j p^{-j}, \sum_{j=1}^n i_j p^{-j} + p^{-n})$.

1.1 Terminology

1. If $(x_n)_{n\in\mathbb{Z}}$ is a sequence, we use x_j^l as a shortened notation for the subsequence $(x_j, x_{j+1}, \dots, x_l)$.

2. For every space of the form $X^{\mathbb{Z}}$ (or $X^{\mathbb{N}}$), $\boldsymbol{\sigma}$ will denote the left-shift operator, defined by
$$(\boldsymbol{\sigma}(\bar{x}))_n = x_{n+1}$$
and $\boldsymbol{\sigma}_-$ will denote the right-shift operator on $X^{\mathbb{Z}}$ (or $X^{\mathbb{Z}_-}$) defined by
$$(\boldsymbol{\sigma}(\bar{x}))_n = x_{n-1}$$

3. All the spaces we discuss are seperable metric spaces, and we always take the Borel σ-algebra on these spaces, which we will denote by $\mathbb{B}$.

4. Recall that if $(X, \mathbb{B}, \mu)$ is a measure space, $(Y, \mathcal{A})$ is a measurable space, and $\rho : X \to Y$ is a measurable function, then *the pushforward of μ by ρ* is the measure ν on Y defined by

$$\int f(y) d\nu(y) = \int f \circ \rho(x) d\mu(x)$$

   We will denote this measure by $\rho\mu$ or $d\rho\mu$.

   If $g : X \to X$ is a non-negative measurable function, then the *multiplication of μ by g*, is the measure ν on X, defined by

$$\int f(x) d\nu(x) = \int f(x) \cdot g(x) d\mu(x)$$

   We denote this measure by $gd\mu$.

5. If $(X, \mathbb{B}, \mu)$ is a probability space, $\mathcal{A} \subseteq \mathbb{B}$ is a sub-σ-algebra, and $B \in \mathbb{B}$, then $\mathbb{P}_\mu(B|\mathcal{A})$ is the function $\mathbb{E}_\mu(1_B|\mathcal{A})$.

   If $Z_1, \ldots, Z_d$ are random variables into a measurable space $(Y, \mathcal{A})$, then we define

$$\mathbb{P}_\mu\left(B|Z_1^d\right) = \mathbb{P}_\mu\left(B|\sigma(Z_1^{-1}\mathcal{A}, \ldots, Z_d^{-1}\mathcal{A})\right)$$

   If $A, B \in \mathbb{B}$, and $\mu(A) > 0$ then $\mathbb{P}_\mu(B|A)$ is not a function, but rather the number $\mathbb{P}_\mu(A|B) = \frac{\mu(A \cap B)}{\mu(A)}$.

1.2 Extended ECPS

In the following we will define Extended ECPS. For an explanation of the relation between them and Furstenberg's ECPS, see [6] or Section 5.

The measures we will perform our 'zooming in' on will be members of $\mathbb{M}(\mathbb{R})$, the space of (positive) Radon measures on $\mathbb{R}$, with the weak topology[2]. The following terminology will be helpful for the construction we will soon describe:

Definition 1.1. *The restriction of $\nu \in \mathbb{M}(\mathbb{R})$ to a Borel set $A \subseteq \mathbb{R}$ will be the measure $1_A d\nu \in \mathbb{M}(\mathbb{R})$.*

Definition 1.2. *We say that ν_1 and ν_2 are similar , if there is an orientation preserving homothety[3] ρ and $\lambda > 0$ such that $d\nu_1 = \lambda d\rho\nu_2$.*

We will discuss measures ν which are normalized so that their restriction to $[0, 1)$ is either zero or a probability measure. More specifically, we define a 'normalizing map' as follows

[2] This is a metrizable, seperable topology in which $\mu_n \to \mu$ iff for every $f \in C_c(\mathbb{R})$, $\int f d\mu_n \to \int f d\mu$. For more details see [8].

[3] In simpler words, ρ is a map of the form $\rho(x) = ax + b$ where $a > 0, b \in \mathbb{R}$

Definition 1.3. *For every Radon measure $\nu \neq 0$ on $\mathbb{R}$, we define*

$$\psi(\nu) = \min\{n \in \mathbb{N} : \nu[-(n-1), n) > 0\}$$

For every $\nu \in \mathbb{M}(\mathbb{R})$ we define

$$N\nu = \begin{cases} \frac{\nu}{\nu[-(\psi(\nu)-1), \psi(\nu))} & \text{if } \nu \neq 0 \\ 0 & \text{if } \nu = 0 \end{cases}$$

We will only discuss measures in the space of normalized measures $N\mathbb{M}(\mathbb{R})$, which is just the image of $\mathbb{M}(\mathbb{R})$ under N.

Now fix a $p \in \mathbb{N} \setminus \{1\}$.

The interval $[0,1)$ can be divided into p intervals of the form $[\frac{i}{p}, \frac{i+1}{p})$. To simplify notation, for $i \in \{0, 1, \dots, p-1\}$ we will write $[i)_p$ instead of $[\frac{i}{p}, \frac{i+1}{p})$. Similarly, for any $n \in \mathbb{N}$, $[0,1)$ can be divided into p^n intervals of equal length, and for $(i_1, i_2, \dots, i_n) \in \{0, 1, \dots, p-1\}^n$ we write

$$[i_1, i_2, \dots, i_n)_{p^n} = [\sum_{k=1}^{n} \frac{i_k}{p^k}, \sum_{k=1}^{n} \frac{i_k}{p^k} + \frac{1}{p^n})$$

We also define ρ_i to be the orientation preserving homothety that takes $[i)_p$ to $[0,1)$, i.e.

$$\rho_i(x) = px - i$$

We now add 'indexes' to the space of normalized measures described above, and consider the space $N\mathbb{M}(\mathbb{R}) \times \{0, 1, \dots, p-1\}^{\mathbb{Z}}$. On this space we define the projection maps

$$\hat{\nu}(\nu, (i_k)_{k \in \mathbb{Z}}) = \nu$$
$$\hat{i}_n(\nu, (i_k)_{k \in \mathbb{Z}}) = i_n$$

We can now define our 'zooming in' map

Definition 1.4. *The 'zooming in' map $M_p^{ext} : N\mathbb{M}(\mathbb{R}) \times \{0, 1, \dots, p-1\}^{\mathbb{Z}} \to N\mathbb{M}(\mathbb{R}) \times \{0, 1, \dots, p-1\}^{\mathbb{Z}}$ will be the map*

$$M_p^{ext}(\nu, (i_k)_{k \in \mathbb{Z}}) = (N\rho_{i_1}\nu, \boldsymbol{\sigma}((i_k)_{k \in \mathbb{Z}}))$$

Note that the restriction of $N\rho_{i_1}\nu$ to $[0,1)$ is a probability measure (or zero) which is similar to the restriction of ν to $[i_1)_{(p)}$. More generally, $\hat{\nu} \circ (M_p^{\text{ext}})^n(\nu, (i_k)_{k \in \mathbb{Z}})$ will be a measure whose restriction to $[0,1)$ is a probability measure (or zero) which is similar to the restriction of ν to $[i_1, i_2, \dots, i_n)_{p^n}$.

We can now define an Extended ECPS

Definition 1.5. *A probability distribution Q^{ext} on $(N\mathbb{M}(\mathbb{R}) \times \{0, 1, \dots, p-1\}^{\mathbb{Z}}, \mathbb{B})$ is called an* adapted distribution*, or a* CP distribution*, if for every $j \in \{0, 1, \dots, p-1\}$ and Q^{ext} almost every $(\nu, (i_n)_{n \leq 0})$,*

$$\mathbb{P}_{Q^{ext}}\left(\hat{i}_1 = j | (\hat{\nu}, (\hat{i}_n)_{n \leq 0})\right)(\nu, (i_n)_{n \in \mathbb{Z}}) = \nu[j)_p$$

In other words, the probability of ‘zooming in’ to $\nu\big|_{[j)_p}$ is exactly $\nu[j)_p$.

Definition 1.6. *A probability distribution Q^{ext} on $(N\mathbb{M}(\mathbb{R}) \times \{0,1,\dots,p-1\}^{\mathbb{Z}}, \mathbb{B})$ is called an* Extended Ergodic CP Distribution *(Extended ECPD) if it is invariant and ergodic with respect to M_p^{ext}, and it is adapted. If this holds,*
$(N\mathbb{M}(\mathbb{R}) \times \{0,1,\dots,p-1\}^{\mathbb{Z}}, M_p^{ext}, Q^{ext})$ is called an Extended ergodic CP system *(Extended ECPS).*

1.3 Main Results

As we discussed earlier, we are interested in the behavior of the measure component of Extended ECPS under translations. We will now define this more precisely:

Definition 1.7. *for every $x \in \mathbb{R}$, let $t_x : \mathbb{R} \to \mathbb{R}$ be the map*

$$t_x(y) = y - x$$

$t_x\nu$ will be the the pushforward of ν by t_x. We also define $t_x^*\nu = Nt_x\nu$.

The maps $\{t_x^*\}_{x\in\mathbb{R}}$ define an action of $\mathbb{R}$ on $N\mathbb{M}(\mathbb{R})$, and Q^{ext} induces a distribution $\hat{\nu}Q^{\text{ext}}$ on $N\mathbb{M}(\mathbb{R})$.

Definition 1.8. *if G is a group which acts measurably on a Borel probability space $(X, \mathbb{B}, \mu)$, where X is a seperable metric space, then*

1. *We say that the action of G on X is conservative with respect to μ, if for every $A \in \mathbb{B}$ with $\mu(A) > 0$, there is a $g \in G \setminus \{1_G\}$, such that $\mu(A \cap gA) > 0$.*
2. *We say that the action of G on X is strictly singular with respect to μ, if for all $g \in G \setminus \{1_G\}$, $g\mu \perp \mu$.*
3. *We say that the action of G on X is recurrent with respect to μ, if for μ almost every x, there is a sequence $(g_n)_{n\in\mathbb{N}} \subseteq G \setminus \{1_G\}$, such that $g_n x \to x$.*

Conservativity implies recurrence, and if G is countable, then conservativity also implies that the action of G on X isn’t strictly singular with respect to μ. However, recurrence, or failing to be strictly singular, does not necessarily imply conservativity[4].

The conservativity of the $\mathbb{Z}$ action on $N\mathbb{M}(\mathbb{R})$ is determined by a property we call determinism:

[4] For example, the action of $\mathbb{Z}\left[\frac{1}{p}\right]$ on $(\mathbb{R}, \mathbb{B}, \frac{1}{2}(\delta_0 + \delta_1))$ by addition is recurrent, is not strictly singular, and yet is not conservative.

Definition 1.9. *We say that a distribution Q^{ext} on $N\mathbb{M}(\mathbb{R}) \times \{0,1,\dots,p-1\}^{\mathbb{Z}}$ is deterministic, if for Q^{ext} almost every $(\nu,(i_l)_{l\in\mathbb{Z}})$,*

$$\mathbb{P}_{Q^{ext}}\left(\hat{i}_1 = i_1 | (\hat{\nu}, (\hat{i}_n)_{n\leq 0})\right)((\nu,(i_l)_{l\in\mathbb{Z}})) = 1$$

Otherwise we say that the distribution is non-deterministic.

We can now state our results. Though our interest is mainly in ECPD, our first result holds also for distributions which are not adapted:

Theorem 1.10. *Let Q^{ext} be a distribution invariant and ergodic under M_p^{ext}. If Q^{ext} is non-deterministic then the translation action of $\mathbb{Z}$ on $N\mathbb{M}(\mathbb{R})$ is conservative with respect to $\hat{\nu}Q^{ext}$.*

As conservativity implies recurrence, one can immediately conclude

Corollary 1.11. *If Q^{ext} is a non-deterministic distribution invariant and ergodic under M_p^{ext}, then for almost every ν, there is a sequence $k_n \in \mathbb{Z}$ such that $t_{k_n}^*\nu \to \nu$ in the weak topology.*

When Q^{ext} is an ECPD (which satisfies an additional technical condition described later), determinism of Q^{ext} implies that Q^{ext} almost every μ is supported on a single point in $[0,1)$, and in particular the $\mathbb{Z}$ action on $N\mathbb{M}(\mathbb{R})$ is strictly singular with respect to $\hat{\nu}Q^{\text{ext}}$, and so in the case of ECPD the converse of Theorem 1.10 also holds.

While the $\mathbb{Z}$ action on $N\mathbb{M}(\mathbb{R})$ is strictly singular with respect to deterministic ECPD, the action of

$$\mathbb{Z}\left[\frac{1}{p}\right] = \{kp^{-n} : \ k \in \mathbb{Z}, n \in \mathbb{N}\}$$

on $N\mathbb{M}(\mathbb{R})$ may be deterministic, and in fact for the $\mathbb{Z}\left[\frac{1}{p}\right]$ action our results can be phrased in the following simple form

Corollary 1.12. *The action of $\mathbb{Z}\left[\frac{1}{p}\right]$ on $N\mathbb{M}(\mathbb{R})$ is conservative with respect to an ECPD (deterministic or non-determinstic) iff the Kalmagorov-Sinai entropy of the ECPD is positive.*

We use the conservativity from the former Theorem to obtain discrete and continuous Pointwise Ergodic Theorems for Extended ECPD:

Theorem 1.13. *If Q^{ext} is a non-deterministic Extended ECPD, then for every $f \in L^1(N\mathbb{M}(\mathbb{R}), \hat{\nu}Q^{ext})$ and $\hat{\nu}Q^{ext}$ almost every ν we have*

$$\lim_N \frac{1}{\nu([0,N))} \sum_{n=0}^{N-1} f(t_n^*\nu)\nu([n,n+1)) = \mathbb{E}_{\hat{\nu}Q^{ext}}(f|\mathbb{J})(\nu)$$

where $\mathbb{J}$ is the σ algebra invariant under a translation map defined in Section 4.

We also obtain a continuous version

Theorem 1.14. *Let Q^{ext} be a non-deterministic Extended ECPD. For every bounded measurable $f : N\mathbb{M}(\mathbb{R}) \to \mathbb{C}$ define $F^f = \int_0^1 f(t_x^*\nu)d\nu(x)$, then for $\hat{\nu}Q^{ext}$ almost every ν,*

$$\lim \frac{1}{\nu[0,T)} \int_0^T f(t_x^*\nu)d\nu(x) = \mathbb{E}_{\hat{\nu}Q^{ext}}(F^f|\mathbb{J})(\nu)$$

Finally, we consider the question of conservativity of the action of $\mathbb{R}$ on $\mathbb{R}$ defined by the translation maps $\{t_x\}_{x\in\mathbb{R}}$, with respect to Q^{ext}-typical measures $\nu \in N\mathbb{M}(\mathbb{R})$. We think of this as a 'pointwise analogue' of the question of conservativity of the action of $\mathbb{R}$ on $N\mathbb{M}(\mathbb{R})$ discussed earlier.

As before, we focus on the action of a countable subgroup $\mathbb{Z}\left[\frac{1}{p}\right]$. The 'pointwise conservativity' of this action is determined by a property we call bilateral determinism. This phrase was introduced by Weiss and Ornstein ([11]) in the context of shift-invariant, ergodic measures on $\{0, 1, \ldots, p-1\}^{\mathbb{Z}}$, to denote the situation in which the 'past' and 'future' of a point determine the 'present', as opposed to the notion of determinism, where the 'past' alone determines the 'present'. In our context:

Definition 1.15. *We say that an Extended ECPD Q^{ext} is bilaterally deterministic, if for every $k_0 \in \mathbb{N}$, and almost every $(\nu, (i_l)_{l\in\mathbb{Z}})$,*

$$\mathbb{P}_{Q^{ext}}\left(\hat{i}_1 = i_1|\hat{\nu}, (\hat{i}_n)_{n\le 0}, (\hat{i}_k)_{k\ge k_0}\right)((\nu, (i_l)_{l\in\mathbb{Z}})) = 1$$

Theorem 1.16. *Let Q^{ext} be an Extended ECPD. If Q^{ext} is bilaterally deterministic, then for Q^{ext} almost every ν, the $\mathbb{Z}\left[\frac{1}{p}\right]$ action on $\mathbb{R}$ is strictly singular with respect to ν.*

Otherwise, for Q^{ext} almost every ν, the $\mathbb{Z}\left[\frac{1}{p}\right]$ action on $\mathbb{R}$ is conservative with respect to ν.

Note that this implies that, while in the general case, the action of $\mathbb{Z}\left[\frac{1}{p}\right]$ may be neither conservative nor strictly singular with respect to some probability measure, in the case of Extended ECPS, the $\mathbb{Z}\left[\frac{1}{p}\right]$ action will always be either strictly singular or conservative, with respect to typical measures.

A canonical example of an Extended ECPS, is an Extended ECPS which arises from (and is isomorphic to) a shift invariant ergodic measure on a symbolic space (see [4],[3], and Subsection 5.3). In this case the concept of bilateral determinism for the shift invariant measure and bilateral determinism for the induced Extended ECPD coincide. One conclusion from this fact is that bilateral determinism of a shift-invariant measure μ is exactly the criterion that determines whether the (typical) prediction measures arising from μ will be conservative with respect to finite coordinate changes. Another conclusion is that since Weiss and Ornstein proved that every shift-invariant ergodic system is isomorphic to a bilaterally deterministic shift invariant ergodic system, it follows that every Extended ECPD arising from a shift invariant measure, is isomorphic to a bilaterally deterministic Extended ECPD, and thus the class of bilaterally deterministic Extended ECPD is 'large'.

We note that determinism implies bilateral determinism, and so 'pointwise conservativity' occurs 'less often' than conservativity. Indeed, while the $\mathbb{Z}$ action on $N\mathbb{M}(\mathbb{R})$ will be conservative with respect to an Extended ECPD Q^{ext}, unless Q^{ext} almost every μ is trivial, in the sense that it is supported on a single point, we saw that we can give a 'large' class of examples of Extended ECPD which are bilaterally deterministic. In fact, we give an example of an Extended ECPD Q^{ext}, which is supported on a family of 'random Cantor measures', and yet not only the action of $\mathbb{Z}\left[\frac{1}{p}\right]$ on $\mathbb{R}$, but also the action of all of $\mathbb{R}$ on $\mathbb{R}$, is strictly singular with respect Q^{ext} almost every μ. Additionally, this singularity is 'strong', in the sense that for every $x \neq 0$, $t_x\mu(\text{supp}\mu) = 0$.

In Section 2 we describe additional 'machinery' needed for proving conservativity and the Ergodic Theorems. We then prove conservativity in Section 3 and the ergodic theorems in Section 4. Finally we discuss 'pointwise conservativity' in Section 5.

2 ECPS Chains

In this Section we introduce dynamical systems we call ECPS Chains, which are (in a sense we will describe soon) equivalent to Extended ECPS. We will then use this equivalence for the proofs of conservativity and the Ergodic Theorems in Sections 3 and 4. In this Section we only give an overview of how this equivalence is established, and we leave the proofs of all Lemmas stated in this Section to Appendix A.

Let $\mathbb{P}_0[0,1) \subseteq N\mathbb{M}(\mathbb{R})$ be the space of measures supported in $[0,1)$, which are either 0 or probability measures. Let $R : N\mathbb{M}(\mathbb{R}) \to \mathbb{P}_0[0,1)$ be the restriction to $[0,1)$, i.e. $R\nu = 1_{[0,1)}\nu$.

For every $\mu \in \mathbb{P}_0[0,1)$ and $0 \leq i \leq p-1$, we define

$$\mu^i = RN\rho_i\mu$$

(note the resemblance to the definition of M_p^{ext}). Define $LS \subseteq (\mathbb{P}_0[0,1) \times \Lambda)^{\mathbb{Z}}$, the space of 'legal sequences', to be

$$LS = \{(\mu_k, i_k)_{k\in\mathbb{Z}} : \ \mu_{k+1} = \mu_k^{i_{k+1}}\}$$

For every $n \in \mathbb{Z}$ we define the projection maps

$$\hat{i}_n((\mu_k, i_k)_{k\in\mathbb{Z}}) = i_n$$

$$\hat{\mu}_n((\mu_k, i_k)_{k\in\mathbb{Z}}) = \mu_n$$

(technically this is problematic since $\hat{i}_n$ is also defined on $N\mathbb{M}(\mathbb{R}) \times \{0, 1, \ldots, p-1\}^{\mathbb{Z}}$, but this should not cause any confusion).

Definition 2.1. *A distribution Q^{chain} on LS is called adapted, if for every $j \in \{0, 1, \dots, p-1\}$,*

$$\mathbb{P}_{Q^{chain}}\left(\hat{i}_1 = j | (\hat{\mu}_n, \hat{i}_n)_{n \le 0}\right)((\mu_n, (i_n)_{n\in\mathbb{Z}})) = \mu_0[j)_p$$

We remark that if Q^{chain} is shift-invariant and adapted then for every $k \in \mathbb{Z}$,

$$\mathbb{P}_{Q^{chain}}\left(\hat{i}_{k+1} = j | (\hat{\mu}_n, \hat{i}_n)_{n \le k}\right)((\mu_n, i_n)_{n\in\mathbb{Z}}) = \mu_k[j)_p$$

Additionally, for every $k \in \mathbb{Z}$, $l > 0$ and $(j_1, \dots, j_l) \in \{0, 1, \dots, p-1\}^l$,

$$\mathbb{P}_{Q^{chain}}\left(\hat{i}_{k+1} = j_1, \dots, \hat{i}_{k+l} = j_l | (\hat{\mu}_n, \hat{i}_n)_{n \le k}\right)((\mu_n, i_n)_{n\in\mathbb{Z}}) = \mu_k[j_1, \dots, j_l)_{p^l} \tag{2.1}$$

(for a proof of this see [4])

Definition 2.2. *A distribution Q^{chain} on LS is called an* Ergodic CP chain distribution *(Chain ECPD) if Q^{chain} is adapted, and invariant and ergodic with respect to the shift operator $\boldsymbol{\sigma}$. If this holds, $(LS, \mathbb{B}, \boldsymbol{\sigma}, Q^{chain})$ is called an* Ergodic CP chain system *(Chain ECPS).*

We now describe a technical condition which we will need to assume in order to get the correspondence we are interested in.

For every $j \in \{0, 1, \dots, p-1\}$, Define[5]

$$E_j = \{(i_k)_{k\in\mathbb{Z}} : \ i_k = j \text{ for all negative enough } k\}$$

We say that a probability distribution Q^{chain} on LS fulfills the Non-Constant Sequence Condition if $\mathbb{P}_0[0,1) \times E_0$ and $\mathbb{P}_0[0,1) \times E_{p-1}$ have zero probability, and we say that a probability distribution Q^{ext} on $N\mathbb{M}(\mathbb{R}) \times \{0, 1, \dots, p-1\}^{\mathbb{Z}}$ fulfills the Non-Constant Sequence Condition if $N\mathbb{M}(\mathbb{R}) \times E_0$ and $N\mathbb{M}(\mathbb{R}) \times E_{p-1}$ have probability zero.

We note that if Q^{ext} is non-deterministic, then it necessarily fulfills the Non-Constant Sequence Condition , since if (w.l.o.g) $N\mathbb{M}(\mathbb{R}) \times E_0$ is not a null set, then the M_p^{ext} invariance and ergodicity can be used to show that necessarily $Q^{\text{ext}}(N\mathbb{M}(\mathbb{R}) \times \{\overrightarrow{0}\}) = 1$ (where $\overrightarrow{0}$ is the point which is 0 in all its coordinates), and in particular for Q^{ext} almost every $(\nu, i_n)_{n\le 0}$, $i_1 = 0$ and

$$\mathbb{P}_{Q^{\text{ext}}}\left(\hat{i}_1 = 0 | (\hat{\nu}, (\hat{i}_n)_{n\le 0})\right)(\nu, (i_n)_{n\in\mathbb{Z}}) = 1$$

which contradicts non-determinism.

[5]A property holds for all 'negative enough' j if there is an $n_0 \in \mathbb{Z}$ such that the property holds for all $j < n_0$.

In order to describe how one can pass from Chain ECPS to ECPS, we will need to introduce some additional terminology. Let $\mathcal{I}$ denote the set of intervals of the form $[a, b)$ (where $b > a$). Every interval $I = [a, b) \in \mathcal{I}$ can be divided into p disjoint intervals in $\mathcal{I}$ with diameter $\frac{b-a}{p}$. For $0 \leq j \leq p-1$ we define I^j to be the $j-th$ interval. (For example, for $I = [0, 1)$, $I^j = [j)_p$).

Definition 2.3. *1. We say that $(I_n)_{n\leq 0} \subseteq \mathcal{I}$ is compatible with $(i_n)_{n\leq 0} \subseteq \{0, 1, \ldots, p-1\}^{\mathbb{Z}}$ if for every $n < 0$,*

$$I_n^{i_{n+1}} = I_{n+1}$$

2. We say that $(I_n)_{n\leq 0}$ is well based if $I_0 = [0, 1)$.

We note that for every sequence $(i_n)_{n\leq 0}$, there is a unique sequence $(I_n)_{n\leq 0}$ which is well based and compatible with $(i_n)_{n\leq 0}$.

Let us denote the projection of E_j onto the non-positive coordinates by E_j^-, i.e.

$$E_j^- = \{(i_n)_{n\leq 0} : \ i_n = j \text{ for all negative enough } n\}$$

We note that if $(i_n)_{n\leq 0} \notin E_0^- \cup E_{p-1}^-$, then the sequence of intervals $(I_n)_{n\leq 0}$ well based and compatible with $(i_n)_{n\leq 0}$ has $\cup_{n\leq 0} I_n = \mathbb{R}$.

Let $\mathcal{H}$ be the group of orientation-preserving homotheties. We note that for every $I, J \in \mathcal{I}$, there is a unique $\rho \in \mathcal{H}$ such that $\rho(I) = J$. We denote this homothety by ρ_I^J.

Now, for every point $(\mu_n, i_n)_{n\in\mathbb{Z}} \in (\mathbb{P}_0[0, 1) \times \{0, 1, \ldots, p-1\})^{\mathbb{Z}}$ we define a measure $\nu \in N\mathbb{M}(\mathbb{R})$ which in fact depends only on the non-positive coordinates $\nu = \nu((\mu_n, i_n)_{n\leq 0})$. This measure 'preserves the information stored in the sequence $(\mu_n)_{n\leq 0}$'.

For a given point $(\mu_n, i_n)_{n\in\mathbb{Z}}$, we first pick the unique sequence $(I_n)_{n\leq 0} \subseteq \mathcal{I}$ which is well based and compatible with $(i_n)_{n\leq 0}$, and then define for every $n \leq 0$,

$$\tilde{\mu}_n = N\rho_{I_0}^{I_n}\mu_n$$

Lemma 2.4. *For every $k \leq n < 0$, there is a $\lambda = \lambda(k) > 0$ such that*

$$\tilde{\mu}_{k-1}\big|_{I_n} = \lambda\tilde{\mu}_k\big|_{I_n}$$

Moreover there is an n_0 such that for all $k < n_0$, $\lambda(k) = 1$.

The idea is that we first define $\tilde{\mu}_0$ to be μ_0. Now the fact that $\mu_{-1}^{i_0} = \mu_0$ enables us to pick a measure $\tilde{\mu}_{-1}$ which is supported on I_{-1}, is similar to μ_{-1}, and its restriction to $I_0 = I_{-1}^{i_0}$ agrees with $\tilde{\mu}_0$ up to a multiplicative constant. We next define $\tilde{\mu}_{-2}$ whose restriction to I_{-1} agrees with $\tilde{\mu}_{-1}$ up to a multiplicative constant etc.

In the following example we show the first three steps of this construction.

Example 2.5. *Let μ_C be the 'Cantor Measure' on $[0,1)$ defined by the property that for every $n > 0$ and every $i_1, i_2, \ldots, i_n \in \{0,1,2\}^n$,*

$$\mu_C[i_1, \ldots, i_n)_{p^n} = \begin{cases} 2^{-n} & \text{if for all } 1 \leq j \leq n,\ i_j \in \{0,2\} \\ 0 & \text{otherwise} \end{cases}$$

Then the Chain distribution $(\delta_{\mu_C} \times (\frac{1}{2}\delta_0 + \frac{1}{2}\delta_2))^{\mathbb{Z}}$ is a shift invariant, ergodic, adapted distribution.

Let $(\mu_n, i_n)_{n\in\mathbb{Z}}$ be a 'typical point', i.e. a point with $\mu_n = \mu_C$ and $i_n \in \{0,2\}$ for every $n \in \mathbb{Z}$. Let us examine the first stages of the construction in the case where $i_0 = 0$ and $i_{-1} = 2$. Let $(I_n)_{n\leq 0}$ be the sequence of intervals which is well based and compatible with $(i_n)_{n\leq 0}$.

Since $(I_n)_{n\leq 0}$ is well based, $I_0 = [0,1)$ and $\tilde{\mu}_0 = \mu_C$.

Since $i_0 = 0$, $I_{-1} = [0,3)$ and

$$\tilde{\mu}_1 = N\rho_{I_0}^{I_{-1}}\mu_{-1} = \mu_C + t_{-2}\mu_C$$

Since $i_{-1} = 2$, $I_{-2} = [-6,3)]$ and

$$\tilde{\mu}_{-2} = (\mu_C + t_{-2}\mu_C) + t_6(\mu_C + t_{-2}\mu_C)$$

$\square$

For every n, m with $n < m < n_0$ (this is the n_0 from Lemma 2.4), we know that for every Borel set $A \subseteq \mathbb{R}$, $\tilde{\mu}_n(A) \geq \tilde{\mu}_m(A)$. Therefore we can define our measure ν by

$$\nu(A) = \lim_{n\to-\infty} \tilde{\mu}_n(A)$$

Define $\Phi : LS \to N\mathbb{M}(\mathbb{R}) \times \{0,1,\ldots,p-1\}^{\mathbb{Z}}$ by

$$\Phi((\mu_n, i_n)_{n\in\mathbb{Z}}) = (\nu((\mu_n, i_n)_{n<0}), (i_n)_{n\in\mathbb{Z}})$$

The correspondence between distributions on $N\mathbb{M}(\mathbb{R}) \times \{0,1,\ldots,p-1\}^{\mathbb{Z}}$ and Chain Distributions is based on the following Lemmas.

Lemma 2.6. $\Phi\boldsymbol{\sigma} = M_p^{ext}\Phi$

Lemma 2.7. *The restriction of Φ to the ($\boldsymbol{\sigma}$-invariant) set*

$$\{(\mu_n, i_n)_{n\in\mathbb{Z}} \in LS :\ (i_n)_{n\in\mathbb{Z}} \notin E_0 \cup E_{p-1}\}$$

is a measurable bijection onto (the M_p^{ext} invariant set)

$$\{(\nu, (i_n)_{n\in\mathbb{Z}}) :\ (i_n)_{n\in\mathbb{Z}} \notin E_0 \cup E_{p-1}\}$$

Now, we note that if Q^{chain} fulfills the Non-Constant Sequence Condition then so does ΦQ^{chain}, and similarly, if Q^{chain} is adapted then so is ΦQ^{chain}. Furthermore, if Q^{chain} fulfills the Non-Constant Sequence Condition then according to the last two Lemmas $(LS, \mathbb{B}, \boldsymbol{\sigma}, Q^{chain})$ and $(N\mathbb{M}(\mathbb{R}) \times \{0, 1, \ldots, p-1\}^{\mathbb{Z}}, \mathbb{B}, M_p^{\text{ext}}, \Phi Q^{chain})$ are measure theoretically isomorphic. Similarly, if Q^{ext} fulfills the Non-Constant Sequence Condition , then Φ^{-1} is well defined on a set of full measure, and the process described above can be reversed.

For our proofs later on, it will be useful to erase the non-positive coordinates in the spaces we just discussed. Let X be the set of points of the form $(\mu_n, i_n)_{n\leq 0} \in (\mathbb{P}_0[0,1) \times \{0, 1, \ldots, p-1\})^{\mathbb{Z}_-}$ satisfying

1. For every $n < 0$, $\mu_{n+1} = \mu_n^{i_{n+1}}$.
2. $(i_n)_{n\leq 0} \notin E_0^- \cup E_{p-1}^-$.

If Q^{chain} is a distribution which fulfills the Non-Constant Sequence Condition, then the subset

$$\{(\mu_n, i_n)_{n\in\mathbb{Z}} \in LS : \ (i_n)_{n\in\mathbb{Z}} \notin E_0 \cup E_{p-1}\}$$

is $\boldsymbol{\sigma}$-invariant and has full measure, and the restriction of the projection onto the non-negative coordinates defined by

$$\pi_-((\mu_n, i_n)_{n\in\mathbb{Z}}) = (\mu_n, i_n)_{n\leq 0}$$

to this subset induces a distribution $Q = \pi_- Q^{chain}$ on X. If Q^{chain} is invariant and ergodic, then Q is invariant and ergodic with respect to the right-shift operator which we denote by $\boldsymbol{\sigma}_-$.

Similarly, let $\tilde{X}$ be the set of points of the form $(\nu, (i_n)_{n\leq 0}) \in N\mathbb{M}(\mathbb{R}) \times \{0, 1, \ldots, p-1\}^{-\mathbb{Z}}$ such that$(i_n)_{n\leq 0} \notin E_0^- \cup E_{p-1}^-$.

If Q^{ext} is a distribution which fulfills the Non-Constant Sequence Condition, then the subset

$$\{(\nu, (i_n)_{n\in\mathbb{Z}}) : \ (i_n)_{n\in\mathbb{Z}} \notin E_0 \cup E_{p-1}\}$$

is a M_p^{ext} invariant subset of full measure, and the restriction of the projection onto the non-negative coordinates $\tilde{\pi}_-$ defined by

$$\tilde{\pi}_-(\nu, (i_n)_{n\in\mathbb{Z}}) = (\nu, (i_n)_{n\leq 0})$$

to this subset induces a distribution $\tilde{Q} = \tilde{\pi}_- Q^{\text{ext}}$ on $\tilde{X}$.

For future reference we note that the map

$$\theta((\mu_n, i_n)_{n\leq 0}) = (\nu((\mu_n, i_n)_{n\leq 0}), (i_n)_{n\leq 0})$$

from X to $\tilde{X}$ is a bijection (the proof is the same as the proof of Lemma 2.7) and has the property

$$\tilde{\pi}_- \Phi = \theta \pi_- \tag{2.2}$$

Additionally, the map $\tilde{\pi}_{\mathbb{M}}(\nu, (i_n)_{n\leq 0}) = \nu$ from $\tilde{X}$ to $N\mathbb{M}(\mathbb{R})$ satisfies

$$\hat{\nu} = \tilde{\pi}_{\mathbb{M}}\tilde{\pi}_{-} \tag{2.3}$$

The following diagram illustrates the relations between the spaces and distributions described above.

$$\begin{tikzcd}
\left(LS, Q^{chain}\right) \arrow[rr, "\simeq", "\Phi"'] \arrow[d, "\pi_-"] & & \left(N\mathbb{M}\left(\mathbb{R}\right) \times \{0,\dots,p-1\}^{\mathbb{Z}}, Q^{ext}\right) \arrow[d, "\tilde{\pi}_-"] \arrow[dd, bend left, "\hat{\nu}"] \\
(X, Q) \arrow[rr, "\simeq", "\theta"'] & & \left(\tilde{X}, \tilde{Q}\right) \arrow[d, "\tilde{\pi}_{\mathbb{M}}"] \\
& & \left(N\mathbb{M}\left(\mathbb{R}\right), \hat{\nu}Q^{ext}\right)
\end{tikzcd}$$

2.1 Translation Maps

In this section we define 'translation maps' on X and $\tilde{X}$.

We recall that we defined translation maps $\{t^*_k\}_{k\in\mathbb{Z}}$ on $N\mathbb{M}(\mathbb{R})$ in Subsection 1.3. We now extend this definition by defining maps s_k on $\{0, 1, \dots, p-1\}^{\mathbb{Z}_-}$, and then define translation maps $\{T_k\}_{k\in\mathbb{Z}}$ on $\tilde{X}$ by

$$T_k(\nu, (i_n)_{n\leq 0}) = (t^*_k\nu, s_k((i_n)_{n\leq 0}))$$

We will then define analogous 'translation maps' on X.

Fix some sequence $(i_n)_{n\leq 0} \notin E^-_0 \cup E^-_{p-1}$. Then the sequence of intervals $(I_n)_{n\leq 0}$ which is compatible with $(i_n)_{n\leq 0}$ has $\cup I_n = \mathbb{R}$, and therefore, for a given $k \in \mathbb{Z}$, there is some $n_0 \leq 0$ such that $[k, k+1) \subseteq I_{n_0}$. There is a $(j_{n_0+1}, j_{n_0+2}, \dots, j_0) \in \{0, 1, \dots, p-1\}^{n_0}$ such that

$$[k, k+1) = (\dots((I_{n_0}^{j_{n_0+1}})^{j_{n_0+2}})\dots)^{j_0}$$

We now define

$$s_k((i_n)_{n\leq 0})_m = \begin{cases} i_m & \text{if } m \leq n_0 \\ j_m & \text{if } m > n_0 \end{cases}$$

which completes the definition of $\{T_k\}_{k\in\mathbb{Z}}$. We note that s_k is defined so that the sequences $(I_n)_{n\leq 0}$ and $(J_n)_{n\leq 0}$ which are well based and compatible with $(i_n)_{n\leq 0}$ and $s_k((i_n)_{n\leq 0})$ respectively, have the property that for all negative enough n, $J_n = I_n - k$. In fact $s_k((i_n)_{n\leq 0})$ is the unique sequence which has this property. We conclude our discussion of translation maps on $\tilde{X}$ with

Lemma 2.8. *The maps $\{T_k\}_{k\in\mathbb{Z}}$ define an action of $\mathbb{Z}$ on $\tilde{X}$.*

The group which will act on X is not $\mathbb{Z}$ but rather a different group we will now describe. $\{0, 1, \ldots, p-1\}^{\mathbb{Z}_-}$ is a group with respect to the addition

$$((i_n)_{n\leq 0} + (j_n)_{n\leq 0})_m = i_m + j_m \mod p$$

and $G = \{(i_n)_{n\leq 0} : \ i_n = 0 \text{ for all negative enough } n\}$ is a subgroup, which is the union of the groups

$$G_m = \{(i_n) : \ i_n = 0 \text{ for every } n < m\}$$

defined for all $m \leq 0$. Returning to the definition of s_k, we see that $s_k((i_n)_{n\leq 0})$ and $(i_n)_{n\leq 0}$ agree for every $n \leq n_0$, and therefore there is an $a \in G_{n_0+1}$ such that $s_k((i_n)_{n\leq 0}) = (i_n)_{n\leq 0} + a$. This fact gives the motivation for the following definition:

For every $m \leq 0$ and $a \in G_m$, the map $S_a : X \to X$ will be defined by

$$S_a((\mu_n, i_n)_{n\leq 0}) = (\nu_n, j_n)_{n\leq 0}$$

where

$$(j_n)_{n\leq 0} = (i_n)_{n\leq 0} + a$$

and for every $n < m$, $\nu_n = \mu_n$, and $\nu_m, \nu_{m+1}, \ldots \nu_0$ are defined recursively by

$$\nu_m = \nu_{m-1}^{j_m}, \nu_{m+1} = \nu_m^{j_{m+1}}, \ldots \nu_0 = \nu_1^{j_0}$$

In other words, $(\nu_n)_{n\leq 0}$ is the unique sequence of measures with $\nu_n = \mu_n$ for all negative enough n, and additionally $(\nu_n, j_n)_{n\leq 0} \in X$.

The relation between the action of G on X and the action of $\mathbb{Z}$ on $\tilde{X}$ is given by

Lemma 2.9. *For every $(\mu_n, i_n)_{n\leq 0} \in X$, the following holds*

1. *For every $k \in \mathbb{Z}$, there is an $a \in G$ such that*

$$\theta S_a((\mu_n, i_n)_{n\leq 0}) = T_k \theta((\mu_n, i_n)_{n\leq 0}) \tag{2.4}$$

2. *For every $a \in G$, there is a $k \in \mathbb{Z}$ such that Equation 2.4 holds.*

3. *Assume that for $a \in G$, $k \in \mathbb{Z}$, Equation 2.4 holds. Then $a = 0$ iff $k = 0$.*

This can be used to show

Lemma 2.10. *The action of G on X is conservative with respect to a probability distribution Q on X, iff the action of $\mathbb{Z}$ on $\tilde{X}$ is conservative with respect to θQ.*

Also, the following holds

Lemma 2.11. *If the action of $\mathbb{Z}$ on $\tilde{X}$ is conservative with respect to $\tilde{Q}$, then the action of $\mathbb{Z}$ on $N\mathbb{M}(\mathbb{R})$ is conservative with respect to $\tilde{\pi}_{\mathbb{M}}\tilde{Q}$.*

Finally, we introduce a notion of determinism for distributions on X. We say that a distribution Q on X is deterministic, if for Q almost every $(\mu_n, i_n)_{n\leq 0}$,

$$\mathbb{P}_Q\left(\hat{i}_0 = i_0 | (\hat{\mu}_n, \hat{i}_n)_{n\leq -1}\right)((\mu_n, i_n)_{n\leq 0}) = 1$$

Otherwise we say that Q is non-deterministic.

The conclusion of this Section is that

Proposition 2.12. *let Q^{ext} be a non-deterministic M_p^{ext} invariant and ergodic probability distribution. If the action of G on X is conservative with respect to $\pi_-\Phi^{-1}Q^{ext}$, then the action of $\mathbb{Z}$ on $N\mathbb{M}(\mathbb{R})$ is conservative with respect to $\hat{\nu}Q^{ext}$.*

Therefore, for the proof of Theorem 1.10 it is sufficient to prove

Proposition 2.13. *Let Q be a right-shift invariant and ergodic distribution on X. Then the action of G on X is conservative with respect to Q iff Q is non-deterministic.*

The proof of Proposition 2.13 is the objective of the next Section. Proposition 2.12 follows immediately from our discussion up to now, as we shall now see. Let Q^{ext} be a non-deterministic M_p^{ext} invariant and ergodic probability distribution. Then in particular Q^{ext} fulfills the Non-Constant Sequence Condition , and therefore the distribution $\pi_-\Phi^{-1}Q^{\text{ext}}$ on X is well defined. if the action of G on X is conservative with respect to $\pi_-\Phi^{-1}Q^{\text{ext}}$, then it follows from Lemma 2.10 that the action of $\mathbb{Z}$ on $\tilde{X}$ is conservative with respect to

$$\theta\pi_-\Phi^{-1}Q^{\text{ext}} =^{Eq.2.2} \tilde{\pi}_-\Phi\Phi^{-1}Q^{\text{ext}} = \tilde{\pi}_-Q^{\text{ext}}$$

Which, by Lemma 2.11 implies that the action of $\mathbb{Z}$ on $N\mathbb{M}(\mathbb{R})$ is conservative with respect to

$$\tilde{\pi}_{\mathbb{M}}\tilde{\pi}_-Q^{\text{ext}} =^{Eq.2.3} \hat{\nu}Q^{\text{ext}}$$

3 Conservativity

In this Section we prove Proposition 2.13, which will conclude the proof of Theorem 1.10, as we discussed in the previous Section. We end the Section with a discussion of conservativity for deterministic distributions.

This proof is an adaptation of Host's proof in [7] that if μ is invariant and ergodic with respect to multiplication by p, then the action of $D = \{kp^{-n} :\ n \in \mathbb{N},\ 0 \leq k < p^n\}$ on $[0, 1)$ is conservative with respect to μ, iff the Kalmagorov-Sinai entropy of μ is larger than zero.

3.1 Conservativity for Increasing Finite Operator Groups

We begin with a general discussion of conservativity for increasing finite operator groups.

Let $(Y, \mathbb{B}, P)$ be a probability space, and for any $n \geq 0$, let G_n be finite groups of measurable maps from Y to Y, such that $G_n \subseteq G_{n+1}$.

Additionally, assume that for every n, there is a measurable finite partition $\gamma^{(n)} = \{R_i^{(n)}\}_{i \leq N_n}$ such that for every $Id \neq T \in G_n$ and $i \leq N_n$, $T(R_i^{(n)}) \cap R_i^{(n)} = \emptyset$.

Define measures P_n by $\int f(y) dP_n(y) = \sum_{T \in G_n} \int f(Ty) dP(y)$. Clearly for any $A \in \mathbb{B}$, $P(A) < P_0(A) < P_1(A) \ldots$ and therefore in particular we have Radon-Nikodym derivatives $\phi_n = \frac{dP}{dP_n}$ and ϕ_n is non-increasing P almost everywhere. Additionally $0 \leq \phi_n \leq 1$.

Finally, define $\mathcal{A}_n$ to be the σ algebra invariant under all $T \in G_n$.

Lemma 3.1. *1. P is strictly singular under G iff for every n, $\phi_n = 1$ P almost everywhere.*

2. P is conservative under G iff $\phi_n \longrightarrow 0$ P a.e.

3. $\mathbb{E}_P(f|\mathcal{A}_n)(y) = \sum_{T \in G_n} f(Ty)\phi_n(Ty)$, in particular for almost every $y \in R_i^{(n)}$, $\phi_n(y) = \mathbb{P}_P(R_i^{(n)}|\mathcal{A}_n)(y)$.

Proof. 1. Assume $A = \{y \in Y : \ \forall n, \phi_n(y) = 1\}$ has $P(A) = 1$. Then for every $n \in \mathbb{N}$,

$$\int 1_A dP = \int 1_A \phi_n dP_n = \int 1_A dP_n = \int 1_A dP + \sum_{T \in G_n \setminus \{Id\}} \int 1_A dTP$$

and therefore for every $T \in G_n \setminus \{Id\}$, $TP(A) = 0$.

Now assume that P is strictly singular under G. Then for every $T \in G \setminus \{Id\}$, there is a Borel set A_T with $P(A_T) = 1$ and $TP(A_T) = 0$. The intersection A of all such sets had $P(A) = 1$ and $TP(A) = 0$ for every $T \in G \setminus \{Id\}$. We claim that for every $n \in \mathbb{N}$ $\phi_n = 1_A$ and therefore $\phi_n = 1$ P almost everywhere, since for every $f \in L^1(P)$ and every $n \in \mathbb{N}$,

$$\int 1_A f dP_n = \int 1_A f dP = \int f dP$$

2. Note that for every n, $\phi_n > 0$ P a.e. Define $K = \{x : \ \forall n, \ \phi_n(x) > 0\}$, then $P(K) = 1$ and $1_K dP_n = \phi_n^{-1} dP$.

Suppose that ϕ_n doesn't tend to zero P a.e., then there exists a $C \subseteq K$ with $P(C) > 0$, and $\epsilon > 0$, such that for every n, $\phi_n > \epsilon$ on C. Therefore for every n

$$\frac{P(C)}{\epsilon} \geq \int_C \phi_n^{-1} dP = P_n(C) = \int \sum_{T \in G_n} 1_C(Tx) dP(x)$$

The sequence $\sum_{T \in G_n} 1_C(Tx)$ is therefore a.e. finite, and since it only accepts integer values, it follows easily that there is an n and $B \subseteq C$ of positive measure s.t. $1_C(Tx) = 0$ for every $T \in G \setminus G_n$ and $x \in B$.

For this n, there is an $i \leq N_n$ such that $A = B \cap R_i^{(n)}$ has positive measure, and since for every $T \in G_n \setminus \{1_G\}$, $R_i^{(n)} \cap TR_i^{(n)} = \emptyset$, we have $\forall T \neq Id,\ A \cap TA = \emptyset$, and so the action of G isn't conservative with respect to P.

Conversely, suppose there is a B of positive measure with

$$\forall T \neq Id,\ P(B \cap TB) = 0 \tag{3.1}$$

Since $P(K) = 1$, $\tilde{B} = B \cap K$ has positive measure while Eq 3.1 still holds when we replace B by $\tilde{B}$. The set $A = \tilde{B} \setminus \cup_{T \neq Id} T\tilde{B}$ is then a set of positive measure contained in K with

$$\forall T \neq Id\ A \cap TA = \emptyset$$

For all n,

$$1 \geq P(\cup_{T \in G_n} TA) = \sum_{T \in G_n} P(TA) = P_n(A) = \int_A \phi_n^{-1} dP$$

and therefore necessarily for P almost every x, $\phi_n(x) \not\to 0$.

3. $F_f = \sum_{T \in G_n} f(Ty)\phi_n(Ty)$ is clearly invariant under any $T \in G_n$ and therefore $F_f \in L^1(\mathcal{A}_n)$. Additionally, for any $A \in \mathcal{A}_n$

$$\int 1_A(y) F_f(y) dP(y) = \int 1_A(y) \sum_{T \in G_n} f(Ty)\phi_n(Ty) dP(y) =$$

$$\int \sum_{T \in G_n} 1_A(Ty) f(Ty)\phi_n(Ty) dP(y) = \int 1_A(y) f(y)\phi_n(y) dP_n(y)$$

$$= \int 1_A(y) f(y) dP(y) \tag{3.2}$$

which proves that indeed $\mathbb{E}(f|\mathcal{A}_n) = F_f$. (Equation 3 can also be used to show that F_f is well defined, i.e. that if $f = g$ P a.e. then $F_f = F_g$ P a.e.)

□

3.2 Proof of Proposition 2.13

We now use the general discussion from the previous Subsection, for the action of the increasing groups G_n on X defined in Subsection 2.1 (Note that increasing here is in the sense $G_{-1} \subseteq G_{-2} \subseteq \ldots$). For every $b = (\ldots, 0, 0, b_n, b_{n+1}, \ldots, b_0) \in G_n$ we define

$$R_b = \{(i_l)_{l \leq 0} :\ \hat{i}_n, \ldots, \hat{i}_0((i_l)_{l \leq 0}) = b_n, \ldots, b_0\}$$

and $\gamma^{(n)} = \{R_b\}_{b \in G_n}$ This is a finite partition of X, and for every $b \in G_n$ and $a \in G_n \setminus \{0\}$,

$$S_a R_b \cap R_b = \emptyset$$

and therefore according to Lemma 3.1, for every $n \leq 0$ and Q almost every $(\mu_l, i_l)_{l\leq 0}$, $\phi_n = \frac{dQ}{d\sum_{a\in G_n} S_a Q}$ is given by

$$\phi_n((\mu_l, i_l)_{l\leq 0}) = P_Q(\hat{i}_n, \ldots, \hat{i}_0 = i_n, \ldots, i_0 | \mathcal{A}_n)((\mu_l, i_l)_{l\leq 0})$$

where $\mathcal{A}_n$ is the σ-algebra invariant under G_n. We note that $\mathcal{A}_n = \boldsymbol{\sigma}_-{}^{-(|n|+1)}\mathbb{B}$ and therefore we may also write

$$\begin{aligned}
\phi_n((\mu_l, i_l)_{l\leq 0}) &= \mathbb{P}_Q\left(\hat{i}_n, \ldots, \hat{i}_0 = i_n, \ldots, i_0 | (\hat{\mu}_l, \hat{i}_l)_{l<n}\right)((\mu_l, i_l)_{l\leq 0}) \\
&= \prod_{j=n}^{0} \mathbb{P}_Q\left(\hat{i}_j = i_j | \hat{i}_n, \ldots, \hat{i}_{j-1}, (\hat{\mu}_l, \hat{i}_l)_{l<n}\right)((\mu_l, i_l)_{l\leq 0}) \\
&=^{(*)} \prod_{j=n}^{0} \mathbb{P}_Q\left(\hat{i}_j = i_j | (\hat{\mu}_l, \hat{i}_l)_{l<j}\right)((\mu_l, i_l)_{l\leq 0}) \\
&=^{(**)} \prod_{j=n}^{0} \phi_0 \circ \boldsymbol{\sigma}_-^{|j|}((\mu_l, i_l)_{l\leq 0})
\end{aligned} \tag{3.3}$$

Where $(*)$ follows from the fact that if $\hat{\mu}_{n-1} = \mu_{n-1}$ and $(\hat{i}_n, \ldots, \hat{i}_0) = (i_n, \ldots, i_0)$ then necessarily

$$\hat{\mu}_n = \mu_{n-1}^{i_n} = \mu_n,\ \hat{\mu}_{n+1} = \mu_n^{i_{n+1}} = \mu_{n+1},\ \text{etc.}$$

and $(**)$ follows from the fact that Q is invariant under the right shift.

According to the Ergodic Theorem, for almost every $(\mu_j, i_j)_{j\leq 0}$,

$$-\lim_{n\to-\infty} \frac{1}{|n|} \log \phi_n(\mu_l, i_l)_{l\leq 0} = -\lim \frac{1}{|n|} \sum_{j=0}^{|n|} \log \phi_0 \circ \boldsymbol{\sigma}_-^j(\mu_l, i_l)_{l\leq 0} = \int -\log \phi_0 dQ$$

and so one of two options hold

1. $-\int \log \phi_0 dQ = 0$ and therefore since we already know that $0 \leq \phi_0 \leq 1$ almost everywhere, necessarily $\phi_0 = 1$ almost everywhere, and so Q is deterministic by definition. Equation 3.3 implies that for every $n \leq 0$, $\phi_n = 1$ Q almost everywhere, and therefore the action of G is strictly singular with respect to Q (by Lemma 3.1).

2. $-\lim_{n\to-\infty} \frac{1}{|n|} \log \phi_n(\mu_j, i_j)_{j\leq 0} = -\int \log \phi_0 dQ > 0$, and therefore the set of points $(\mu_l, i_l)_{l\leq 0}$ for which $\phi_n((\mu_l, i_l)_{l\leq 0}) \to 0$ is necessarily a null set, and so Q isn't deterministic, and the action of G is conservative with respect to Q.

This concludes the proof of Proposition 2.13. □

3.3 Deterministic Distributions

Let us assume that Q^{ext} is a M_p^{ext} invariant and ergodic distribution, which is deterministic and fulfills the Non-Constant Sequence Condition . Then according to Proposition 2.13 the action of G on X isn't conservative with respect to $\pi_{-}\Phi^{-1}Q^{\text{ext}}$, and therefore according to Lemma 2.10 the action of $\mathbb{Z}$ on $\tilde{X}$ isn't conservative with respect to $\theta\pi_{-}\Phi^{-1}Q^{\text{ext}} = \tilde{\pi}_{-}Q^{\text{ext}}$. This however does not automatically imply that the action of $\mathbb{Z}$ on $N\mathbb{M}(\mathbb{R})$ isn't conservative with respect to $\hat{\nu}Q^{\text{ext}}$. Indeed, when Q^{ext} isn't adapted this is not necessarily the case, as the following example shows.

Example 3.2. *Assume $p > 2$, and pick some $2 \leq i < p-1$. Denote the point $(i_n)_{n\in\mathbb{Z}}$ with $i_n = i$ for every $n \in \mathbb{Z}$, by $\vec{i}$. Then*

$$M_p^{ext}(Leb, \vec{i}) = (Leb, \vec{i})$$

and therefore the distribution $\delta_{(Leb,\vec{i})}$ is deterministic, invariant, ergodic, not adapted, and fulfills the Non-Constant Sequence Condition . Since for every $x \in \mathbb{R}$, $t_x^* Leb = Leb$ it follows that the $\mathbb{Z}$ action on $N\mathbb{M}(\mathbb{R})$ is conservative with respect to $\hat{\nu}\delta_{(Leb,\vec{i})}$.

If Q^{ext} is adapted, then Q^{ext} almost every measure is supported on a single point in $[0,1)$, and so the set

$$A = \{\nu : \ \nu \text{ is supported in } [0,1)\}$$

has full measure. For every $k \in \mathbb{Z} \setminus \{0\}$,

$$t_k^* A \cap A = \emptyset$$

and therefore the $\mathbb{Z}$ action on $N\mathbb{M}(\mathbb{R})$ isn't conservative with respect to $\hat{\nu}Q^{\text{ext}}$.

3.4 Sketch of Proof of Corollary 1.12.

Let Q^{ext} be an Extended ECPD. If Q^{ext} is non-deterministic, then Q^{ext} is conservative with respect to the $\mathbb{Z}$ action on $N\mathbb{M}(\mathbb{R})$, and therefore also with respect to the $\mathbb{Z}\left[\frac{1}{p}\right]$ action on $N\mathbb{M}(\mathbb{R})$. In this case the Kalmagorov-Sinai entropy is positive, since for $Q = \pi_{-}\Phi^{-1}Q^{\text{ext}}$, $-\int \log\phi_0 dQ$ as defined above is positive, and can be seen to be the conditional entropy of $(X, \boldsymbol{\sigma}_{-}, Q)$ with respect to the sub σ-algebra $\boldsymbol{\sigma}_{-}^{-1}\mathbb{B}$. Thus, the entropy of Q^{ext}, which is necessarily larger than the conditional entropy of Q, is positive.

If Q^{ext} is deterministic, then a.e. ν is a dirac measure supported on a point in $[0,1)$, and thus typical points are of the form $(\delta_x, (x_n)_{n\in\mathbb{N}})$, where $\sum_{n\in\mathbb{N}} x_n p^{-n} = x$. The map L defined by $L(\delta_x, (x_n)_{n\in\mathbb{N}}) = x$ is a factor map from the Extended ECPS to $([0,1), M_p, LQ^{\text{ext}})$, and (in this case) the factor has the same entropy as the original Extended ECPS. By Host, the action of D on $[0,1)$ is conservative iff LQ^{ext} has positive entropy, and as we discuss in Remark 5.4, for measures supported on $[0,1)$ the action of D on $[0,1)$ (by addition $\mod 1$) is conservative iff the action of $\mathbb{Z}\left[\frac{1}{p}\right]$ on $\mathbb{R}$ is conservative. Finally, the action of $\mathbb{Z}\left[\frac{1}{p}\right]$ on $\mathbb{R}$ is conservative with respect to LQ^{ext} iff the action of $\mathbb{Z}\left[\frac{1}{p}\right]$ on $N\mathbb{M}(\mathbb{R})$ is conservative with respect to Q^{ext}. $\square$

4 Ergodic Theorems

To prove Ergodic theorems for the translation maps described in Section 3, we will use Hurewicz's Ergodic Theorem and the Chacon-Ornstein Lemma for non-singular transformations. We give a brief summary of what we will need from the theory of non-singular transformations, for reference and more details see [1].

4.1 Non-Singular Transformations

Let $(X,\mathbb{B},\mu)$ be a probability space[6], and $T : X \to X$ a measurable invertible operator.

Definition 4.1. *We say that T is non-singular if $T\mu \sim \mu$.*

We note that T is non-singular iff T^{-1} is non-singular. Note also that it is possible that $T\mu$ and μ are not mutually singular, and yet T is still not non-singular. A non singular-transformation T induces an isometry $U_T : L^1(\mu) \to L^1(\mu)$ through[7]

$$U_T(f) = \frac{dT^{-1}\mu}{d\mu} f \circ T$$

We say that T is conservative if the $\mathbb{Z}$ action defined by $\{T^n :\ n \in \mathbb{Z}\}$ is conservative in the sense of Definition 1.8.

We can now state Hurewicz's Ergodic Theorem:

Theorem 4.2 (Hurewicz)**.** *Let $(X,\mathbb{B},\mu)$ be a probability space, and T a conservative, non-singular transformation. Then for every $f,p \in L^1(\mu)$ with $p > 0$*

$$\frac{\sum_{k=0}^{n} U_T^k f(x)}{\sum_{k=0}^{n} U_T^k p(x)} \to \mathbb{E}_{\mu_p}(\frac{f}{p}|\mathbb{J})$$

for a.e. $x \in X$, where $d\mu_p = pd\mu$ and $\mathbb{J}$ is the σ-algebra invariant under T.

We will also use the Chacon-Ornstein Lemma

Lemma 4.3 (Chacon-Ornstein)**.** *For every $f,p \in L^1(\mu)$ with $p > 0$*

$$\frac{U_T^n f(x)}{\sum_{k=0}^{n} U_T^k p(x)} \to 0$$

for a.e. $x \in X$.

---

[6]For the purpose of the discussion in this Subsection only, X will be some general space, and not the space we defined earlier.

[7]In Aaronson's notation in [1] this is $\hat{T}^{-1}$

A simple calculation show that U_T^n is given by

$$U_T^n f = (f \circ T^n)(\frac{dT^{-1}\mu}{d\mu})(\frac{dT^{-1}\mu}{d\mu} \circ T) \dots (\frac{dT^{-1}\mu}{d\mu} \circ T^{n-1}) \tag{4.1}$$

4.2 Proof of Ergodic Theorems

We now prove the Ergodic Theorems described in Subsection 1.3. Fix some non-deterministic Extended ECPD Q^{ext}. Q^{ext} induces a distribution $\tilde{Q} = \tilde{\pi}_{-} Q^{\text{ext}}$ on $\tilde{X}$, and a distribution $Q = \theta^{-1}\tilde{Q}$ on X.

The translation maps $\{T_n\}_{n\in\mathbb{Z}}$ are not necessarily non-singular with respect to $\tilde{Q}$. To see this, we first note that the adaptedness of $\tilde{Q}$ implies that $\tilde{Q}$ and Q give the sets

$$\tilde{Y} = \{(\nu, (i_n)_{n\leq 0}) : \ \nu \neq 0\}$$

$$Y = \{(\mu_n, i_n)_{n\leq 0} : \ \mu_0 \neq 0\}$$

full measure. In Example 2.5 we described the chain distribution $(\delta_{\mu_C} \times (\frac{1}{2}\delta_0 + \frac{1}{2}\delta_2))^{\mathbb{Z}}$ which induces a non-deterministic ECPD which gives the set $A = \{(\nu, (i_n)_{n\leq 0}) : \nu[1, 2) = 0\}$ full measure while

$$T_1 A \subseteq (\tilde{Y})^c$$

and so $T_1\tilde{Q} \perp \tilde{Q}$. However, the set $\tilde{Y}$ and the map T_1 induce a map $T : \tilde{Y} \to \tilde{Y}$ which is non-singular. To define T, we first define

$$\tau(\nu) = \min\{n \in \mathbb{N} : \ T_n(\nu, (i_n)_{n\leq 0}) \in \tilde{Y}\}$$

Lemma 4.4. *For almost every ν, $\tau(\nu) < \infty$.*

Proof. Define

$$g(\nu) = \max\{n \in \mathbb{Z} : \ \nu[n, n + 1) > 0\}$$

Since $\tilde{Y}$ has full measure it is sufficient to show that,

$$\tilde{Q}(\{(\nu, (i_n)_{n\leq 0}) : \ \tau(\nu) = \infty\} \cap \tilde{Y}) = \tilde{\pi}_{\mathbb{M}}\tilde{Q}(\{\nu : \ \nu[0, 1) > 0 \text{ and } \tau(\nu) = \infty\}) =$$

$$\tilde{\pi}_{\mathbb{M}}\tilde{Q}(\{\nu : \ g(\nu) = 0\}) = 0$$

Indeed, this must be the case as otherwise, since we know the action of $\mathbb{Z}$ on $N\mathbb{M}(\mathbb{R})$ is conservative with respect to $\tilde{\pi}_{\mathbb{M}}\tilde{Q}$, there is some $k \in \mathbb{Z} \setminus \{0\}$ with

$$\tilde{\pi}_{\mathbb{M}}\tilde{Q}(\{\nu : \ g(\nu) = 0\} \cap t_k^*\{\nu : \ g(\nu) = 0\}) > 0$$

but $t_k^*\{\nu : \ g(\nu) = 0\} = \{\nu : \ g(\nu) = -k\}$ and therefore the intersection above is empty, which gives a contradiction. $\square$

We now define

$$T(\nu, (i_n)_{n\leq 0}) = T_{\tau(\nu)}(\nu, (i_n)_{n\leq 0})$$

Similarly, we define

$$\tau_-(\nu) = \min\{n \in \mathbb{N} : \ T_{-n}(\nu, (i_n)_{n\leq 0}) \in \tilde{Y}\}$$

For almost every ν, $\tau_-(\nu) > -\infty$ and it can be verified that the inverse of T is given by

$$T^{-1}(\nu, (i_n)_{n\leq 0}) = T_{-\tau_-(\nu)}(\nu, (i_n)_{n\leq 0})$$

At the end of this section we will prove

Lemma 4.5. *T is a non-singular transformation. Moreover*

$$\varphi \equiv \frac{dT^{-1}\tilde{Q}}{d\tilde{Q}}(\nu, (i_n)_{n\leq 0}) = \nu[\tau(\nu), \tau(\nu)+1)$$

Since $\{T_n : \ n \in \mathbb{Z}\}$ is conservative, also T is conservative. And thus we can use the Hurewicz Ergodic Theorem. To do so we will first calculate $U_T^n f$.

Define $\tau_n(\nu)$ recursively by $\tau_1 = \tau$ and

$$\tau_n(\nu) = \tau_{n-1}(\nu) + \tau(t^*_{\tau_{n-1}(\nu)}\nu)$$

Note that if $k \in \mathbb{N}$ and $k \leq \tau_n(\nu)$, then $\nu[k, k+1) > 0$ iff there is a $1 \leq j \leq n$ such that $k = \tau_j(\nu)$. To keep notation uncluttered, we write τ_n instead of $\tau_n(\nu)$. Note that

$$\varphi \circ T^n(\nu, (i_n)_{n\leq 0}) = \varphi(t^*_{\tau_n}\nu) = t^*_{\tau_n}\nu[\tau(t^*_{\tau_n}\nu), \tau(t^*_{\tau_n}\nu)+1) = \frac{\nu[\tau_{n+1}, \tau_{n+1}+1)}{\nu[\tau_n, \tau_n+1)}$$

and so Eq 4.1 gives us

$$U_T^n f(\nu, (i_k)_{k\leq 0}) = f \circ T^n(\nu, (i_k)_{k\leq 0}) \cdot \nu[\tau_1, \tau_1+1) \prod_{j=1}^{n-1} \frac{\nu[\tau_{j+1}, \tau_{j+1}+1)}{\nu[\tau_j, \tau_j+1)}$$

$$= f \circ T^n(\nu, (i_k)_{k\leq 0}) \cdot \nu[\tau_n, \tau_n+1)$$

Now, applying the Hurewicz Theorem for $p = 1$ we get for a.e. $(\nu, (i_n)_{n\leq 0})$,

$$\tilde{A}_n^f(\nu, (i_n)_{n\leq 0}) \equiv \frac{1}{\nu[0, \tau_n)} \sum_{k=0}^{n-1} f \circ T^k(\nu, (i_n)_{n\leq 0}) \cdot \nu(\tau_k, \tau_{k+1}) \to \mathbb{E}_{\tilde{Q}}(f|\mathbb{J})(\nu, (i_n)_{n\leq 0})$$

Define

$$A_m^f(\nu, (i_n)_{n\leq 0}) = \frac{1}{\nu[0, m)} \sum_{j=0}^{m-1} f \circ T_j(\nu, (i_n)_{n\leq 0}) \cdot \nu[j, j+1)$$

which is well defined whenever $m \geq \tau_1$. If $\tau_n \leq m < \tau_{n+1}$ then $A^f_m = \tilde{A}^f_n$ and so for a.e. $(\nu, (i_n)_{n\leq 0})$,

$$A^f_m(\nu, (i_n)_{n\leq 0}) \to \mathbb{E}_{\tilde{Q}}(f|\mathbb{J})(\nu, (i_n)_{n\leq 0})$$

Theorem 1.13 is just the special case in which $f(\nu, (i_n)_{n\leq 0}) = f(\nu)$.

Let $f = f(\nu)$ be a bounded measurable function, for $a < b \in \mathbb{R}$ define

$$A^b_a(f)(\nu) = \frac{1}{\nu[a,b)} \int_a^b f(t^*_x \nu) d\nu(x)$$

We will show that A^x_0 converges a.e. (Theorem 1.14). We define $F^f = A^1_0$ and note that it too is bounded and measurable. However, if $f = g$ $\tilde{Q}$ a.e., it is not necessarily true that $F^f = F^g$ $\tilde{Q}$ a.e. To see this we return to the non-deterministic ECPD induced by the chain distribution $(\delta_{\mu_C} \times (\frac{1}{2}\delta_0 + \frac{1}{2}\delta_2))^{\mathbb{Z}}$ from Example 2.5, and note that the set

$$B = \{\nu : \ \forall \epsilon > 0, \ \nu[0,\epsilon), \ \nu(1-\epsilon, \epsilon) > 0\}$$

has full measure, but for every $x \in \mathbb{R} \setminus \mathbb{Z}$, $B + x \cap B$ is a null set. Thus we can pick $f = 1$ and $g = f 1_B$, and obtain $F^f = 1$ and $F^g = 0$.

Nonetheless, we can still apply the ergodic theorem we just proved to F^f, we obtain

$$A^N_0(f)(\nu) = A^{Ff}_N(\nu) \to \mathbb{E}_{\tilde{Q}}(F^f|\mathbb{J})(\nu)$$

for a.e. ν. To extend this to any $x \in \mathbb{R}$, note that if we take $f = p = 1$ in the Chacon-Ornstein Lemma, we obtain

$$\frac{\nu[\tau_n, \tau_n + 1)}{\nu[0, \tau_n + 1)} \to 0$$

and therefore

$$\frac{\nu[x, \lceil x \rceil)}{\nu[0, \lceil x \rceil)} \to 0$$

using this and the fact that for any $0 < x < N$,

$$A^N_0(f)(\nu) = \frac{\nu[0,x)}{\nu[0,N)} A^x_0(f)(\nu) + \frac{\nu[x,N)}{\nu[0,N)} A^N_x(f)(\nu)$$

it is not difficult to see that A^x_0 converges to the same limit as A^N_0 does.

Proof of Lemma 4.5. Notation in this proof: We denote the multiplication of f and g by fg and their composition by $f \circ g$. Recall that the pushforward of a measure P by f we will denote by fP or dfP, while the multiplication of P by f will be fdP.

It is sufficient to show that

$$\frac{dT^{-1}\tilde{Q}}{d\tilde{Q}}(\nu, (i_n)_{n\leq 0}) = \nu[\tau(\nu), \tau(\nu) + 1)$$

and then since $\frac{dT^{-1}\tilde{Q}}{d\tilde{Q}} > 0$ it follows that $T^{-1}\tilde{Q} \sim \tilde{Q}$.

As in the proof of conservativity, in our calculation of $\frac{dT^{-1}\tilde{Q}}{d\tilde{Q}}$ we use the 'translation maps' $\{S_a\}_{a\in G}$ on X. This can be done since the set of full measure $\tilde{Y}$ can be divided into a countable number of disjoint sets

$$\tilde{Y}_a = \{(\nu, (i_n)_{n\le 0}) : \ T^{-1}(\nu, (i_n)_{n\le 0}) = \theta \circ S_a \circ \theta^{-1}(\nu, (i_n)_{n\le 0})\}$$

Fix some $a \in G$, and pick an $n \le 0$ so that $a \in G_n$. We recall that, for $Q_n = \sum_{b\in G_n} S_b Q$ and $\phi_n = \frac{dQ}{dQ_n}$, ϕ_n is given by

$$\begin{aligned}\phi_n((\mu_l, i_l)_{l\le 0}) &=^{\text{Eq 3.3}} \mathbb{P}_Q(\hat{i}_n, \ldots, \hat{i}_0 = i_n, \ldots, i_0 | (\hat{\mu}_l, \hat{i}_l)_{l<n})((\mu_l, i_l)_{l\le 0}) \\ &=^{\text{Eq 2.1}} \mu_{n-1}[i_n, i_{n+1}, \ldots, i_0)_{p^{|n|+1}}\end{aligned}$$

Since Q_n is S_a invariant,

$$\frac{dS_aQ}{dQ_n} = \phi_n \circ S_{-a}$$

Additionally, note that for all $(\mu_l, i_l)_{l\le 0} \in Y$, $\phi_n((\mu_l, i_l)_{l\le 0}) > 0$, and therefore

$$1_Y dS_aQ \ll 1_Y dQ_n \sim 1_Y dQ = dQ$$

and so

$$1_Y dS_aQ = \frac{dS_aQ}{dQ_n}(\frac{dQ}{dQ_n})^{-1} 1_Y dQ = \frac{\phi_n \circ S_{-a}}{\phi_n} dQ \tag{4.2}$$

This can be used to show that for every $n \le 0$ and $a \in G_n$,

$$1_{T^{-1}\tilde{Y}_a} dT^{-1}\tilde{Q} = 1_{T^{-1}\tilde{Y}_a} \frac{\phi_n \circ \theta^{-1} \circ T}{\phi_n \circ \theta^{-1}} d\tilde{Q} \tag{4.3}$$

We leave this computation to the end of the proof. We now obtain $\frac{dT^{-1}\tilde{Q}}{d\tilde{Q}}$ through

$$\begin{aligned}dT^{-1}\tilde{Q} = \sum_{a\in G} 1_{T^{-1}\tilde{Y}_a} dT^{-1}\tilde{Q} = \sum_{a\in G} 1_{T^{-1}\tilde{Y}_a} \frac{\phi_n \circ \theta^{-1} \circ T}{\phi_n \circ \theta^{-1}} d\tilde{Q} = \\ \frac{\phi_n \circ \theta^{-1} \circ T}{\phi_n \circ \theta^{-1}} d\tilde{Q}\end{aligned}$$

We now compute $\frac{\phi_n \circ \theta^{-1} \circ T}{\phi_n \circ \theta^{-1}}$. Let $(\nu, (i_l)_{l\le 0}) = \theta((\mu_l, i_l)_{l\le 0})$ be some point in Y. and let $(I_l)_{l\le 0}$ be the sequence of intervals which is well based and compatible with $(i_l)_{l\le 0}$. Then

$$\phi_n \circ \theta^{-1}(\nu, (i_l)_{l\le 0}) = \mu_{n-1}[i_n, \ldots, i_0)_{p^{|n|+1}} = \frac{\nu[0,1)}{\nu(I_{n-1})} = \frac{1}{\nu(I_{n-1})}$$

and (for $\tau = \tau(\nu)$)

$$\phi \circ \theta^{-1} \circ T(\nu, (i_l)_{l\le 0}) = \frac{1}{t_\tau^* \nu(I_{n-1} - \tau)} = \frac{\nu[\tau, \tau+1)}{\nu(I_{n-1})}$$

Which shows that indeed $\frac{dT^{-1}\tilde{Q}}{d\tilde{Q}}(\nu,(i_n)_{n\leq 0}) = \nu[\tau(\nu),\tau(\nu)+1)$.

We now go back to proving Equation 4.3. For every $f \in L^1(\tilde{Q})$ we have

$$\int 1_{T^{-1}\tilde{Y}_a} \cdot f dT^{-1}\tilde{Q} = \int 1_{\tilde{Y}_\alpha} \cdot (f \circ T^{-1}) d\tilde{Q} =$$
$$\int 1_{\tilde{Y}_a} \cdot (f \circ \theta \circ S_a \circ \theta^{-1}) d\theta Q = \int (1_{\tilde{Y}_a} \circ \theta) \cdot (f \circ \theta \circ S_a) dQ =$$
$$\int (1_{\tilde{Y}_a} \circ \theta \circ S_{-a}) \cdot (f \circ \theta) dS_a Q =^{(*)} \int (1_{\tilde{Y}_a} \circ \theta \circ S_{-a}) \cdot (f \circ \theta) \cdot 1_Y dS_a Q$$
$$=^{Eq.\ 4.2} \int (1_{\tilde{Y}_a} \circ \theta \circ S_{-a}) \cdot (f \circ \theta) \cdot (\frac{\phi_n \circ S_{-a}}{\phi_n}) dQ$$
$$= \int (1_{\tilde{Y}_a} \circ \theta \circ S_{-a} \circ \theta^{-1}) \cdot f \cdot (\frac{\phi_n \circ S_{-a} \circ \theta^{-1}}{\phi_n \circ \theta^{-1}}) d\tilde{Q}$$
$$=^{(**)} \int 1_{T^{-1}\tilde{Y}_a} \cdot f \cdot (\frac{\phi_n \circ \theta^{-1} \circ T}{\phi_n \circ \theta^{-1}}) d\tilde{Q}$$

Equality $(*)$ follows from the fact that

$$1_{\tilde{Y}_a} \circ \theta \circ S_{-a} = 1_{S_a \theta^{-1}\tilde{Y}_a} = 1_{\theta^{-1}T^{-1}\tilde{Y}_a} \tag{4.4}$$

and

$$\theta^{-1}T^{-1}\tilde{Y}_a \subseteq \theta^{-1}T^{-1}\tilde{Y} \subseteq \theta^{-1}\tilde{Y} \subseteq Y$$

and therefore

$$(1_{\tilde{Y}_a} \circ \theta \circ S_{-a}) \cdot 1_Y = 1_{\tilde{Y}_a} \circ \theta \circ S_{-a}$$

Equality $(**)$ follows from the fact that

$$1_{\tilde{Y}_a} \circ \theta \circ S_{-a} \circ \theta^{-1} =^{\text{Eq 4.4}} 1_{\theta^{-1}T^{-1}\tilde{Y}_a} \circ \theta^{-1} = 1_{T^{-1}\tilde{Y}_a}$$

and from the fact that if $(\nu,(i_l)_{l\leq 0}) \in T^{-1}\tilde{Y}_a$ then

$$\theta^{-1} \circ T((\nu,(i_l)_{l\leq 0})) = S_{-a} \circ \theta^{-1}((\nu,(i_l)_{l\leq 0}))$$

□

5 Pointwise Conservativity and Strict Singularity

Our aims in this Section is to prove Theorem 1.16, which shows that the condition for conservativity of the action of $\mathbb{Z}\left[\frac{1}{p}\right]$ on $\mathbb{R}$ with respect to typical measures is bilateral determinism, to discuss the connection between bilateral determinism as defined here and bilateral determinism as defined by Ornstein and Weiss for shift-invariant measures, in the case of an Extended ECPD induced by a shift invariant ergodic measure, and finally to discuss and present the example described in Subsection 1.3 of an Extended ECPD with the property that for almost every ν, the action of $\mathbb{R}$ on $\mathbb{R}$ is strictly singular with respect to ν.

We begin this Section by defining ECPS (Ergodic Conditional Probability Systems), which are the original[8] systems defined by Furstenberg in [4]. We will then discuss the relationship between ECPS and Extended ECPS, and show how the statements described in the previous paragraph can be proved via proving analogous claims for ECPS. We will then discuss these claims on ECPS.

5.1 ECPS

Fix $p \in \mathbb{N} \setminus \{1\}$. Define $M_p : [0,1) \to [0,1)$ by

$$M_p(x) = px \mod p$$

We also define

$$E_{p-1}^{+} = \{(i_n)_{n\in\mathbb{N}} : \quad \text{for all large enough } n, i_n = p-1\}$$

Every number $x \in [0,1)$ has a unique representation $x = \sum_{n\in\mathbb{N}} x_n p^{-n}$ such that the sequence $(x_n)_{n\in\mathbb{N}}$ is not in E_{p-1}^{+}. We call this representation the *standard representation of x in base p* .

Consider the space $\mathbb{P}_0[0,1) \times [0,1)$. Define the projection $\hat{\mu}$ on $\mathbb{P}_0[0,1) \times [0,1)$ by $\hat{\mu}(\mu,x) = \mu$, and let $\hat{x}_n(\mu,x)$ be the n-th coordinate in the standard representation of x. Now define a map $M_p^{CP} : \mathbb{P}_0[0,1) \times [0,1) \to \mathbb{P}_0[0,1) \times [0,1)$ by

$$M_p^{CP}(\mu,x) = (\mu^{x_1}, M_p x_1)$$

A distribution P on $\mathbb{P}_0[0,1) \times [0,1)$ is called adapted, if for every $j \in \{0,1,\ldots,p-1\}$, for P almost every (μ,x),

$$\mathbb{P}_P\left(\hat{x}_1 = j|\hat{\mu}\right)(\mu,x) = \mu[j)_p$$

If P is also M_p^{CP} invariant, this property implies that for every $(j_1,j_2,\ldots,j_n)$ in $\{0,1,\ldots,p-1\}^n$, for P almost every (μ,x),

$$\mathbb{P}_P\left(\hat{x}_1,\ldots,\hat{x}_n = j_1,\ldots,j_n|\hat{\mu}\right)(\mu,x) = \mu[j_1,\ldots,j_n)_{p^n} \tag{5.1}$$

Note that the set $\{(\mu,x) : \quad \hat{x}_1(\mu,x),\ldots,\hat{x}_n(\mu,x) = j_1,\ldots,j_n\}$ is exactly the set $\mathbb{P}_0[0,1) \times [j_1,\ldots,j_n)_{p^n}$ and therefore Equation 5.1 implies that the conditional measures of P with respect to the σ algebra generated by $\hat{\mu}$ at the point (μ,x), are exactly μ. In other words, for every $f \in L^1(P)$,

$$\int\int f(\mu,x)d\mu(x)dP(\mu) = \int f(\mu,x)dP(\mu,x)$$

[8]The definition we give here is not exactly the same is Furstenberg's, but the definitions conincide, with the exception of the distribution $\delta_{(\delta_1,1)}$ which is an ECPD in Furstenberg's definitions, while in our definition $(\delta_1,1)$ is not in our space.

In particular, P gives a set $A \subseteq \mathbb{P}_0[0,1) \times [0,1)$ zero mass, iff for P almost every μ, μ gives the set $A(\mu) = \{x \in [0,1) : \ (\mu, x) \in A\}$ zero mass. Thus P gives the set

$$\{(\mu, x) : \text{ for all } n, \mu[x_1, \ldots, x_n)_{p^n} > 0\}$$

full measure.

Definition 5.1. *A distribution P on $\mathbb{P}_0[0,1) \times [0,1)$ is called an Ergodic Conditional Probability Distribution (ECPD), if it is an adapted distribution, which is invariant and ergodic with respect to M_p^{CP}.*

If P is an ECPD we call $(\mathbb{P}_0[0,1) \times [0,1), \mathbb{B}, P, M_p^{CP})$ an Ergodic Conditional Probability System (ECPS).

Finally,

Definition 5.2. *We say that an ECPD P is bilaterally deterministic, if for P almost every (μ, x), and every $0 \leq j \leq p-1$,*

$$\mathbb{P}_P\left(\hat{x}_1 = x_1 | \hat{\mu}, \cap_{n=2}^{\infty} \sigma(\hat{x}_n^{\infty})\right)((\mu, x)) = \mathbb{P}_\mu\left(\hat{x}_1 = x_1 | \cap_{n=2}^{\infty} \sigma(\hat{x}_n^{\infty})\right)((\mu, x)) = 1$$

(Here also we use an abuse of notation, and think of $\hat{x}_n$ also as maps on $[0,1)$, and not only as maps on $\mathbb{P}_0[0,1) \times [0,1)$).

The natural invertible extension of an ECPS can be realized as a Chain ECPS (as we do not need this, we do not discuss this in detail). On the other hand, consider the map $\Psi : (LS \setminus \{(\mu_l, i_l)_{l \in \mathbb{Z}} : \ (\hat{i}_n((\mu_l, i_l)_{l \in \mathbb{Z}}))_{n \in \mathbb{N}} \in E_{p-1}^+\}) \to \mathbb{P}_0[0,1) \times [0,1)$ defined by

$$\Psi((\mu_l, i_l)_{l \in \mathbb{Z}}) = (\mu_0, \sum_{j=1}^{\infty} i_j p^{-j})$$

Then $\Psi \boldsymbol{\sigma} = M_p^{CP} \Psi$. If Q^{chain} is a Chain ECPD which satisfies the Non-Constant Sequence Condition, then (using the Shift invariance of Q^{chain}) for almost every $(\mu_l, i_l)_{l \in \mathbb{Z}}$,

$$(\hat{i}_n((\mu_l, i_l)_{l \in \mathbb{Z}}))_{n \in \mathbb{N}} \notin E_{p-1}^+$$

and therefore Q^{chain} is supported on the domain of Ψ, and so ΨQ^{chain} is well defined, and is an ECPD. Thus there is a one-to-one correspondence between ECPD and Chain-ECPD.

The interval $[0,1)$ is a group with the addition operation

$$x +_{\mathbb{T}} y = x + y \mod p$$

accordingly, we define $t_x^{\mathbb{T}} : [0,1) \to [0,1)$ by $t_x^{\mathbb{T}}(y) = y -_{\mathbb{T}} x$. This defines an action of subgroups $G \subseteq [0,1)$ on $[0,1)$ via the maps $\{t_x^{\mathbb{T}}\}_{x \in G}$. Recall that we defined D to be the group

$$D = \{kp^{-n} : \ n \in \mathbb{N}, 0 \leq k \leq p^n - 1\}$$

In Subsection 5.2 we will show

Lemma 5.3. *If P is a bilaterally deterministic ECPD, then for P almost every μ, the action of D on $[0,1)$ is strictly singular with respect to μ. Otherwise, for P almost every μ, the action of D on $[0,1)$ is conservative with respect to μ.*

We will additionally give an example of an ECPD P_{sing}, such that for P_{sing} almost every measure μ, the action of $[0,1)$ on $[0,1)$ is strictly singular with respect to μ, and moreover, for all $x \neq 0$, $t_x^{\mathbb{T}}\mu(\text{supp}\mu) = 0$.

Lemma 5.3 can be used to show that Theorem 1.16 holds, and similarly the Extended ECPS which arises as a realization of the natural invertible extension of the ECPS $(\mathbb{P}_0[0,1) \times [0,1), \mathbb{B}, M_p^{CP}, P_{sing})$, where P_{sing} is the ECPD described above, has the property that for almost every ν, the action of $\mathbb{R}$ on $\mathbb{R}$ is strictly singular with respect to ν, and moreover, for all $x \neq 0$, $t_x\nu(\text{supp}\nu) = 0$.

We do not give a full proof of how to pass from our results on translations on $\mathbb{T}$, to the results regarding translations on $\mathbb{R}$, but we mention that such a proof is based on the following observations

Remark 5.4.

1. *For every $A \subseteq [0,1)$ and $\alpha \in [0,1)$,*

$$(A + \alpha + \mathbb{Z}) \cap [0,1) = A +_{\mathbb{T}} \alpha$$

   This can be used to show, that if μ is supported in $[0,1)$, then the action of some group $G \subseteq \mathbb{T}$ (by addition in $\mathbb{T}$) is conservative/strictly singular with respect to μ iff the action of $G + \mathbb{Z}$ (by addition in $\mathbb{R}$) on $\mathbb{R}$ is conservative/strictly singular with respect to μ. Note that $D + \mathbb{Z} = \mathbb{Z}\left[\frac{1}{p}\right]$ and $[0,1) + \mathbb{Z} = \mathbb{R}$.

2. *Let $A \subseteq \mathbb{R}$ be a Borel set, $\mu \in \mathbb{M}(\mathbb{R})$, $x \in \mathbb{R}$, and let $\rho : \mathbb{R} \to \mathbb{R}$ be of the form $\rho(x) = p^n x + b$, then*

$$\rho\mu(\rho A + p^n x) = \mu(A + x)$$

   This can be used to show that if $G \subseteq \mathbb{R}$ is invariant under multiplication by p (and this is indeed the case for both $\mathbb{Z}\left[\frac{1}{p}\right]$ and $\mathbb{R}$), then if the action of G on $\mathbb{R}$ is conservative/strictly singular with respect to μ, then it will also be conservative/strictly singular with respect to $\rho\mu$.

We now explain how this can be used to show that if Q^{ext} is not a bilaterally deterministic distribution, then the action of $\mathbb{Z}\left[\frac{1}{p}\right]$ on $\mathbb{R}$ is conservative with respect to Q^{ext} almost every ν.

Assume Q^{ext} is not bilaterally deterministic. Then $(N\mathbb{M}(\mathbb{R}) \times \{0,1,\dots,p-1\}^{\mathbb{Z}}, \mathbb{B}, M_p^{\text{ext}}, Q^{\text{ext}})$ is measure theoretically isomorphic to $(LS, \mathbb{B}, \boldsymbol{\sigma}, Q^{chain} = \Phi^{-1}Q^{\text{ext}})$, and $(\mathbb{P}_0[0,1) \times [0,1), \mathbb{B}, M_p^{CP}, P = \Psi Q^{chain})$ is a factor of both isomorphic systems, and the fact that Q^{ext} isn't bilaterally deterministic, implies that P is not bilaterally deterministic either. Therefore by Lemma 5.3, for typical (μ, x), the action of D on $[0,1)$ is conservative with respect to μ, and by the first remark above, this implies that the action of $D + \mathbb{Z} = \mathbb{Z}\left[\frac{1}{p}\right]$ on $\mathbb{R}$ is conservative with repsect to typical μ. This implies that for typical $(\mu_l, i_l)_{l\in\mathbb{Z}}$, the action of $\mathbb{Z}\left[\frac{1}{p}\right]$ on $\mathbb{R}$ is conservative with respect to μ_0, and therefore also with respect to all μ_n by shift invariance. Now typical $(\nu, (i_n)_{n\in\mathbb{Z}}) \in N\mathbb{M}(\mathbb{R}) \times \{0, 1, \ldots, p-1\}^{\mathbb{Z}}$ are of the form $\Phi((\mu_l, i_l)_{l\in\mathbb{Z}})$, where for every n, conservativity with respect to μ_n holds. Using the second remark, the action of $\mathbb{Z}\left[\frac{1}{p}\right]$ on $\mathbb{R}$ will also be conservative with respect to $\tilde{\mu}_n$, which are normalized version of measures of the form $\rho\mu_n$, and therefore also with respect to $\nu = \lim \tilde{\mu}_n$.

5.2 Conservativity and Strict Singularity for ECPS

Proof of Lemma 5.3 Let P be an ECPD. For every $n \in \mathbb{N}$ define $D_{p^n} = \{kp^{-n} : 0 \le k \le p^n - 1\}$. This is an increasing sequence of finite groups, whose union is the group D defined in the previous section.

We also define for every $n \in \mathbb{N}$ a partition $\mathbb{D}_{p^n} = \{R_a\}_{a\in\{0,1,\ldots,p-1\}^n}$ of $[0, 1)$ where

$$R_a = \{x : x_1^n = a\}$$

note that for all $a, b \in \{0, 1, \ldots, p-1\}^n$ with $b \neq 0$, $t_b R_a \cap R_a = \emptyset$.

Now, for every (μ, x), and every $n \in \mathbb{N}$ define the functions

$$\phi_n(\mu, x) = \frac{d\mu}{d\sum_{a\in D_{p^n}} t_a\mu}$$

For every μ, and for μ almost every x, $\phi_n(\mu, x)$, is a non-increasing sequence, and therefore we can also define

$$\phi = \lim_{n\to\infty} \phi_n$$

Note that for every n, the σ-algebra invariant under the action of D_{p^n} is exactly the σ-algebra generated by $\hat{x}_n, \hat{x}_{n+1}, \ldots$ We can now apply Lemma 3.1 (in fact this is exactly the context in which this lemma was originally formulated by Host) to deduce that

$$\phi_n((\mu, x)) = \mathbb{P}_\mu\left(\hat{x}_1^n = x_1^n | \hat{x}_{n+1}^\infty\right)(\mu, x)$$

and additionally, the action of D on μ is strictly singular iff for all $n \in \mathbb{N}$, $\phi_n(\mu, \cdot) = 1$ μ almost everywhere (which holds iff $\phi(\mu, \cdot) = 1$ μ almost everywhere, as ϕ_n is non-increasing), and is conservative iff $\phi(\mu, \cdot) = 0$ μ almost everywhere.

For $n, k > 0$, define

$$\phi_n^k((\mu, x)) = \mathbb{P}_\mu\left(\hat{x}_1^n = x_1^n | \hat{x}_{n+1}^{n+k}\right)(x)$$

By the Martingale Theorem, for P almost every (μ, x),

$$\phi_n((\mu, x)) = \lim_k \phi_n^k((\mu, x))$$

Note that

$$\phi_n^k \circ M_p^{CP}((\mu, x)) = \phi_n^k((\mu^{x_1}, M_p x)) = \mathbb{P}_{\mu^{x_1}}\left(\hat{x}_1^n = x_2^{n+1} | \hat{x}_{n+1}^{n+k}\right)(M_p x)$$

$$= \frac{\mu^{x_1}(\hat{x}_1^{n+k} = x_2^{n+k+1})}{\mu^{x_1}(\hat{x}_{n+1}^{n+k} = x_{n+2}^{n+k+1})} = \frac{\mu(\hat{x}_1 = x_1, \hat{x}_2^{n+k+1} = x_2^{n+k+1})}{\mu(\hat{x}_1 = x_1, \hat{x}_{n+2}^{n+k+1} = x_{n+2}^{n+k+1})}$$

$$= \frac{\mu(\hat{x}_1^{n+k+1} = x_1^{n+k+1})}{\mu(\hat{x}_{n+2}^{n+k+1} = x_{n+2}^{n+k+1})} \frac{\mu(\hat{x}_{n+2}^{n+k+1} = x_{n+2}^{n+k+1})}{\mu(\hat{x}_1 = x_1, \hat{x}_{n+2}^{n+k+1} = x_{n+2}^{n+k+1})}$$

$$= \phi_{n+1}^k(\mu, x)[\mathbb{P}_\mu\left(\hat{x}_1 = x_1 | \hat{x}_{n+2}^{n+k+1}\right)(x)]^{-1}$$

Taking the limit $k \to \infty$ we see that for almost every (μ, x),

$$\phi_n \circ M_p^{CP}(\mu, x) = \phi_{n+1}(\mu, x)[\mathbb{P}_\mu(\hat{x}_1 = x_1 | \hat{x}_{n+2}^\infty)]^{-1} \tag{5.2}$$

and if we now take n to infinity, we get, for P almost every (μ, x),

$$\phi \circ M_p^{CP}(\mu, x) = \phi(\mu, x)[\mathbb{P}_\mu\left(\hat{x}_1 = x_1 | \cap_{n=3}^\infty \sigma(\hat{x}_n^\infty)\right)(x)]^{-1} \tag{5.3}$$

This implies that $\phi \circ M_p^{CP} \geq \phi$ P almost everywhere, and in a measure preserving ergodic system, this implies that ϕ is an M_p^{CP} invariant function, and therefore there is some $0 \leq \beta \leq 1$ such that $\phi = \beta$ P almost everywhere. Equation 5.3 implies that for P almost every (μ, x),

$$\beta = \beta[\mathbb{P}_\mu\left(\hat{x}_1 = x_1 | \cap_{n=3}^\infty \sigma(\hat{x}_n^\infty)\right)(x)]^{-1}$$

Assume $\beta > 0$, then $\mathbb{P}_\mu\left(\hat{x}_1 = x_1 | \cap_{n=3}^\infty \sigma(\hat{x}_n^\infty)\right)(x) = 1$ P almost everywhere, which means that P is a bilaterally deterministic ECPD. For every $n \geq 2$, the σ-algebra generated by $\hat{x}_n^\infty$ contains $\cap_{n=3}^\infty \sigma(\hat{x}_n^\infty)$, and therefore for P almost every (μ, x),

$$\mathbb{P}_\mu\left(\hat{x}_1 = x_1 | \hat{x}_n^\infty\right)(x) = 1$$

In particular

$$\phi_1(\mu, x) = \mathbb{P}_\mu\left(\hat{x}_1 = x_1 | \hat{x}_2^\infty\right)(x) = 1$$

almost everywhere, and according to Equation 5.2,

$$\phi_{n+1} = \phi_n \circ M_p$$

almost everywhere, and the last two equations imply that for every n, $\phi_n = 1$ almost everywhere, and so $\beta = 1$. Thus we have shown that either $\beta = 0$, and so the D action is conservative with respect to typical measures, or P is a bilaterally deterministic ECPD and $\beta = 1$, and so the D action is strictly singular with respect to typical measures. $\square$

5.3 Extended ECPD Induced by Shift-Invariant Measures

Let $\hat{j}_n : \{0, 1, \ldots, p-1\}^{\mathbb{Z}} \to \{0, 1, \ldots, p-1\}^{\mathbb{Z}}$ be the projection maps

$$\hat{j}_n((i_l)_{l \in \mathbb{Z}}) = i_n$$

we say that a measure μ on $(\{0,1,\ldots,p-1\}^{\mathbb{Z}}, \mathbb{B}^{symbolic})$ is bilaterally deterministic if

$$\mathbb{B}^{symbolic} =^{\mu} \left(\cap_{n\in\mathbb{Z}_-}\sigma(\hat{j}_{-\infty}^{n})\right) \vee \left(\cap_{n\in\mathbb{N}}\sigma(\hat{j}_{n}^{\infty})\right)$$

As we discussed in the Introduction (in Subsection 1.3), Ornstein and Weiss introduced this concept in [11], and showed that any symbolic space endowed with a measure μ which is shift invariant and ergodic, is measure theoretically isomorphic to another symbolic space which is bilaterally deterministic.

We now present the canonical example of an Extended ECPD Q^{ext} induced by, and measure theoretically isomorphic to, a shift-invariant ergodic measure on a symbolic space (this example appears in [3] and [4]), and show that in this example, Q^{ext} is bilaterally deterministic as an ECPD, iff μ is bilaterally deterministic as a shift invariant measure. As we discussed in the introduction, this implies that there are 'many' bilaterally deterministic ECPD, since the result obtained by Ornstein and Weiss implies that every ECPS generated by a shift-invariant ergodic measure will be measure theoretically isomorphic to an ECPS which is bilaterally deterministic. Additionally, Theorem 1.16 implies that the bilateral determinism of a shift invariant ergodic measure μ is the property that determines whether the Extended ECPS generated by this measure will have the 'pointwise conservativity' property. In the language of symbolic spaces, this is equivalent to saying that almost every conditional measure of μ will be conservative with respect to the action of finite coordinate changes, iff μ is bilaterally deterministic.

Let μ be a shift-invariant ergodic measure on $\{0,1,\ldots,p-1\}^{\mathbb{Z}}$. We assume additionally that $\mu \neq \delta_{\overrightarrow{p-1}}, \delta_{\vec{0}}$. For some point $(i_n)_{n\in\mathbb{Z}}$, let us denote the measure obtained from conditioning on all coordinates smaller or equal to n by $\mu^{(i_{-\infty}^{n})}$. We call such measures prediction measures. Any prediction measure is concentrated on the set $\{(j_k)_{k\in\mathbb{Z}} : j_{-\infty}^{n} = i_{-\infty}^{n}\}$ and we can therefore think of it as a measure on $\{0,1,\ldots,p-1\}^{\mathbb{N}}$. Our assumption that $\mu \neq \delta_{\overrightarrow{p-1}}$ implies that E_{p-1}^{+} is a null set with respect to almost every prediction measure, and a measure on $\{0,1,\ldots,p-1\}^{\mathbb{N}}$ which gives E_{p-1}^{+} zero measure can be identified with a measure on $[0,1)$ (via the map $(i_n)_{n\in\mathbb{N}} \mapsto \sum_{n\in\mathbb{N}} i_n p^{-n}$). using this identification of prediction measures $\mu^{(i_{-\infty}^{n})}$ with measures $\tilde{\mu}^{(i_{-\infty}^{n})}$ on $[0,1)$, we can define a map $\chi : \{0,1,\ldots,p-1\}^{\mathbb{Z}} \to (\mathbb{P}_0[0,1) \times \{0,1,\ldots,p-1\})^{\mathbb{Z}}$ by

$$\chi((i_n)_{n\in\mathbb{Z}}) = (\tilde{\mu}^{(i_{-\infty}^{n})}, i_n)_{n\in\mathbb{Z}}$$

Lemma 5.5. *$Q^{chain} = \chi\mu$ is supported on LS, is an adapted distribution, and $(LS, \mathbb{B}, \boldsymbol{\sigma}, \chi\mu)$ is measure theoretically isomorphic to $(\{0,1,\ldots,p-1\}^{\mathbb{Z}}, \mathbb{B}, \boldsymbol{\sigma}, \mu)$ (in particular Q^{chain} is a chain ECPD).*

The proof of this Lemma appears in Appendix B.

Our assumption that $\mu \neq \delta_{\vec{0}}, \delta_{\overrightarrow{p-1}}$ implies that Q^{chain} fulfills the Non-Constant Sequence Condition, and therefore the distribution $Q^{\text{ext}} = \Phi Q^{chain}$ is an Extended

ECPD, and the Extended ECPS $(N\mathbb{M}(\mathbb{R}) \times \{0,1,\dots,p-1\}^{\mathbb{Z}}, \mathbb{B}^{ext}, M_p^{\text{ext}}, Q^{\text{ext}})$ is measure theoretically isomorphic to $(LS, \mathbb{B}^{chain}, \boldsymbol{\sigma}, Q^{chain})$ and thus also to $(\{0,1,\dots,p-1\}^{\mathbb{Z}}, \mathbb{B}^{symbolic}, \boldsymbol{\sigma}, \mu)$. Recall that Q^{ext} is bilaterally deterministic if for every $k_0 \in \mathbb{N}$, and Q^{ext} almost every $(\nu, (i_l)_{l\in\mathbb{Z}})$,

$$\mathbb{P}_{Q^{\text{ext}}}\left(\hat{i}_1 = i_1 | \hat{\nu}, (\hat{i}_n)_{n\leq 0}, (\hat{i}_k)_{k\geq k_0}\right)((\nu, (i_l)_{l\in\mathbb{Z}})) = 1$$

which in our case occurs iff for every $k_0 \in \mathbb{N}$ and Q^{chain} almost every $(\mu_l, i_l)_{l\in\mathbb{Z}}$,

$$\mathbb{P}_{Q^{chain}}\left(\hat{i}_1 = i_1 | (\hat{\mu}_n, \hat{i}_n)_{n\leq 0}, (\hat{i}_k)_{k\geq k_0}\right)((\mu_l, i_l)_{l\in\mathbb{Z}}) = 1$$

Using the shift invariance of Q^{chain}, this can be shown to be equivalent to

$$\mathbb{B}^{chain} =^{Q^{chain}} \left(\cap_{n\in\mathbb{Z}_-}\sigma((\hat{\mu}^n_{-\infty}, \hat{i}^n_{-\infty}))\right) \vee \left(\cap_{n\in\mathbb{N}}\sigma(\hat{i}^{\infty}_n)\right)$$

which in turn is equivalent to bilateral determinism of μ, i.e. to

$$\mathbb{B}^{symbolic} =^{\mu} \left(\cap_{n\in\mathbb{Z}_-}\sigma(\hat{j}^n_{-\infty})\right) \vee \left(\cap_{n\in\mathbb{N}}\sigma(\hat{j}^{\infty}_n)\right)$$

5.4 Example of a 'Strongly Singular' ECPS

Fix $p = 10$. We now describe an ECPD P_{sing} satisfying the singularity properties described earlier (i.e. for P_{sing} almost every μ, the action of $[0,1)$ on $[0,1)$ is strictly singular with respect to μ, and moreover, for every $x \in [0,1) \setminus \{0\}$, $t^{\mathbb{T}}_x\mu(\text{supp}\mu) = 0$). For every n, denote words in $\{L,M,R\}^n$ by $W = (W_1,\dots,W_n)$, and for some $A \in \{L,M,R\}^l$, and $W \in \{L,M,R\}^n$, $W \circ A$ will be the word

$$(W_1,\dots,W_n,A_1,\dots,A_l) \in \{L,M,R\}^{n+l}$$

A measure in our ECPS will be selected by a random recursive process, in which the intervals which the measure will be supported on will be selected.

Stage 1: Select $I_L = [2)_{10}$, and $I_R = [9)_{10}$. I_M will be the interval $[5)_{10}$ with probability $\frac{1}{2}$, and $[6)_{10}$ with probability $\frac{1}{2}$ (independently of all future choices).

Now assume that the intervals I_W have been defined for all $W \in \{L,M,R\}^n$.

Stage $n+1$: For every $W \in \{L,M,R\}^n$, $I_{W\circ L}$ is the intersection of the interval I_W selected in the previous stage and the set $\{x : x_{n+1} = 2\}$, $I_{W\circ R}$ is the interval $I_W \cap \{x : x_{n+1} = 9\}$, and $I_{W\circ M}$ is, with equal probability, either $I_W \cap \{x : x_{n+1} = 5\}$ or $I_W \cap \{x : x_{n+1} = 6\}$.

We now define μ to be the unique measure which for every $n \in \mathbb{N}$, and every $W \in \{L,M,R\}^n$, has $\mu(I_W) = 3^{-n}$. Given μ, a point $x \in [0,1)$ will be selected according to μ, i.e. for every n, x will be with equal probability, in one of the intervals $\{I_W\}_{W\in\{L,M,R\}^n}$.

We give a more formal definition of the ECPS described above, and prove that it is a ECPS, in Appendix B. Our objective is to prove the following

Proposition 5.6. *For every μ selected in the manner described above, and every $x \in [0,1) \setminus D$, $t_x^{\mathbb{T}}\mu(supp\mu) = 0$.*

Proposition 5.7. *For almost every μ, the action of D on $[0,1)$ is strictly singular with respect to μ. Moreover, for every $\alpha \in D \setminus \{0\}$, $t_\alpha^{\mathbb{T}}\mu(supp\mu) = 0$*

Proof of Proposition 5.6. We will actually show that for every $s \in \mathbb{R} \setminus \mathbb{Z}\left[\frac{1}{p}\right]$,

$$t_s\mu(\text{supp}\mu) = 0$$

and therefore using the same reasoning as in the first part of Remark 5.4, for every $s \in [0,1) \setminus D$, $t_s^{\mathbb{T}}\mu(\text{supp}\mu) = 0$.

Fix some μ and $s \in \mathbb{R} \setminus D$. We will show that for every $s \in (0,1]$, $t_s\mu(\text{supp}\mu) = 0$, and the same reasoning can be used to show the same for $s \in [-1,0)$. It is unnecessary to consider $s \notin [-1,1]$ since in that case $t_s\mu[0,1) = 0$.

Define $K_n = K_n(\mu)$ by $K_n = \uplus_{W \in \{L,M,R\}^n} \bar{I}_W$. K_n is a decreasing sequence of compact sets, and $K = \cap K_n$. is the support of μ.

Intervals $I, J \in \mathbb{D}_{10^n}$ (the definition of $\mathbb{D}_{p^n}$ appears in the beginning of Subsection 5.2) will be called overlapping if $I = J$, and we will say that I precedes J if $I + 10^{-n} = J$.

The s we fixed has a unique decimal representation $s = \sum_{i=1}^{\infty} s_i 10^{-i}$. For $j \in \mathbb{N}$, $j \le k \le \infty$ we define $[s]_j^k = \sum_{i=j}^{k} s_i 10^{-i}$.

Let $G_n = G_n(\mu)$ be the set

$$\{W \in \{L,M,R\}^n : \exists Z \in \{L,M,R\}^n \; I_W + [s]_1^n \text{ overlaps or precedes } I_Z\}$$

Note that if $I, J \in \mathbb{D}_{10^n}$ and $I + [s]_1^n$ doesn't overlap or precede J, then necessarily for every $n \le k \le \infty$,

$$I + [s]_1^k \cap J = \emptyset$$

which implies that $\bar{I} + [s]_1^k \cap \bar{J}$ has in it at most one point, and therefore has measure zero as μ is non-atomic. We conclude that

$$\mu(K_n \cap (K_n - s)) \le |G_n| 3^{-n}$$

and that if $W \notin G_n$ then for any $A \in \{L,M,R\}$, $W \circ A \notin G_{n+1}$, and so

$$|G_{n+1}| \le 3|G_n|$$

The fact that $s \notin D$ implies that there is an infinite number of $l \in \mathbb{N}$ such that either $s_{l+1} \notin \{0,9\}$ or $(s_{l+1}, s_{l+2}) = (0,9)$, and therefore

$$9 \cdot 10^{-(l+1)} > [s]_{l+1}^{\infty} \ge [s]_{l+1}^{l+2} \ge 9 \cdot 10^{-(l+2)} \tag{5.4}$$

We pick a subsequence n_k with the above property, and $n_{k+1} - n_k \ge 2$.

Fix an l that fulfills Eq 5.4, and $W \in G_l$. There is a unique $Z \in G_l$ such that $I_W + [s]_1^l$ overlaps or precedes I_Z.

Let us first assume $I_W + [s]_1^l$ overlaps I_Z. Then $I_{WRR} + [s]_1^l + 10^{-(l+2)}$ is 'to the right of I_Z', and since $[s]_{l+1}^{l+2} > 10^{-(l+2)}$, necessarily

$$(I_{WRR} + [s]_1^{l+2}) \cap I_Z = \emptyset$$

On the other hand, $I_{WRR} \subseteq I_W$ cannot intersect any other I_B for $B \in \{L, M, R\}^l$ since I_W does not overlap or precede I_B. Thus $I_{WRR} + [s]_1^{l+2}$ does not intersect any of the l-th generation intervals $\{I_B\}_{B\in\{L,M,R\}^l}$ and therefore it cannot intersect the $l+2$ generation intervals $\{I_B\}_{B\in\{L,M,R\}^{l+2}}$, and so by definition $WRR \notin G_{l+2}$.

Now let us assume that $I_W + [s]_1^l$ precedes I_Z. Then it can be shown similarly that $I_{WLL} + [s]_1^{l+2}$ doesn't intersect any I_B for $B \in \{L, M, R\}^l$ and so $WLL \notin G_{l+2}$. Thus in both cases we obtain

$$|G_{l+2}| \leq 8|G_l|$$

Using this inequality recursively for the sequence n_k described above, we get $|G_{n_k}| \leq 3^{n_k} \left(\frac{8}{9}\right)^k |G_1|$ which implies

$$\mu(K \cap (K - s)) \leq \mu(K_{n_k} \cap (K_{n_k} - s)) \leq 3^{-n_k} 3^{n_k} \left(\frac{8}{9}\right)^k \to 0$$

Thus we have shown that $K = \text{supp}\mu$ satisfies $\mu(K) = 1$ and $t_{-s}\mu(K) = 0$.

$\square$

Proof of Proposition 5.7. We show that for P_{sing} almost every (μ, x), there is a $k \in \mathbb{N}$ such that $\phi_1^k((\mu, x)) = 1$. This occurs iff μ and x_2^{k+1} determine x_1, i.e. if $\mu([x_1, x_2, \ldots, x_{k+1})_{p^{k+1}} > 0$, but for every $j \neq x_1$, $\mu([j, x_2 \ldots, x_{k+1})_{p^{k+1}} = 0$. Therefore, if $\phi_1^k((\mu, x)) = 1$ then necessarily also $\phi_1^{k+1}((\mu, x)) = 1$, and so the fact that typical (μ, x) have $\phi_1^k((\mu, x)) = 1$ for some k, imply that $\phi_1 = \lim \phi_1^k = 1$ almost everywhere, and thus this property is stronger that the strict singularity we want to prove, and in fact (though we do not elaborate on this point) implies that for all $\alpha \in D \setminus \{0\}$, $t_\alpha^{\mathbb{T}}\mu(\text{supp}\mu) = 0$.

For any $a = (a_1, \ldots, a_{k+1}) \in \{2, 5, 6, 9\}^{k+1}$, we estimate

$$\mathbb{P}_{P_{sing}} \left(\{(\mu, x) : \phi_1^k(\mu, x) = 1\} | \{\hat{x}_1^{k+1} = a\} \right)$$

For every n, and $a \in \{2, 5, 6, 9\}^n$, let $r_n(a)$ be the number of a_i such that $a_i \in \{5, 6\}$. Note that

$$\mathbb{P}_{P_{sing}} \left(\{\mu : \mu[a)_{p^n} > 0\} \right)) = 2^{-r_n(a)}$$

For any $a \in \{2, 5, 6, 9\}^{k+1}$, and $i \in \{2, 5, 6, 9\}$, define

$$E_i^a = \{(\mu, x) : x_1^{k+1} = a \text{ and } \mu[i, a_2, \ldots, a_{k+1})_{10^{k+1}} > 0\}$$

Then from the discussion above,

$$\mathbb{P}_{P_{sing}}\left(\{(\mu,x):\phi_1^k(\mu,x)=1\}|\{\hat{x}_1^{k+1}=a\}\right)=\mathbb{P}_{P_{sing}}\left(E_{a_1}^a\setminus\cup_{i\neq a_1}E_i^a|\{\hat{x}_1^{k+1}=a\}\right)\quad(5.5)$$

We note that $\mathbb{P}_{P_{sing}}\left(E_{a_1}^a|\{\hat{x}_1^{k+1}=a\}\right)=1$, since for almost every (μ,x), $\mu[x_1,\dots,x_{k+1})_{10^{k+1}}>0$. For $i\neq a_1$, if $i,a_1\in\{5,6\}$, then $\mathbb{P}_{P_{sing}}\left(E_i^a|\{\hat{x}_1^{k+1}=a\}\right)=\mathbb{P}_{P_{sing}}\left(E_i^a\cap E_{a_1}^a|\{\hat{x}_1^{k+1}=a\}\right)=0$, and otherwise

$$\begin{aligned}\mathbb{P}_{P_{sing}}\left(E_i^a|\{\hat{x}_1^{k+1}=a\}\right)=\mathbb{P}_{P_{sing}}\left(\{\mu:\ \mu[i,a_2,\dots,a_{k+1})_{10^{k+1}}>0\}\right)=2^{-r_{k+1}(i,a_2,\dots,a_{k+1})}\\ \le 2^{-r_{k+1}(a)+1}\end{aligned}$$

and therefore, returning to Equation 5.5 we obtain

$$\mathbb{P}_{P_{sing}}\left(\{(\mu,x):\phi_1^k(\mu,x)=1\}|\{\hat{x}_1^{k+1}=a\}\right)\ge 1-3\cdot 2^{-r_{k+1}(a)+1}$$

For every k, let A_k be the set $A_k=\{(\mu,x):r_{k+1}(x_1^{k+1})>\frac{k+1}{6}\}$, then by the law of large numbers, $P_{sing}(A_k)\to 1$. Therefore,

$$\begin{aligned}\mathbb{P}_{P_{sing}}\left(\{(\mu,x):\ \phi_1^k(\mu,x)=1\}\right)&\ge\sum_{a:\ r_{k+1}(a)\ge\frac{k+1}{6}}\mathbb{P}_{P_{sing}}\left(\phi_1^k=1|\{\hat{x}_1^{k+1}=a\}\right)\mathbb{P}_{P_{sing}}(\{\hat{x}_1^{k+1}=a\})\\ &\ge(1-3\cdot 2^{-\frac{k+1}{6}+1})P_{sing}(A_k)\to 1\end{aligned}$$

□

Appendices

A Proofs for Section 2

In this Section we give the proofs omitted in Section 2. We begin with Subsection A.1 in which we present some facts which will be very useful for most of the proofs of Section 2. We then present the proofs themselves in Subsection A.2.

A.1 Preliminaries

Properties of N In Subsection 1.2 we defined for every $\nu\in\mathbb{M}(\mathbb{R})\setminus\{0\}$

$$\psi(\nu)=\min\{n\in\mathbb{N}:\nu[-(n-1),n)>0\}$$

and then defined a normalization map on $\mathbb{M}(\mathbb{R})$ by

$$N\nu=\begin{cases}\frac{\nu}{\nu[-(\psi(\nu)-1),\psi(\nu))} & \text{if } \nu\neq 0\\ 0 & \text{if } \nu=0\end{cases}$$

This map has the following properties

Remark A.1. *1. For every μ, there is a $\lambda(\mu) > 0$ such that $N\mu = \lambda\mu$. Moreover, if*

$$\mu_1\big|_{[-\psi(\mu_1)+1,\psi(\mu))} = \mu_2\big|_{[-\psi(\mu_1)+1,\psi(\mu_1))}$$

then $\lambda(\mu_1) = \lambda(\mu_2)$.

2. For every $\lambda > 0$, $N(\lambda\mu) = N\mu$. If $\rho : \mathbb{R} \to \mathbb{R}$ is a measurable function then (for $\lambda = \lambda(\mu)$)

$$N\rho N\mu = N\rho\lambda\mu = N\lambda\rho\mu = N\rho\mu$$

In the following we will discuss sequences of measures $(\mu_n)_{n>0}$, with the property that there is an increasing sequence of intervals of the form $I_n = [a_n, b_n)$ with $a_n, b_n \in \mathbb{Z}$ such that μ_n is supported in I_n (in fact we will discuss sequences with non-positive indexes, i.e. $(\mu_n)_{n\le 0}$, $(I_n)_{n\le 0}$ etc. but of course this makes no difference). The following Lemma will be useful

Lemma A.2. *In the setup described above,*

1. *If for every n, for every $k > 0$ there is a $\eta_k > 0$ such that $\eta_k\mu_n\big|_{I_n} = \mu_{n+k}\big|_{I_n}$, then for all large enough n, $N\mu_n\big|_{I_n} = N\mu_{n+k}\big|_{I_n}$.*
2. *If for all large enough n, for every $k > 0$, $\mu_n\big|_{I_n} = \mu_{n+k}\big|_{I_n}$, then the sequence $\lambda(\mu_n)$ is constant for all large enough n. Additionally, for every A, $\lim \mu_n(A)$ is well defined and defines a measure $\lim \mu_n$ with the property that*

$$\lim N\mu_n = N \lim \mu_n$$

3. *If ρ is an invertible continuous function whose inverse is also continuous, then $\lim \rho\mu_n = \rho \lim \mu_n$ (this is in fact correct for every sequence μ_n which converges in the weak topology).*

Proof. 1. If for every n, $\mu_n = 0$, then the claim is trivial. Otherwise, there is an $n \in \mathbb{N}$ such that $\tilde{\mu}_n(I_n) > 0$, and therefore there is a minimal $m = \psi(\mu) \in \mathbb{N}$ with $\tilde{\mu}_n[-m+1, m) > 0$. There is now a large enough n_0 with the property that

$$[-m+1, m) \cap (\cup_{l>0} I_l) \subseteq I_{n_0} \tag{A.1}$$

which implies that for every $n \ge n_0$, $k > 0$, if $\eta_k\mu_n$ and μ_{n+k} agree on I_n, then they will also agree on $[-m+1, m)$. It follows from Remark A.1 that $\lambda(\eta_k\mu_n) = \lambda(\mu_{n+k})$ and

$$N(\mu_n) = N(\eta_k\mu_n) = \eta_k\lambda(\mu_{n+k})\mu_n$$

and therefore

$$N(\mu_n)\big|_{I_n} = \eta_k\lambda(\mu_{n+k})\mu_n\big|_{I_n} = \lambda(\mu_{n+k})\mu_{n+k}\big|_{I_n} = (N\mu_{n+k})\big|_{I_n}$$

2. If for every n, $\mu_n = 0$, then the claim is trivial. Otherwise, There is an $n_0 \in \mathbb{N}$ for which Equation A.1 holds, and for every $n > n_0$, $\lambda(\mu_n) = \lambda(\mu_{n_0})$. We note that also $\nu = \lim \mu_n$ agrees with μ_{n_0} on I_{n_0} and therefore also $\lambda(\nu) = \lambda(\mu_{n_0})$, and so
$$\lim N\mu_n = \lim \lambda(\mu_{n_0})\mu_n = \lambda(\mu_{n_0})\nu = N\nu$$

3. Assume $\mu_n \to \mu$ (in the weak topology). Let f be a continuous function supported on a compact set K. Then $f \circ \rho$ is a continuous function, supported in $\rho^{-1}K$ which is a compact set since we assume ρ^{-1} is a continuous function, and therefore $\int f d\rho\mu_n = \int (f \circ \rho) d\mu_n \to \int (f \circ \rho) d\mu = \int f d\rho\mu$.

□

Properties of $\mathcal{H}$ We recall that we denoted the set of orientation preserving homotheties by $\mathcal{H}$, and the intervals of the form $[a, b)$ by $\mathcal{I}$, and defined ρ_I^J to be the unique orientation preserving homothety which takes the interval $I \in \mathcal{I}$ to the interval $J \in \mathcal{I}$.

For any $m \in \mathbb{Z} \cup \{\infty\}$, we said that $(I_n)_{n<m}$ is compatible with $(i_n)_{n<m}$ if for every $n < m$, $I_n = I_{n-1}^{i_n}$. We record the following properties of $\mathcal{H}$:

1. For $I, J, K \in \mathcal{I}$, $\rho_J^K \rho_I^J = \rho_I^K$, since these are both homotheties in $\mathcal{H}$ which take I to K, and therefore uniqueness implies that they are equal.

2. for $\rho \in \mathcal{H}$, $I \in \mathcal{I}$ and $i \in \{0, 1, \ldots, p-1\}$, $\rho(I^i) = (\rho(I))^i$. It follows that $\rho_I^J = \rho_{I^i}^{J^i}$, and that for $m \in \mathbb{Z} \cup \{\infty\}$, if $(I_n)_{n<m}$ is compatible with $(i_n)_{n<m}$, then so is $(\rho I_n)_{n<m}$.

A.2 Proofs

We now present the proofs of the Lemmas from Section 2. As we mentioned before, the proofs will often use the observations presented in the previous Subsection, and we will not necessarily give a reference each time one of these observations is used.

Recall that we defined for every $(\mu_n, i_n)_{n\in\mathbb{Z}} \in LS$ measures $\tilde{\mu}_n$ obtained by taking the sequence of intervals $(I_n)_{n\leq 0}$ which is well ordered and compatible with $(i_n)_{n\leq 0}$, and defining $\tilde{\mu}_n = N\rho_{I_0}^{I_n}\mu_n$.

Lemma 2.4. *For every $k \leq n < 0$, there is a $\lambda = \lambda(k) > 0$ such that*
$$\tilde{\mu}_{k-1}\big|_{I_n} = \lambda\tilde{\mu}_k\big|_{I_n}$$
Moreover there is an n_0 such that for all $k < n_0$, $\lambda(k) = 1$.

Proof. It is sufficient to prove the Lemma for $k = n$, and the case where $k < n$ follows immediately from the fact that $I_n \subseteq I_k$.

Note that since $\mu_n = \mu_{n-1}^{i_n}$, μ_n and $\rho_{I_0^{i_n}}^{I_0}\mu_{n-1}$ agree on $I_0 = [0, 1)$ up to a multiplicative constant, and therefore if we pushforward both measures by $\rho_{I_0}^{I_n}$, they will agree on I_n up to a multiplicative constant. Note that $\rho_{I_0}^{I_n}\mu_n$ is, up to normalization, equal to $\tilde{\mu}_n$, and since

$$\rho_{I_0}^{I_n}\rho_{I_0^{i_n}}^{I_0} = \rho_{I_0^{i_n}}^{I_n} = \rho_{I_0^{i_n}}^{I_{n-1}^{i_n}} = \rho_{I_0}^{I_{n-1}}$$

it follows that $\rho_{I_0}^{I_n}\rho_{I_0^{i_n}}^{I_0}\mu_{n-1} = \rho_{I_0}^{I_{n-1}}\mu_{n-1}$, which is equal to $\tilde{\mu}_{n-1}$ up to normalization. Thus we proved that $\tilde{\mu}_{n-1}$ and $\tilde{\mu}_n$ agree on I_n up to a multiplicative constant.

Finally, the fact that for negative enough n, $\lambda(n) = 1$, follows immediately from Lemma A.2.

□

The last Lemma enabled us to define a measure $\nu = \nu((\mu_n, i_n)_{n\leq 0})$ by $\nu(A) = \lim_{n\to-\infty}\tilde{\mu}_n(A)$ and $\Phi : LS \to N\mathbb{M}(\mathbb{R}) \times \{0, 1, \dots, p-1\}^{\mathbb{Z}}$ by

$$\Phi((\mu_n, i_n)_{n\in\mathbb{Z}}) = (\nu((\mu_n, i_n)_{n\leq 0}), (i_n)_{n\in\mathbb{Z}})$$

We then claimed

Lemma 2.6. $\Phi\boldsymbol{\sigma} = M_p^{ext}\Phi$

Recall that we defined M_p^{ext} on $N\mathbb{M}(\mathbb{R}) \times \{0, 1, \dots, p-1\}^{\mathbb{Z}}$ by

$$M_p^{\text{ext}}(\nu, (i_n)_{n\in\mathbb{Z}}) = (N\rho_{i_1}\nu, \boldsymbol{\sigma}((i_n)_{n\in\mathbb{Z}}))$$

where in our terminology here, if $(I_n)_{n\leq 1}$ is the sequence of intervals well based and compatible with $(i_n)_{n\leq 1}$, then $\rho_{i_1} = \rho_{I_1}^{I_0}$.

Proof. For every $(\mu_n, i_n)_{n\in\mathbb{Z}} \in LS$,

$$\Phi \circ \boldsymbol{\sigma}((\mu_n, i_n)_{n\in\mathbb{Z}}) = \Phi((\mu_{n+1}, i_{n+1})_{n\in\mathbb{Z}}) = (\nu((\mu_{n+1}, i_{n+1})_{n\leq 0}), (i_{n+1})_{n\in\mathbb{Z}})$$

while

$$\begin{aligned} M_p^{\text{ext}} \circ \Phi((\mu_n, i_n)_{n\in\mathbb{Z}}) &= M_p^{\text{ext}}(\nu((\mu_n, i_n)_{n\in\mathbb{Z}}), (i_n)_{n\in\mathbb{Z}}) \\ &= (N\rho_{i_1}\nu((\mu_n, i_n)_{n\leq 0}), (i_{n+1})_{n\in\mathbb{Z}})) \end{aligned}$$

therefore, we need to prove that

$$\nu((\mu_{n+1}, i_{n+1})_{n\leq 0}) = N\rho_{i_1}\nu((\mu_n, i_n)_{n\leq 0})$$

Let $(I_n)_{n\leq 1}$ be the sequence of intervals which is well based and compatible with $(i_n)_{n\leq 1}$. It follows that $(I_{n+1})_{n\leq 0}$ is still compatible with $(i_{n+1})_{n\leq 0}$. As homotheties preserve compatibility, $(J_n)_{n\leq 0} = (\rho_{I_1}^{I_0}I_{n+1})_{n\leq 0}$ is also compatible with $(i_{n+1})_{n\leq 0}$, and it is also well based since $\rho_{I_1}^{I_0}I_1 = I_0$. Therefore,

$$\nu((\mu_{n+1}, i_{n+1})_{n<0}) = \lim_{n\to-\infty} N\rho_{J_0}^{J_n}\mu_{n+1} \tag{A.2}$$

Now, as $J_0 = [0,1) = I_0$, and the homothety in $\mathcal{H}$ which takes I_{n+1} to $J_n = \rho_{I_1}^{I_0} I_{n+1}$ is just $\rho_{I_1}^{I_0}$ by definition,

$$\rho_{J_0}^{J_n} = \rho_{I_0}^{J_n} = \rho_{I_{n+1}}^{J_n}\rho_{I_0}^{I_{n+1}} = \rho_{I_1}^{I_0}\rho_{I_0}^{I_{n+1}}$$

and so, returning to Eq A.2 we obtain

$$\nu((\mu_{n+1}, i_{n+1})_{n<0}) = \lim_{n\to-\infty} N\rho_{I_1}^{I_0}\rho_{I_0}^{I_{n+1}}\mu_{n+1} = \lim_{n\to-\infty} N\rho_{I_1}^{I_0} N\rho_{I_0}^{I_{n+1}}\mu_{n+1}$$

$$= \lim_{n\to-\infty} N\rho_{I_1}^{I_0}\tilde{\mu}_{n+1} = N \lim_{n\to-\infty} \rho_{I_1}^{I_0}\tilde{\mu}_{n+1}$$

$$= N\rho_{I_1}^{I_0} \lim_{n\to-\infty} \tilde{\mu}_{n+1} = N\rho_{I_1}^{I_0}\nu((\mu_n, i_n)_{n\le 0})$$

□

Lemma 2.7. *The restriction of Φ to the ($\boldsymbol{\sigma}$-invariant) set*

$$\{(\mu_n, i_n)_{n\in\mathbb{Z}} \in LS :\ (i_n)_{n\in\mathbb{Z}} \notin E_0 \cup E_{p-1}\}$$

is a measurable bijection onto (the M_p^{ext} invariant set)

$$\{(\nu, (i_n)_{n\in\mathbb{Z}}) :\ (i_n)_{n\in\mathbb{Z}} \notin E_0 \cup E_{p-1}\}$$

Proof. Using the tools provided in Subsection A.1 in the same manner as we did above, it can be shown that

$$\Phi^{-1}(\nu, (i_n)_{n\in\mathbb{Z}}) = (N\rho_{I_n}^{I_0}(\nu\big|_{I_n}), i_n)_{n\in\mathbb{Z}}$$

where $(I_n)_{n\le 0}$ is the sequence which is well based and compatible with $(i_n)_{n\le 0}$. □

We recall that for $k \in \mathbb{Z}$, we defined $T_k : \tilde{X} \to \tilde{X}$ by

$$T_k(\nu, (i_n)_{n\le 0}) = (Nt_k\nu, s_k((i_n)_{n\le 0}))$$

where $s_k(i_n)$ is the unique sequence with the property that, for $(I_n)_{n\le 0}$ and $(J_n)_{n\le 0}$, the sequences of intervals which are well based and compatible with $(i_n)_{n\le 0}$ and $s_k((i_n)_{n\le 0})$ respectively, for all negative enough n, $J_n = I_n - k$.

Lemma 2.8. *The maps $\{T_k\}_{k\in\mathbb{Z}}$ define an action of $\mathbb{Z}$ on $\tilde{X}$.*

Proof. We need to prove that for all $l, k \in \mathbb{Z}$, $Nt_lNt_k = Nt_{l+k}$ and $s_ls_k = s_{l+k}$. Since for all $k, l \in \mathbb{Z}$, $t_k, t_l \in \mathcal{H}$, it follows that for every $\nu \in N\mathbb{M}(\mathbb{R})$,

$$Nt_lNt_k\nu = Nt_lt_k\nu = Nt_{l+k}\nu$$

If $(I_n)_{n\le 0}$, $(J_n)_{n\le 0}$ and $(F_n)_{n\le 0}$ are well based and compatible with $(i_n)_{n\le 0}$, $s_k((i_n)_{n\le 0})$ and $s_l(s_k((i_n)_{n\le 0}))$ respectively, then for all negative enough n, $J_n = I_n - k$ and $F_n = J_n - l$ and therefore for all negative enough n, $F_n = I_n - (l + k)$. Therefore, uniqueness implies that

$$s_{l+k}((i_n)_{n\le 0}) = s_l(s_k((i_n)_{n\le 0}))$$

□

We now recall that we defined a group

$$G = \{(i_n)_{n\leq 0} :\ i_n = 0 \text{ for all negative enough } n\}$$

and we defined 'translation maps' $\{S_a\}_{a\in G}$ on X by $S_a((\mu_n, i_n)_{n\leq 0}) = (\nu_n, j_n)$ where $(j_n)_{n\leq 0} = (i_n)_{n\leq 0} + a$ and ν_n is the unique sequence with $\mu_n = \nu_n$ for all negative enough n, and additionally $(\nu_n, j_n)_{n\leq 0} \in X$.

Lemma 2.9. *For every $(\mu_n, i_n)_{n\leq 0} \in X$, the following holds*

1. *For every $k \in \mathbb{Z}$, there is an $a \in G$ such that*

$$\theta S_a((\mu_n, i_n)_{n\leq 0}) = T_k\theta((\mu_n, i_n)_{n\leq 0}) \tag{2.4}$$

2. *For every $a \in G$, there is a $k \in \mathbb{Z}$ such that Equation 2.4 holds.*

3. *Assume that for $a \in G$, $k \in \mathbb{Z}$, Equation 2.4 holds. Then $a = 0$ iff $k = 0$.*

Proof.

1. For given $k \in \mathbb{Z}$ and $(\mu_n, i_n)_{n\leq 0}$, as $(i_n)_{n\leq 0}$ and $s_k((i_n)_{n\leq 0})$ disagree on a finite number of coordinates, there is some $a \in G$ such that $s_k((i_n)_{n\leq 0}) = (i_n)_{n\leq 0} + a$. We claim that for this a, equality holds in the measure coordinate as well, i.e.

$$\theta S_a((\mu_n, i_n)_{n\leq 0}) = T_k\theta((\mu_n, i_n)_{n\leq 0}) = (Nt_k\nu((\mu_n, i_n)_{n\leq 0}), s_k((i_n)_{n\leq 0}))$$

Let $(I_n)_{n\leq 0}$ and $(J_n)_{n\leq 0}$ be the sequences of intervals which are well based and compatible with $(i_n)_{n\leq 0}$ and $s_k((i_n)_{n\leq 0}) = (i_n)_{n\leq 0} + a$, then

$$Nt_k\nu((\mu_n, i_n)_{n\leq 0}) = Nt_k(\lim_{n\to-\infty} N\rho_{I_0}^{I_n}\mu_n) = \lim_{n\to-\infty} Nt_kN\rho_{I_0}^{I_n}\mu_n$$

$$= \lim_{n\to-\infty} Nt_k\rho_{I_0}^{I_n}\mu_n = \lim_{n\to-\infty} N\rho_{I_0}^{t_kI_n}\mu_n =^{(*)} \nu(S_a((\mu_l, i_l)_{l\leq 0}))$$

where $(*)$ follows from the fact that for all negative enough n, $J_n = t_kI_n$ and $\hat{\mu}_nS_a((\mu_l, i_l)_{l\leq 0}) = \mu_n$.

2. For given $a \in G$ and $(\mu_n, i_n)_{n\leq 0} \in X$, it is sufficient to find a $k \in \mathbb{Z}$ such that $s_k((i_n)_{n\leq 0}) = (i_n)_{n\leq 0} + a$ and therefore by the proof of the former claim, there is a $b \in G$ such that $\theta S_b((\mu_n, i_n)_{n\leq 0}) = T_k\theta((\mu_n, i_n)_{n\leq 0})$ and in particular $(i_n)_{n\leq 0} + a = (i_n)_{n\leq 0} + b$ and therefore $a = b$.

   To find such a $k \in \mathbb{Z}$, let $(I_l)_{l\leq 0}$ be the sequence well based and compatible with $(i_l)_{l\leq 0}$, and define

$$(j_l)_{l\leq 0} = (i_l)_{l\leq 0} + a$$

   There is some $n \leq 0$ such that for all $m < n$, $i_m = j_m$. Define a sequence $(J_n)_{n\leq 0}$, by requiring that for all $m < n$, $J_m = I_m$, and by recursively requiring that for all $m \geq n$,

$$J_m = J_{m-1}^{j_m}$$

Then $(J_l)_{l\leq 0}$ is compatible with $(j_l)_{l\leq 0}$ but not well based. However, there is a $k\in\mathbb{Z}$ such that $t_kJ_0 = I_0$, and as homotheties preserve compatibility, $(t_kJ_l)_{l\leq 0}$ is well based and compatible with $(j_l)_{l\leq 0}$. Since for all negative enough l, $t_kJ_l = t_kI_l$, $(j_l)_{l\leq 0}$ fulfills the property that determines $s_k((i_l)_{l\leq 0})$ uniquely, and therefore

$$s_k((i_l)_{l\leq 0}) = (j_l)_{l\leq 0} = (i_l)_{l\leq 0} + a$$

3. This just follows from the fact that $(i_l)_{l\leq 0} + a = (i_l)_{l\leq 0}$ iff $a = 0$, and similarly $s_k((i_l)_{l\leq 0}) = (i_l)_{l\leq 0}$ iff $k = 0$.

□

The following can now be easily shown

Lemma 2.10. *The action of G on X is conservative with respect to a probability distribution Q on X, iff the action of $\mathbb{Z}$ on $\tilde{X}$ is conservative with respect to θQ.*

Proof. Assume that the action of $\mathbb{Z}$ on $\tilde{X}$ is conservative with respect to θQ. Let $A\subseteq X$ be a Borel set, with $Q(A) > 0$, we want to prove that $Q(A\cap(\cup_{a\in G\setminus\{0\}}S_aA)) > 0$.

$$Q(A\cap(\cup_{a\in G\setminus\{0\}}S_aA)) = \theta Q(\theta A\cap(\cup_{a\in G\setminus\{0\}}\theta S_aA)) = \theta Q(\theta A\cap(\cup_{k\in\mathbb{Z}\setminus\{0\}}T_k\theta A)) > 0$$

Where $\theta A\cap(\cup_{k\in\mathbb{Z}\setminus\{0\}}T_k\theta A)$ has positive measure since we assumed the action of $\mathbb{Z}$ on $\tilde{X}$ is conservative with respect to θQ. The other direction can be prove in exactly the same way. □

Lemma 2.11. *If the action of $\mathbb{Z}$ on $\tilde{X}$ is conservative with respect to $\tilde{Q}$, then the action of $\mathbb{Z}$ on $N\mathbb{M}(\mathbb{R})$ is conservative with respect to $\tilde{\pi}_{\mathbb{M}}\tilde{Q}$.*

Proof. If $A\subseteq N\mathbb{M}(\mathbb{R})$ is a Borel set with $\tilde{\pi}_{\mathbb{M}}\tilde{Q}(A) > 0$, then

$$\tilde{Q}(\{(\nu,(i_n)_{n\leq 0})\in\tilde{X} :\ \nu\in A\}) = \tilde{\pi}_{\mathbb{M}}\tilde{Q}(A) > 0$$

and therefore the conservativity in the assumption implies that there is some $k\neq 0$ such that

$$0 < \tilde{Q}(\{(\nu,(i_n)_{n\leq 0})\in\tilde{X} :\ \nu\in A\}\cap T_k\{(\nu,(i_n)_{n\leq 0})\in\tilde{X} :\ \nu\in A\}) =$$

$$\tilde{\pi}_{\mathbb{M}}\tilde{Q}(A\cap t_k^*A)$$

□

B Formal Description of the ECPS from Subsection 5.3 and Subsection 5.4

We begin this Appendix by proving that the Chain ECPD which arises from a shift invariant ergodic measure on a symbolic space, as described in Subsection 5.3, is indeed a Chain ECPD. We then give a more formal description of the ECPD described in Subsection 5.4 and prove that it is indeed an ECPD.

Lemma 5.5. *$Q^{chain} = \chi\mu$ is supported on LS, is an adapted distribution, and $(LS, \mathbb{B}, \boldsymbol{\sigma}, \chi\mu)$ is measure theoretically isomorphic to $(\{0,1,\dots,p-1\}^{\mathbb{Z}}, \mathbb{B}, \boldsymbol{\sigma}, \mu)$ (in particular Q^{chain} is a chain ECPD).*

Proof. Before we give the proof, let us denote the map $(i_k)_{k\in\mathbb{N}} \mapsto \sum_{k\in\mathbb{N}} i_k p^{-k}$ by π. Note that Q^{chain} almost every μ gives $\{0,1,\dots,p-1\}^{\mathbb{N}} \setminus E^+_{p-1}$ full measure, and therefore we think of π as defined on this set only. The advantage of this is that π is now a bijection. For every $n \in \mathbb{Z}$ let us denote by f_n the maps which identify $\{(j_k)_{k\in\mathbb{Z}} : \ j^n_{-\infty} = i^n_{-\infty}\}$ with $\{0,1,\dots,p-1\}^{\mathbb{N}}$. We can now give a formal definition for the identification of prediction measures with measures on $[0,1)$, by $\tilde{\mu}^{(i^n_{-\infty})} = \pi f_n \mu^{(i^n_{-\infty})}$.

The main part of the proof is to show that Q^{chain} is supported on LS, or in other words, to show that $(\tilde{\mu}^{(i^n_{-\infty})})^{i_{n+1}} = (\tilde{\mu}^{(i^{n+1}_{-\infty})})$. It is sufficient to show that for every $r \in \mathbb{N}$ and $a \in \{0,1,\dots,p-1\}^r$, the measures are equal on the set $\{x : \hat{x}^r_1 = a\}$, since the collection of finite disjoint unions of such sets are an algebra that generates the Borel σ-algebra.

Note that for every $\mu \in \mathbb{P}_0[0,1)$ and $i \in \{0,1,\dots,p-1\}$,

$$\mu^i\left(\{x : \hat{x}^r_1 = a\}\right) = \mathbb{P}_\mu\left(\hat{x}^{r+1}_2 = a | \hat{x}_1 = i\right) \tag{B.1}$$

Additionally, note that since π is not defined on E^+_{p-1},

$$\pi^{-1}\left(\{x : \hat{x}^r_1 = a\}\right) = \{(i_k)_{k\in\mathbb{N}} : \ i^r_1 = a\}$$

and therefore

$$f_n^{-1}\pi^{-1}\left(\{x : \hat{x}^r_1 = a\}\right) = \{(j_k)_{k\in\mathbb{Z}} : \ j^n_{-\infty} = i^n_{-\infty}, j^{n+r}_{n+1} = a\} \tag{B.2}$$

Thus, using Eq. B.1 and Eq. B.2, we obtain

$$\begin{aligned}(\tilde{\mu}^{(i^n_{-\infty})})^{i_{n+1}}\left(\{x : \hat{x}^r_1 = a\}\right) &=^{Eq.\ B.1} \mathbb{P}_{\tilde{\mu}^{(i^n_{-\infty})}}\left(\hat{x}^{r+1}_2 = a | \hat{x}_1 = i\right) \\ &=^{Eq.\ B.2} \mathbb{P}_{\mu^{(i^n_{-\infty})}}\left(\hat{j}^{n+r+1}_{n+2} = a | \hat{j}^n_{-\infty} = i^n_{-\infty}, \hat{j}_{n+1} = i_{n+1}\right) \\ &= \mathbb{P}_\mu\left(\hat{j}^{n+r+1}_{n+2} = a | \hat{j}^n_{-\infty} = i^n_{-\infty}, \hat{j}_{n+1} = i_{n+1}\right) \\ &=^{Eq.\ B.2} \tilde{\mu}^{(i^{n+1}_{-\infty})}\left(\{x : \hat{x}^r_1 = a\}\right)\end{aligned}$$

Thus we have shown that Q^{chain} is supported on LS. It is now immediate to check that χ is a bijective factor map. Finally, Q^{chain} is adapted since for every $j \in \{0,1,\dots,p-1\}$ and Q^{chain} almost every $(\mu_n, i_n)_{n\in\mathbb{Z}}$,

$$\begin{aligned}\mathbb{P}_{Q^{chain}}\left(\hat{i}_1 = j | (\hat{\mu_n}, \hat{i}_n)_{n\le 0}\right)((\mu_n, i_n)_{n\in\mathbb{Z}}) &= \mathbb{P}_\mu\left(\hat{j}_1 = j | (\hat{j}_n)_{n\le 0}\right)((i_n)_{n\in\mathbb{Z}}) \\ &= \mu^{(i^0_{-\infty})}(\{(j_n)_{n\in\mathbb{Z}} : \hat{j}_1 = j\}) = \tilde{\mu}^{(i^0_{-\infty})}[j)_p = \mu_o[j)_p\end{aligned}$$

□

We now show a more formal way to define the ECPS from Subsection 5.4 as a factor of a product of symbolic spaces, and show that it is indeed an ECPS.

Consider the countable set

$$\{L,M,R\}^* = \{\emptyset\} \cup (\cup_{n\in\mathbb{N}}\{L,M,R\}^n)$$

For $(x,y) \in \{5,6\}^{\{L,M,R\}^*} \times \{L,M,R\}^{\mathbb{N}}$ and $W = (W_1, W_2, \ldots, W_d) \in \{L,M,R\}^d$ we define the cylinder sets $I_W(x) = [a_1,\ldots,a_d)_{10^d}$, where

$$a_i(x, W_1^{i-1}) = \begin{cases} 2 & \text{if } W_i = L \\ x_{W_1^{i-1}} & \text{if } W_i = M \\ 9 & \text{if } W_i = R \end{cases}$$

Thus the intervals I_W are chosen with the same distribution as in Subsection 5.4. We now take μ_x to be the unique measure that, for every $W \in \{L,M,R\}^d$, has $\mu(I_W) = 3^{-d}$. Finally, we pick $z(x,y) \in [0,1)$ by requiring that for every n, z will be in the interval $I_W(x)$, where $W = \{y_1,\ldots,y_n\}$. Thus z will be in the interval $I_W(x)$ with probability 3^{-n}, which is exactly the probability that μ_x gives that interval. π is now defined via

$$\pi(x,y) = (\mu_x, z(x,y))$$

For all $x \in \{5,6\}^{\{L,M,R\}^*}$ and $A \in \{L,M,R\}$, define x^A by setting $x^A_W = x_{A\circ W}$ for every $W \in \{L,M,R\}^*$.

We now define a (continuous) map $S : \{5,6\}^{\{L,M,R\}^*} \times \{L,M,R\}^{\mathbb{N}} \to \{5,6\}^{\{L,M,R\}^*} \times \{L,M,R\}^{\mathbb{N}}$ by

$$S(x,y) = (x^{y_1}, \boldsymbol{\sigma}(y))$$

and a measure λ on $\{5,6\}^{\{L,M,R\}^*} \times \{L,M,R\}^{\mathbb{N}}$ which is just the product of the uniform measures on the symbolic spaces $\{5,6\}^{\{L,M,R\}^*}$ and $\{L,M,R\}^{\mathbb{N}}$.

Lemma B.1. *1.* $\pi S = M_{10}^{CP}\pi$.

2. $P_{sing} = \pi\lambda$ *is adapted, and invariant and ergodic with respect to* M^s_{10}.

Proof. 1. We need to prove that for λ almost every (x,y),

$$(\mu_{x^{y_1}}, z(x^{y_1}, \boldsymbol{\sigma} y)) = \pi S(x,y) = M_{10}^{CP}\pi(x,y) = (\mu_x^{z_1(x,y)}, M_{10} z(x,y))$$

It follows from the definition of $a_n(x, W_1^n)$ that

$$a_n(x^{W_1}, W_2^{n+1}) = a_{n+1}(x, W_1^{n+1})$$

Which implies that for every $n \in \mathbb{N}$,

$$z_n(x^{y_1}, \boldsymbol{\sigma} y) = a_n(x^{y_1}, (\boldsymbol{\sigma} y)_1^n) = a_{n+1}(x, y_1^{n+1}) = z_{n+1}((x,y)) = (M_{10} z(x,y))_n$$

Additionally, $\mu_{x^{y_1}}$ and $\mu_x^{z_1(x,y)}$ give mass 3^{-n} precisely to 3^n n-th generation intervals of the form $(W \in \{L,M,R\}^n)$

$$I_W(x^{y_1}) = [a_1^n(x^{y_1}, W))_{10^n} = [a_2^{n+1}(x, y_1 \circ W))_{10^n}$$

and therefore they are equal as measures.

2. We will first prove the adaptivity of P_{sing}, and then prove that λ is invariant and ergodic with respect to S and therefore P_{sing} is invariant and ergodic with respect to M^s_{10}.

   Adaptiveness As we explained previously, for every $j \in \{0, 1, \dots, 9\}$,

$$\mathbb{P}_{P_{sing}}(\hat{x}_1 = j|\hat{\mu}) = \mu[j)_{10}$$

   λ **is invariant** We can rewrite S as

$$S(x, y) = (S_y x, \boldsymbol{\sigma} y)$$

   Where $S_y x = x^{y_1}$. As there is only a finite number of S_y it is not difficult to show that the function

$$(x, y) \mapsto S_y x$$

   is measurable. Additionally, $\boldsymbol{\sigma}\lambda_2 = \lambda_2$ and for every y, it can be verified that $S_y \lambda_1 = \lambda_1$, and therefore $(X \times Y, \mathbb{B}_1 \times \mathbb{B}_2, S, \lambda)$ is a skew system, and S is measure preserving (see page 11 in [10] for the definition of skew product we use here).

   λ **is mixing** We give an outline of the proof. Let $\mathcal{A}$ be the collection of cylinder sets in $\mathbb{B}_2$ multiplied by X, and $\mathcal{C}$ be the collection of cylinder sets in $\mathbb{B}_1$ multiplied by Y. It is sufficient to show that for every $A_1, A_2 \in \mathcal{A}$ and $B_1, B_2 \in \mathcal{C}$, for large enough $n \in \mathbb{N}$, $A_1 \cap B_1$ and $S^{-n}(A_2 \cap B_2)$ are independent, since this implies that the same is true for disjoint unions of such intersections, which is an algebra which generates $\mathbb{B}_1 \times \mathbb{B}_2$, and therefore $(X \times Y, \mathbb{B}_1 \times \mathbb{B}_2, S, \lambda)$ is mixing.

   Pick $A_1, A_2 \in \mathcal{A}$ and $B_1, B_2 \in \mathcal{C}$, it can be shown that

   (a) For large enough n, $S^{-n}A_2$ is independent of $A_1 \cap B_1$.

   (b) There is an n_0, such that for every $n > n_0$, and every $A \in \mathcal{A}$, $S^{-n}B_2$ and $A \cap B_1$ are independent.

   Since for every $n \in \mathbb{N}$ and $A \in \mathcal{A}$, $S^{-n}A \in \mathcal{A}$, it follows that for $n > n_0$,

$$\lambda((A_1 \cap B_1) \cap S^{-n}(A_2 \cap B_2)) = \lambda(((A_1 \cap S^{-n}A_2) \cap B_1) \cap S^{-n}B_2)$$

$$= \lambda(A_1 \cap S^{-n}A_2 \cap B_1)\lambda(S^{-n}B_2)$$

   which for large enough n is equal to

$$\lambda(A_1 \cap B_1)\lambda(S^{-n}(A_2))\lambda(S^{-n}(B_2)) = \lambda(A_1 \cap B_1)\lambda(S^{-n}(A_2 \cap B_2))$$

   .

□

References

[1] Jon Aaronson. *An introduction to infinite ergodic theory.* Volume 50 of Mathematical Surveys and Monographs, A.M.S., Providence, RI, 1997.

[2] Harry Furstenberg. *Intersections of Cantor sets and transversality of semigroups.* Problems in Analysis, Princeton Mathematical Series, 31. Ed. R. C. Gunning. Princeton University Press, Princeton, NJ, 41-59, 1970.

[3] Hillel Furstenberg & Benjamin Weiss. *Markov processes and Ramsey theory for trees.* Special Issue on Ramsey Theory. Combinatorics, Probability and Computing, 12, 548-563, 2003.

[4] Hillel Furstenberg. *Ergodic fractal measures and dimension conservation.* Ergodic Theory and Dynamical Systems, 28(2),405-422,2008.

[5] Michael Hochman & Pablo Shmerkin. *Entropy averages and projections of fractal measures.* Annals of Mathematics. 175, 1001-1059, 2012.

[6] Michael Hochman. *Dynamics on fractals and fractal distributions.* In preperation.

[7] Bernard Host. *Nombres normaux, entropie, translations.* Israel Journal of Mathematics, 91(1-3),419-428,1995.

[8] Pertti Mattila, *Geometry of sets and measures in Euclidean space.* volume 44 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1995. Fractals and rectifiability.

[9] Konstantin Medynets & Boris Solomyak. *Second order ergodic theorem for self-similar tiling systems.* Ergodic Theory and Dynamical Systems.Available on CJO 2013 doi:10.1017/etds.2013.27.

[10] Karl Petersen. *Ergodic theory.* Second volume of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1983.

[11] Benjamin Weiss & Donald S. Ornstein, *Every transformation is bilaterally deterministic.* Israel Journal of Mathematics, 21(2-3),154-158,1975.