ISRAEL JOURNAL OF MATHEMATICS 77 (1992), 211–240

MEASURES ON THE CIRCLE INVARIANT UNDER MULTIPLICATION BY A NONLACUNARY SUBSEMIGROUP OF THE INTEGERS

BY

AIMEE S. A. JOHNSON

Department of Mathematics, Tufts University
Medford, MA 02155, USA

ABSTRACT

Let $\mathcal{S}$ be a nonlacunary subsemigroup of the natural numbers and let μ be an $\mathcal{S}$-invariant and ergodic measure. Using entropy arguments on a symbolic representation of the inverse limit of this action, we show that if any element in $\mathcal{S}$ has positive entropy with respect to μ, then μ is Lebesgue.

1. Introduction

In this paper we want to explore multiplication on the interval mod 1. Let $\mathcal{S}$ be a semigroup generated by such maps. Consider the set $\mathcal{M}$ of Borel probability measures, invariant and ergodic for $\mathcal{S}$. If $\mathcal{S}$ is generated by just one map then $\mathcal{M}$ is very large. If $\mathcal{S}$ is generated by two numbers that are powers of the same number, then $\mathcal{M}$ is still large. This is because we obtain all the measures from the singly generated semigroup. But in other situations, $\mathcal{M}$ is quite different.

We can characterize a semigroup Σ of $\mathbb{N}$ to be nonlacunary if it is not contained in a singly generated semigroup. For example, the semigroups generated by 2 and 3, or 6 and 10, are both nonlacunary. In [F] Furstenberg showed that any closed subset of [0, 1) invariant under a nonlacunary semigroup of integers must be finite or all of [0, 1). He conjectured that a stronger result held, that any invariant ergodic Borel probability measure for such a semigroup must be either atomic or Lebesgue.

Under a stronger hypothesis, Lyons [L] obtained the following result. If Σ is generated by two relatively prime integers and if either one of these two elements

Received September 27, 1990 and in revised form May 1, 1991

is exact as a measure-preserving endomorphism then μ is Lebesgue measure. This raised the issue of using entropy, which Rudolph [R] incorporated, to show that if μ is ergodic for a relatively prime pair of integers and if either map has positive entropy with respect to μ then μ is Lebesgue measure. The purpose of this paper is to extend Rudolph's result to the following:

THEOREM A: *Let p and q generate a nonlacunary subsemigroup of the integers. Let $T = \times p \pmod 1, S = \times q \pmod 1$ on the circle. If μ is a Borel probability measure invariant and ergodic for T and S then either μ is Lebesgue measure or $h_\mu(T) = 0 = h_\mu(S)$.*

Theorem A tells us that in the case of 2 maps, $\mathcal{M}$ is a small set, containing only Lebesgue measure and measures of entropy zero. The main body of this paper will prove Theorem A. We will first use Theorem A to prove the general result:

THEOREM B: *Let Σ be any multiplicative nonlacunary subsemigroup of $\mathbb{N}$ and let μ be an invariant Borel probability measure that is ergodic for Σ. Then either μ is Lebesgue measure or $h_\mu(t) = 0$ for every $t \in \Sigma$, where t represents multiplication by t mod 1 on the circle.*

Proof: Let us assume there exists some t with $h_\mu(t) > 0$. Consider this element $t \in \Sigma$. We want to show that there is a doubly generated nonlacunary subsemigroup of Σ that contains t. Since $\Sigma \subseteq \mathbb{N}$ there is a smallest number $a \in \mathbb{N}$ such that t is a power of a. By definition of nonlacunary we can find $s \notin \{a^n\}_{n=1}^{\infty}$, such that the semigroup S generated by s and t is a nonlacunary subsemigroup of Σ as wanted.

By [Ro] there is a decomposition of the measure μ denoted by $\mu = \int \mu_z dz$ where μ_z is ergodic for S. It is well known that $h_\mu(b) = \int h_{\mu_z}(b)dz$ for $b \in S$. Since we are assuming $h_\mu(t) > 0$, the set $\{z : h_{\mu_z}(t) > 0\}$ must have positive measure. By Theorem A such μ_z are Lebesgue. Then we can decompose μ as $\mu = \alpha L + (1-\alpha)\bar{\mu}$ where L is Lebesgue measure and $\alpha > 0$. But both L and μ are invariant and ergodic for Σ and thus we must have $\alpha = 1$. Thus if $h_\mu(t) > 0$ for any t we have $\mu = L$. This in turn shows $h_\mu(t) > 0$ for every $t \in \Sigma$ so in particular we have shown that $h_\mu(t) = 0$ for one t implies $h_\mu(t) = 0$ for every t. This completes the proof of Theorem B. ∎

Notice that Theorem B reduces the Furstenberg conjecture to the entropy zero case.

2. The Symbolic Representation, Measures, and Entropy

In this chapter we review material from [R]. Since the proofs appear in that paper, we will just state the results here. Let T_0 be multiplication by p mod 1 and S_0 be multiplication by q mod 1. We will describe the action of T_0 and S_0 on the circle symbolically.

Partition the circle, [0,1), into pq intervals

$$I_j = \left[\frac{j}{pq}, \frac{j+1}{pq}\right]_{j=0}^{pq-1}.$$

Notice that

$$T_0(I_j) = p \times \left[\frac{j}{pq}, \frac{j+1}{pq}\right] = \left[\frac{pj}{pq}, \frac{p(j+1)}{pq}\right].$$

Let $i = pj$ mod pq. Then we can write this image as

$$\left[\frac{i}{pq}, \frac{i+p}{pq}\right]_{j=0}^{pq-1} = I_i \cup \cdots \cup I_{i+p-1}.$$

Similarly, $S_0(I_j) = I_k \cup \cdots \cup I_{k+q-1}$ where $k = qj$ mod pq. Thus $\{I_j\}_{j=0}^{pq-1}$ forms a Markov partition for both T_0 and S_0. Let

$$V = \left\{x \in [0,1) : x = \frac{t}{p^n q^m}; n, m, t \in \mathbb{N}\right\}.$$

V contains the points for which there exists n and m such that $T_0^{n-1} S_0^{m-1} x$ lie on a boundary of our partition.

Define $F_T(i) = \{j : I_j \subset T_0(I_i)\}$ and similarly let $F_s(i) = \{j : I_j \subset S_0(I_i)\}$. These are the 'followers' of a symbol i for the maps T_0 and S_0, respectively.

Now we can associate to each T_0 and S_0 a $pq \times pq$ transition matrix of 0's and 1's: $M_T = [a_{ij}]$ where $a_{ij} = 1$ iff $j \in F_T(i), M_S = [b_{ij}]$ where $b_{ij} = 1$ iff $j \in F_s(i)$. Let $\Sigma = \{0, 1, \ldots, pq-1\}$ be the state space associated with these matrices.

If $[i_0, i_1, \cdots, i_{n-1}]$ is a finite word of elements of Σ with all $a_{i_k i_{k+1}} = 1$ then $\bigcap_{j=0}^{n-1} T_0^{-j}(I_{i_j})$ is an interval $\left[t/p^n q, (t+1)/p^n q\right]$. Thus to any one-sided infinite M_T-allowed word $\vec{i} = [i_0, i_1, \ldots]$ there corresponds a point $x_{\vec{i}} = \bigcap_{j=0}^{\infty} T_0^{-j}(I_{i_j}) \in [0,1)$.

Similarly for any M_s-allowed word

$$\vec{i} = \begin{bmatrix} i_{m-1} \\ \cdot \\ \cdot \\ \cdot \\ i_1 \\ i_0 \end{bmatrix},$$

$\bigcap_{j=0}^{m-1} S_0^{-j}(I_{i_j})$ is an interval $[t/pq^m, (t+1)/pq^m]$ and to each one-sided infinite M_s-allowed word there corresponds a point $x = \bigcap_{j=0}^{\infty} S_0^{-j}(I_{i_j})$.

Let $\mathbb{N} = \{0, 1, 2, \ldots\}$ and let $Y \subseteq \Sigma^{\mathbb{N}^2}$ consist of all arrays which are M_T-allowed on rows and M_s-allowed on columns. We can think of a point $y \in Y$ as a 'first quadrant' of symbols, where there is a symbol at each nonnegative lattice point. Let T be the left shift and S the down shift. That is, $Ty(i,j) = y(i+1,j)$ and $Sy(i,j) = y(i,j+1)$.

To any point $x \in [0,1) \setminus V$ there corresponds a unique point $y_x \in Y$. Just set $y_x(n,m) = j$ if $T_0^n S_0^m x \in I_j$. Recall that for $x \in V$ there exists k, r, s such that $x = k/p^r q^s$. Thus for all $n \geq r-1$ and $m \geq s-1, T_0^n S_0^m x$ is on the boundary of two I_j's. The symbol at (n,m) could indicate the left or right interval. However, if we specify the left (right) interval at $(r-1, s-1)$ then in order to obey the transition rules we must take the left (right) interval at all (n,m) with $n \geq r-1$ and $m \geq s-1$. So there are two points in Y that represent each $x \in V$.

Remark 2.1: For any symbols $a_0, a_1, a_2, \ldots \in \Sigma$ there exists $y \in Y$ with $y(i,i) = a_i$. All such y will agree on $y(i,j), i \neq j$. ∎

Remark 2.2: Consider the map $\varphi(y) = \bigcap_{i=0}^{\infty} T_0^{-i} S_0^{-i}[I_{y(i,i)}]$. This is a map from (Y, S, T) to $([0,1), S_0, T_0)$ which is 1 to 1 everywhere except on the countable set V where it is 2 to 1. ∎

Put the product topology on $Y; \varphi$ is continuous. Let $V^* \subset Y$ be those countably many points with $\varphi(y) \in V$. Note that V^* is invariant for both T and S.

Remark 2.3: Any M_T-allowed horizontal ray of symbols $i_{(n,m)} i_{(n+1,m)} \cdots$ determines all symbols $y(j,k), j \geq n, k \geq m$ of any $y \in Y$ with $y(\tilde{n}, m) = i_{(\tilde{n},m)}, \tilde{n} \geq n$, as long as $y \notin V^*$.

Similarly any M_S-allowed vertical ray of symbols $i_{(s,t)}i_{(s,t+1)}\cdots$ determines all symbols $y(i,j)$, $i \geq s$, $j \geq t$ of any $y \in Y$ with $y(s,j) = i_{(s,j)}$, $j \geq t$ as long as $y \notin V^*$.

Pictorially this means that a vertical ray of symbols determines all symbols to its right and a horizontal ray of symbols determines all symbols above it. ∎

Let $\hat{Y} \subseteq \Sigma^{\mathbf{Z}^2}$ be those doubly infinite arrays where all rows are M_T-allowed and all columns are M_S-allowed. For $\hat{y} \in \hat{Y}$, let $\varphi(\hat{y})$ be the point in $[0,1)$ associated with the first quadrant. Let T and S still represent left and down shifts. Note that $T_0\varphi = \varphi T$ and $S_0\varphi = \varphi S$.

Next, let $\mathcal{M}$ be the space of all T_0 and S_0 invariant Borel probability measures on $[0,1)$. This is a weakly compact convex space. Let $\mathcal{M}_0 \subset \mathcal{M}$ be the ergodic measures minus the point mass at zero. In this last case it is trivially true that $h_\mu(T_0) = 0 = h_\mu(S_0)$.

Remark 2.4: If $\mu \in \mathcal{M}$ and $x \in V, x \neq 0$ then $\mu(x) = 0$. ∎

Because V is a T_0 and S_0 invariant set, any $\mu \in \mathcal{M}_0$ must give it zero or full measure. Using this remark and that we've already excluded the point mass to zero, it must be that $\mu(V) = 0$.

Remark 2.5: Any measure $\mu \in \mathcal{M}_0$ lifts to a unique T and S invariant Borel probability measure on $\hat{Y}$. ∎

Let $\hat{\mathcal{M}}$ be the T and S invariant Borel probability measures on $\hat{Y}$ and $\hat{\mathcal{M}}_0$ be the ergodic ones (excluding the point mass at 0).

Let P be the partition of $\hat{Y}$ according to the symbol $\hat{y}(0,0)$.

Remark 2.6: For any $\hat{\mu} \in \hat{\mathcal{M}}$, $\bigvee_{j=0}^{\infty} T^{-j}(P) = \bigvee_{j=0}^{\infty} S^{-j}(P)$ $\hat{\mu}$-a.e. ∎

Remark 2.7: For $\hat{\mu} \in \hat{\mathcal{M}}$, $h_{\hat{\mu}}(T) = h_{\hat{\mu}}(T,P)$ and $h_{\hat{\mu}}(S) = h_{\hat{\mu}}(S,P)$. ∎

THEOREM 2.8: *For $\hat{\mu} \in \hat{\mathcal{M}}_0$ and any T and S invariant algebra $\mathcal{A}$,*

$$h_{\hat{\mu}}(T,\mathcal{A}) = \frac{\log p}{\log q} h_{\hat{\mu}}(S,\mathcal{A}).$$

For the rest of the paper we fix an arbitrary $\mu \in \mathcal{M}_0$.

3. The δ_0 Distribution

Recall that p and q are two integers that generate a nonlacunary semigroup. Specifically, we can write

$$p = p_0 \pi_1^{n_1} \cdots \pi_h^{n_h}, \quad q = q_0 \pi_1^{m_1} \ldots \pi_h^{m_h}$$

where $p_0, q_0, \pi_1, \ldots, \pi_h$ are pairwise relatively prime integers, $n_1/m_1 > \cdots > n_h/m_h$ and either $p_0 \neq 1$ or $h \geq 2$. This form comes from writing p and q in their prime number decomposition and grouping the terms as shown.

Fix a point $\hat{y} \in \hat{Y}$. Let $\varphi(\hat{y}) = x \in [0,1)$. Then

$$T_0^{-1}(x) = \left\{ \frac{x}{p} + \frac{i}{p} \right\}_{i=0}^{p-1}.$$

For $\hat{y}$ this corresponds to the p possible symbols that could be at position $(-1,0)$ that are consistent with the first quadrant of $\hat{y}$.

Given this 1–1 correspondence, the terminologies will often be intertwined. In particular, $\{\frac{x}{p} + \frac{i}{p}\}_{i=0}^{p-1}$ will be referred to as the preimages associated to $(-1,0)$.

Define k_{01} to be the smallest integer such that $k_{01}m_1 \geq n_1$. Clearly $k_{01}m_i \geq n_i$ for $i = 1$ to h. Notice that k_{01} is defined so that

$$\frac{q^{k_{01}}}{p} = \frac{q_0^{k_{01}} \pi_1^{m_1 k_{01}} \cdots \pi_h^{m_h k_{01}}}{p_0 \pi_1^{n_1} \ldots \pi_h^{n_h}} = \frac{q_0 \pi_1^{m_1 k_{01} - n_1} \cdots \pi_h^{m_h k_{01} - n_h}}{p_0}.$$

Thus

$$S_0^{k_{01}} T_0^{-1}(x) = \left\{ \frac{q^{k_{01}} x}{p} + \frac{q^{k_{01}} i}{p} \right\}_{i=0}^{p-1} = \left\{ \frac{q^{k_{01}} x}{p} + \frac{i'}{p_0} \right\}_{i'=0}^{p_0-1}.$$

This corresponds in our symbolic representation to the p_0 possible symbols at $(-1, k_{01})$, given a first quadrant of symbols. Notice that

$$S_0^k T_0^{-1}(x) = \left\{ \frac{q^k x}{p} + \frac{i}{p_0} \right\}_{i=0}^{p_0-1} \quad \text{for every } k \geq k_{01}$$

so in fact there are p_0 possible symbols at $(-1,k), k \geq k_{01}$, given a first quadrant of symbols. As before, we will often refer to $\{\frac{q^k x}{p} + \frac{i}{p_0}\}_{i=0}^{p_0-1}$ as the preimages associated with $(-1,k)$. Let x_{01} be that particular $\frac{q^k x}{p} + \frac{i}{p_0}$ which is the actual value of $\varphi(S^{k_{01}}T^{-1}\hat{y})$.

Definition 3.1:

$$\delta_0(\hat{y},1)\left[\frac{i}{p_0}\right] = E_{\hat{\mu}}\left[S^{-k_{01}}T\varphi^{-1}(x_{01}+\frac{i}{p_0})|\bigvee_{i=0}^{\infty}T^{-1}(P)\right](\hat{y}).$$

∎

$\delta_0(\hat{y},1)$ is a distribution on

$$\left\{0,\frac{1}{p_0},\frac{2}{p_0},\ldots,\frac{p_0-1}{p_0}\right\}$$

where the number associated with 0 is the expectation of the actual preimage.

Next consider

$$T_0^{-r}(x) = \left\{\frac{x}{p^r}+\frac{i}{p^r}\right\}_{i=0}^{p^r-1}.$$

Let k_{0r} be the smallest integer such that $k_{0r}m_1 \geq rn_1$. As before this also insures that $k_{0r}m_i \geq rn_i$ for $i = 1,\ldots,h$. Then

$$\begin{aligned} S_0^{k_{0r}}T_0^{-r}(x) &= \left\{\frac{g^{k_{0r}}x}{p^r}+\frac{q^{k_{0r}}i}{p^r}\right\}_{i=0}^{p^r-1} \\ &= \left\{\frac{q^{k_{0r}}x}{p^r}+\frac{q_0^{k_{0r}}\pi_i^{k_{0r}m_1-rn_1}\cdots\pi_h^{k_{0r}m_h-rn_h}i}{p_0^r}\right\}_{i=0}^{p_0^r-1}. \end{aligned}$$

These are the p_0^r possible preimages associated with position $(-r,k_{0r})$. Let x_{0r} be that particular $\frac{q^{k_{0r}}x}{p^r}+\frac{i'}{p_0^r}$ which is the actual value of $\varphi(S^{k_{0r}}T^{-r}\hat{y})$.

Definition 3.1 (General):

$$\delta_0(\hat{y},r)\left[\frac{i}{p_0^r}\right] = E_{\hat{\mu}}\left[S^{-k_{0r}}T^r\varphi^{-1}(x_{0r}+\frac{i}{p_0^r})|\bigvee_{i=0}^{\infty}T^{-i}(P)\right](\hat{y}).$$

∎

This is the probability that $\varphi(\hat{y})$ extends under $S_0^{k_{0r}}T_0^{-r}$ to a point translated by i/p_0^r from $\varphi(S^{k_{0r}}T^{-r}\hat{y})$.

For the next lemma we want to see what happens to this distribution when S is applied to $\hat{y}$.

Notation: $\delta_0(\hat{y},r)$ is the distribution put on the points $\{i/p_0^r\}_{i=0}^{p_0^r-1}$ as just described. Let $\hat{y}_j$ be an element in $\hat{Y}$ with $\varphi(\hat{y}_j) = x/q + j/q \in S_0^{-1}\varphi(\hat{y})$ with $\hat{y}_j(i,\ell+1) = \hat{y}(i,\ell)$ for $\ell \geq 0$ and all i. Let $\bar{x}_{0r} = \varphi(S^{k_{0r}-1}T^{-r}\hat{y})$. Then $\delta_0(\hat{y}_j,r)$ is the distribution put on the points $\{i'/p_0^r\}_{i'=0}^{p_0^r-1}$ where $\delta_0(\hat{y}_j,r)[i'/p_0^r]$ is the expectation of $\bar{x}_{0r}+i'/p_0^r$ given $x/q+j/q$. Define $S\delta_0(\hat{y}_j,r)$ to be the distribution on $\{i/p_0^r\}_{i=0}^{p_0^r-1}$ where $S\delta_0(\hat{y}_j,r)[i/p_0^r]$ is the expectation of $x_{0r}+i/p_0^r$ given $x/q+j/q$; this is just $\delta_0(\hat{y}_j,r)$ rearranged by the correspondence between $\{\bar{x}_{0r} + i'/p_0^r\}$ and $\{x_{0r} + i/p_0^r\}$ given by multiplication by q mod p_0^r, which corresponds in the symbolic space to an application of S. Also, let

$$E\left[\frac{x}{q}+\frac{j}{q}\middle| x\right] = E_{\hat{\mu}}\left[S\varphi^{-1}(\frac{x}{q}+\frac{j}{q})|\bigvee_{i=0}^{\infty}T^{-i}(P)\right](\hat{y}).$$

LEMMA 3.2: *$\delta_0(\hat{y},r) = S\delta_0(\hat{y}_j,r)$ for $j = 0$ to $q-1$.*

Proof: We see $x_{0r} + i/p_0^r$ at $S_0^{k_{0r}}T_0^{-r}(x)$ only if for some j,

$$S_0^{-1}(x) = \frac{x}{q}+\frac{j}{q} \quad \text{and} \quad S_0^{k_{0r}+1}T_0^{-r}(\frac{x}{q}+\frac{j}{q}) = x_{0r} + \frac{i}{p_0^r}.$$

We can write this as

$$\delta_0(\hat{y},r) = \sum_{j=0}^{q-1} E\left[\frac{x}{q}+\frac{j}{q}\middle| x\right] S\delta_0(\hat{y}_j,r).$$

Taking the entropy,

$$\begin{aligned} h[\delta_0(\hat{y},r)] &= h\left[\sum_{j=0}^{q-1} E\left[\frac{x}{q}+\frac{j}{q}\middle| x\right] S\delta_0(\hat{y}_j,r)\right] \\ &\geq \sum_{j=0}^{q-1} E\left[\frac{x}{q}+\frac{j}{q}\middle| x\right] h[S\delta_0(\hat{y}_j,r)] \\ &= \sum_{j=0}^{q-1} E\left[\frac{x}{q}+\frac{j}{q}\middle| x\right] h[\delta_0(\hat{y}_j,r)] \end{aligned}$$

where the inequality is an equality iff $S\delta_0(\hat{y}_j,r)$ is the same for every j and thus in fact equals $\delta_0(\hat{y},r)$. Notice that

$$h[\delta_0(\hat{y}_j,r)] = h[S\delta_0(\hat{y}_j,r)] \leq h[\delta_0(\hat{y},r)]$$

because in the last we are conditioning on less information. But

$$\int h[\delta_0(\hat{y}_j, r)]d\hat{\mu} = \int h[\delta_0(\hat{y}, r)]d\hat{\mu}$$

using the S-invariance of $\hat{\mu}$. Together these give

$$h[\delta_0(\hat{y}_j, r)] = h[\delta_0(\hat{y}, r)] \quad \text{for every } j$$

and we get the necessary equality. ∎

The above lemma says that the expectations of the symbols at $(-r, k_{0r})$ are independent of the symbol at $(0, -1)$.

COROLLARY 3.3:

$$\delta_0(S\hat{y}, r)[\frac{qi}{p_0^r} \bmod 1] = \delta_0(\hat{y}, r)[\frac{i}{p_0^r}].$$

Proof: On the right is the expected value of $x_{0r} + i/p_0^r$, given $\varphi(\hat{y})$. By Lemma 3.2, this is the same as the expected value of $q \times (x_{0r} + i/p_0^r)$ given $\varphi(S\hat{y})$. But $q \times x_{0r} = \varphi(S^{k_{0r}}T^{-r}S\hat{y})$ so this equals the expected value of the real preimage plus qi/p_0^r mod 1, given $\varphi(S\hat{y})$. ∎

Recall that in the symbolic representation, we are given the first quadrant of symbols and we have various posibilities for the symbolic paths in the second quadrant, which correspond to preimages of $x = \varphi(\hat{y})$. For δ_0 we are interested in the preimages that correspond to each position $(-r, k_{0r})$; these are the possible preimages of the point x under the map $S^{k_{0r}}T^{-r}$ and we know there are p_0^r of them. To each preimage associated with $(-r+1, k_{0\,(r-1)})$ there corresponds p_0 possible preimages associated with $(-r, k_{0r})$, altogether making up the total p_0^r. The next lemma states that the probability of a certain preimage corresponding to $(-r+1, k_{0\,(r-1)})$ is exactly the sum of the probabilities of the associated p_0 preimages corresponding to $(-r, k_{0r})$.

LEMMA 3.4:

$$\delta_0(\hat{y}, r-1)\left[\frac{i}{p_0^{r-1}}\right] = \Sigma\delta_0(\hat{y}, r)\left[\frac{j}{p_0^r}\right]$$

where this sum is over all j such that $j\pi_1^{n_1}\cdots\pi_h^{n_h} = iq^{k_{0r}-k_{0(r-1)}}$, mod p_0^{r-1}.

Proof: The expectation of $x_{0r} + j/p_0^r$ equals the expectation of $x_{0\,(r-1)} + i/p_0^{r-1}$ multiplied by the expectation of $x_{0r} + j/p_0^r$ given $x_{0\,(r-1)} + i/p_0^{r-1}$. In other words,

$$\delta_0(\hat{y},r)\left[\frac{j}{p_0^r}\right]=\delta_0(\hat{y},r-1)\left[\frac{i}{p_0^{r-1}}\right]\cdot$$
$$\cdot E_{\hat{\mu}}\left[S^{-k_{0r}}T^r\varphi^{-1}(x_{0r}+\frac{j}{p_0^r})|\bigvee_{i=0}^{\infty}T^{-i}(P)\vee\bigvee_{i=1}^{r-1}S^{-k_{0i}}T^i(P)\right](\hat{y}).$$

This is for i and j such that

$$T_0(x_{0r}+\frac{j}{p_0^r})=S^{k_{0r}-k_{0(r-1)}}(x_{0\,(r-1)}+\frac{i}{p_0^{r-1}}).$$

Fix i, sum over all such valid j. Note that the sum of the second terms is exactly one. The j's thus described are such that

$$\frac{pj}{p_0^r}=\frac{q^{k_{0r}-k_{0\,(r-1)}}i}{p_0^{r-1}}$$

$$\Leftrightarrow\qquad \frac{\pi_1^{n_1}\dots\pi_h^{n_h}j}{p_0^{r-1}}=\frac{q^{k_{0r}-k_{0\,(r-1)}}i}{p_0^{r-1}}$$

as wanted. ∎

COROLLARY 3.5: *$\delta_0(T^r\hat{y},2r)$ determines $\delta_0(T^i\hat{y},r+i)$ for $0\le i\le r$.*

Proof: By Corollary 3.3, $\delta_0(T^r\hat{y},2r)$ determines $\delta_0(S^{-k_{0\,(r-i)}}T^r\hat{y},2r)$. Let

$$w_{2r}=\varphi(S^{k_{0(2r)}}T^{-2r}S^{-k_{0\,(r-i)}}T^r\hat{y})$$

and

$$w_{r-i}=\varphi(S^{k_{0\,(r-i)}}T^{-r+i}S^{-k_{0\,(r-i)}}T^r\hat{y})=\varphi(T^i\hat{y}).$$

The expectation of $w_{2r}+j/p_0^{2r}$, given

$$\varphi(S^{-k_{0\,(r-i)}}T^r\hat{y}),$$

is the expectation of $w_{2r}+j/p_0^{2r}$ given w_{r-i} multiplied by the expectation of w_{r-i} given $\varphi(S^{-k_{0\,(r-i)}}T^r\hat{y})$. We can write this as

$$\delta_0(S^{-k_{0\,(r-i)}}T^r\hat{y},2r)\left[\frac{j}{p_0^{2r}}\right]=$$
$$E_{\hat{\mu}}\left[S^{-k_{0(2r)}}T^{2r}\varphi^{-1}\left(x_{2r}+\frac{j}{p_0^{2r}}\right)\middle|S^{-k_{0\,(r-i)}}T^{r-i}\varphi^{-1}(w_{r-i})\right]$$
$$\times E_{\hat{\mu}}[S^{-k_{0\,(r-i)}}T^{r-i}\varphi^{-1}(w_{r-i})|\varphi^{-1}\varphi(S^{-k_{0\,(r-i)}}T^r(\hat{y})],$$

but only if $w_{2r}+j/p_0^{2r}$ and w_{r-i} are consistent with one another. In other words, j must satisfy

$$p^{r+i}\left(x_{2r}+\frac{j}{p_0^{2r}}\right)=q^{k_{0(2r)}-k_{0(r-i)}}(w_{r-i}),$$

which will occur iff

$$\frac{p^{r+i}j}{p_0^{2r}}=0 \bmod 1 \Leftrightarrow \frac{\pi_1^{n_1(r+i)}\cdots\pi_h^{n_h(r+i)}j}{p_0^{r-i}}=0 \bmod 1$$
$$\Leftrightarrow j=kp_0^{r-i} \quad \text{for } k=0,\cdots,p_0^{r-i}-1.$$

Then

$$\delta_0(S^{-k_{0(r-i)}}T^r\hat{y},2r)\left[\frac{kp_0^{r-i}}{p_0^{2r}}\right]=\delta_0(T^i\hat{y},r+i)\left[\frac{k}{p_0^{r+i}}\right]\delta_0(S^{-k_{0(r-i)}}T^r\hat{y},r-i)[0].$$

But we know the first and last terms, and thus can find the middle term. ∎

LEMMA 3.6: *If $\hat{y}_1$ and $\hat{y}_2$ agree on their first quadrants, then $\delta_0(\hat{y}_1,r)$ and $\delta_0(\hat{y}_2,r)$ differ by a translation mod 1 of size $\varphi(S^{k_{0r}}T^{-r}\hat{y}_2)-\varphi(S^{k_{0r}}T^{-r}\hat{y}_1)$.*

Proof: By definition of δ_0, only the first quadrant is used to find the expectations. The value at $(-r,k_{0r})$ is used only to determine the ordering which begins with the expectation of $\varphi(S^{k_{0r}}T^{-r}\hat{y})$. Thus $\delta_0(\hat{y}_1,r)$ and $\delta_0(\hat{y}_2,r)$ have the same weights but with different starting points. ∎

4. Preimages and Their Movement

Definition 4.1: Define k_{aj} for $a=1,\ldots,h-1$ and $j\geq 1$ to be the smallest integer such that

$$jn_{a+1}-k_{aj}(m_{a+1}n_a-n_{a+1}m_a)\leq 0.$$

[Note that $n_a/m_a>n_{a+1}/m_{a+1}$, so such an integer exists.]

Define $k_{hj}=0$ for every $j\geq 1$.

Definition 4.2:

$$\mathcal{D}_{a-1}=\bigvee_{i=0}^{\infty}T^{-i}(P)\vee\bigvee_{i=1}^{\infty}S^{-k_{0i}}T^i(P)\vee\bigvee_{i=1}^{\infty}\bigvee_{j=1}^{a-1}S^{-k_{ji}n_j}T^{i+k_{ji}m_j}(P).$$

$$\mathcal{D}_0=\bigvee_{i=0}^{\infty}T^{-i}(P)\vee\bigvee_{i=1}^{\infty}S^{-k_{0i}}T^i(P).$$

Let $\mathcal{D}_{a-1}(\hat{y})$ be the array of symbols given by the first quadrant of $\hat{y},\hat{y}(-i,k_{0i})$ for $i\geq 1$, and $\hat{y}(-i-k_{ji}m_j,k_{ji}n_j)$ for $i\geq 1, j=1$ to $a-1$. ∎

In Section 5 we will construct distributions $\delta_a, a = 1, \ldots, h$, that will be similar to δ_0. Instead of moving the preimages up by S, we will move them diagonally by $S^{n_a}T^{-m_a}$ ('up the staircase'). $\mathcal{D}_{a-1}$ holds the information from previous staircases.

The objective for this chapter is to show that given $\mathcal{D}_{a-1}$, the possible preimages corresponding to position $(-r, 0)$ are a coset of the group

$$\left\{\frac{i}{\pi_a^{rn_a}\cdots\pi_h^{rn_h}}\right\}, \quad i = 0, \cdots, (\pi_a^{rn_a}\cdots\pi_h^{rn_h} - 1),$$

the possible preimages corresponding to position $(-r - k_{ar}m_a, k_{ar}n_a)$ are a coset of the group

$$\left\{\frac{i}{\pi_a^{rn_a}}\right\}, \qquad i = 0, \cdots, \pi_a^{rn_a} - 1,$$

and the correspondence between the two is a $\pi_{a+1}^{rn_{a+1}}\cdots\pi_h^{rn_h}$ to one map. Further application of $S^{n_a}T^{-m_a}$ will yield a 1–1 map with movement given by multiplication by an integer that depends only on r and a. ∎

Notation: When the letter d is used in prescribing the range of an index, it always refers to the denominator of the fraction appearing in the expression in question.

If $\mathcal{D}_0(\hat{y})$ is given then it is enough to assume $\mathcal{D}_0(\hat{y}) = \vec{0}$, since the possible preimages in the general case are just a coset of those found in this case. In section 3 we showed that

$$T_0^{-r}(0) = \left\{\frac{i}{p^r}\right\}_{i=0}^{d-1}$$

had

$$S_0^{k_{0r}}T_0^{-r}(0) = \left\{\frac{i'}{p_0^r}\right\}_{i'=0}^{d-1}.$$

If we specify $i' = 0$ (i.e. that $\mathcal{D}_0(\hat{y}) = \vec{0}$) then we are left with

$$T_0^{-r}(0) = \left\{\frac{i}{\pi^{rn_1}\cdots\pi_h^{rn_h}}\right\}_{i=0}^{d-1}.$$

We will prove the objective of this section by induction on a. It will proceed as follows: assume that if $\mathcal{D}_{a-1}(\hat{y}) = \vec{0}$ then the possible preimages associated to $(-r, 0)$ are $\{i/\pi_a^{rn_a}\cdots\pi_h^{rn_h}\}_{i=0}^{d-1}$. This is true for $a = 1$, as stated above. We will show that:

(i) The possibilities associated to $(-r-k_{ar}m_a,\ k_{ar}n_a)$ are $\{i/\pi_a^{rn_a}\}_{i=0}^{d-1}$ and the map between the possible preimages associated to $(-r,0)$ and $(-r-k_{ar}m_a,\ k_{ar}n_a)$ is $\pi_{a+1}^{rn_{a+1}}\cdots\pi_h^{rn_h}$ to 1.

(ii) For the next step in the induction it is enough to consider $\mathcal{D}_a(\hat{y})=\vec{0}$. This will reduce the possible preimages corresponding to $(-r,0)$ down to $\{i/\pi_{a+1}^{rn_{a+1}}\cdots\pi_h^{rn_h}\}_{i=0}^{d-1}$.

LEMMA 4.3: *Given $\mathcal{D}_{a-1}(\hat{y})=\vec{0}$, the preimage associated to $(-r,0)$ determines the preimage at $(-r-km_a,kn_a)$ for every $k\geq 0$.*

Proof: If the preimage associated to $(-r,0)$ is fixed as $b/\pi_a^{n_a r}\cdots\pi_h^{n_h r}$, then the possible preimages associated to $(-r-km_a,0)$ are the

$$\left\{\frac{j}{\pi_a^{n_a(r+km_a)}\cdots\pi_h^{n_h(r+km_a)}}\right\}_{j=0}^{d-1}$$

such that

$$T_0^{km_a}\left(\frac{j}{\pi_a^{n_a(r+km_a)}\cdots\pi_h^{n_h(r+km_a)}}\right)=\frac{b}{\pi_a^{n_a r}\cdots\pi_h^{n_h r}}.$$

Thus the possibilities here have the form $\left\{\bar{u}+i/\pi_a^{n_a km_a}\cdots\pi_h^{n_h km_a}\right\}_{i=0}^{d-1}$. Apply $S_0^{kn_a}$ to get

$$\left\{q^{kn_a}\bar{u}+\frac{q_0\pi_1^{m_1 kn_a}\cdots\pi_a^{m_a kn_a}\cdots\pi_h^{m_h kn_a}i}{\pi_a^{n_a km_a}\cdots\pi_h^{n_h km_a}}\right\}_{i=0}^{d-1}=$$

$$\left\{q^{kn_a}\bar{u}+q'\pi_{a+1}^{k(n_a m_{a+1}-m_a n_{a+1})}\cdots\pi_h^{k(n_a m_h-m_a n_h)}i\right\}=q^{kn_a}\bar{u}\bmod 1.$$

This gives the preimage as wanted.

In order to prove statement (i) consider what happens to $\{i/\pi_a^{rn_a}\cdots\pi_h^{rn_h}\}_{i=0}^{d-1}$ when we 'move up the staircase'. Associated to $(-r-km_a,0)$ we have

$$\left\{\frac{i}{\pi_a^{na(r+km_a)}\cdots\pi_h^{n_h(r+km_a)}}\right\}_{i=0}^{d-1}.$$

Apply $S_0^{kn_a}$ to get

$$\left\{\frac{q_0^{kn_a}\pi_1^{m_1 kn_a}\cdots\pi_h^{m_h kn_a}i}{\pi_a^{n_a(r+km_a)}\cdots\pi_h^{n_h(r+km_a)}}\right\}_{i=0}^{d-1}=$$

$$\left\{ \frac{q_0^{kn_a}(\pi_1^{m_1} \cdots \pi_{a-1}^{m_{a-1}})^{kn_a}(\pi_{a+1}^{m_{a+1}n_a - n_{a+1}m_a} \cdots \pi_h^{m_h n_a - n_h m_a})^k i}{\pi_a^{rn_a} \cdots \pi_h^{rn_h}} \right\}.$$

Since $n_a/m_a > n_i/m_j$ for $j = a+1, \ldots, h$, the terms in the above fraction have positive exponents. Let $k = k_{ar}$. By Definition 4.1 we have

$$k_{ar}(m_{a+1}n_a - n_{a+1}m_a) \geq rn_{a+1}.$$

Claim 4.4: $k_{ar}(m_i n_a - n_i m_a) \geq rn_i$ for $i = a+1, \ldots, h$.

Proof: Since $n_{a+1}/m_{a+1} > n_i/m_i$ for $i \geq a+2$, we get

$$\frac{m_{a+1}}{n_{a+1}}\frac{n_a}{r} - \frac{m_a}{r} < \frac{m_i}{n_i}\frac{n_a}{r} - \frac{m_a}{r} \quad \text{for } i \geq a+2.$$

Using Definition 4.1 we see the left-hand side is $\geq 1/k_{ar}$ so that the right side must also be. Rearrange the terms to yield the result. ∎

Proof of (i): Using Claim 4.4 we see that the only remaining preimages associated with $(-r - k_{ar}m_a, k_{ar}n_a)$ are $\{i'/\pi_a^{rn_a}\}_{i'=0}^{d-1}$, so that the map must be $\pi_{a+1}^{rn_{a+1}} \cdots \pi_h^{rn_h}$ to 1. ∎

Proof of (ii): From Lemma 4.3 it is easy to see that the movement 'up the staircase' is an endomorphism. This tells us two things. First, if $i' = 0$ (i.e. $\mathcal{D}_a(\hat{y}) = \vec{0}$) then the possible preimages left at $(-r, 0)$ form a subgroup $\{i/b\}_{i=0}^{b-1}$. Secondly, if $i' \neq 0$ then the possible preimages at $(-r, 0)$ form a coset of this subgroup. By Lemma 4.3 these cosets are disjoint. Each coset has b elements and their union must equal the original $\pi_a^{rn_a} \cdots \pi_h^{rn_h}$ possible preimages at $(-r, 0)$. Thus $b = \pi_{a+1}^{rn_{a+1}} \cdots \pi_h^{rn_h}$. This completes the induction. ∎

What remains of our objectives for this section is to show that further application of $S^{n_a}T^{-m_a}$ to the preimages corresponding to $(-r - k_{ar}m_a, k_{ar}n_a)$ yields a 1–1 map with movement given by multiplication. We previously showed that the possible preimages associated with $(-r - k_{ar}m_a, k_{ar}n_a)$ are a coset of $\{j/\pi_a^{n_a r}\}_{j=0}^{d-1}$. Similar arguments show this for $(-r - km_a, kn_a)$ for any $k \geq k_{ar}$.

Lemma 4.5: *There is a 1–1 correspondence between the preimages at $(-r - km_a, kn_a)$ and $(-r - (k+1)m_a, (k+1)n_a)$ for $k \geq k_{ar}$.*

Proof: This follows just as the argument for the proof of (ii). Specifying the value at $(-r - (k+1)m_a, (k+1)n_a)$ restricts the values at $(-r - km_a, kn_a)$ to a coset of the possibilities that one gets when the first value is given as zero. Lemma 4.3 tells us all these cosets are disjoint and thus each coset can have only 1 element in it. ∎

The preimages associated to $(-r-km_a, kn_a)$, for $k \geq k_{ar}$, are

$$\varphi(S^{kn_a}T^{-r-km_a}\hat{y}) + \{i/\pi_a^{n_a r}\}$$

and at $(-r-(k+1)m_a, (k+1)n_a)$ they are

$$\varphi(S^{(k+1)n_a}T^{-r-(k+1)m_a}\hat{y}) + \{j/\pi_a^{n_a r}\}.$$

By Lemma 4.5 we can define a 1–1 map from $\{i/\pi_a^{n_a r}\}$ to $\{j/\pi_a^{n_a r}\}$. Denote this map by f_{ak}.

LEMMA 4.6: *f_{ak} is the same map for every $k \geq k_{ar}$. In other words, if*

$$f_{ak}(\frac{i}{\pi_a^{n_a r}}) = \frac{c}{\pi_a^{n_a r}}$$

then

$$f_{ak'}(\frac{i}{\pi_a^{n_a r}}) = \frac{c}{\pi_a^{n_a r}} \quad \text{for all } k' \geq k_{ar}.$$

Proof: Assume not. Then at k_1 we can find j_1 such that

$$\frac{q^{n_a} i}{p^{m_a}\pi_a^{rn_a}} + \frac{q^{n_a} j_1}{p^{m_a}} = \frac{c_1}{\pi_a^{n_a r}}$$

and at k_2 we can find j_2 such that

$$\frac{q^{n_a} i}{p^{m_a}\pi_a^{rn_a}} + \frac{q^{n_a} j_2}{p^{m_a}} = \frac{c_2}{\pi_a^{n_a r}}.$$

But then we have

$$\frac{q^{n_a}(j_1 - j_2)}{p^{m_a}} = \frac{c_1 - c_2}{\pi_a^{n_a r}}.$$

On the left we do not have $\pi_a^{rn_a}$ in the denominator, which is a contradiction. ∎

LEMMA 4.7: *The map from the preimages associated to $(-r-km_a, kn_a)$ to the preimages associated to $(-r-(k+1)m_a, (k+1)n_a)$ is given by multiplication by an integer we will denote by m_{ar}^*.*

Proof: The map f_{ar} is an isomorphism. The only group isomorphisms on the group $\{i/d\}_{i=0}^{d-1}$ is multiplication by an integer relatively prime to d. ∎

In review, this gives us the following. Given $\mathcal{D}_{a-1}(\hat{y})$, the possible preimages associated with $(-r,0)$ are a coset of $\{i/\pi_a^{rn_a}\cdots\pi_h^{rn_h}\}_{i=0}^{d-1}$. Apply the map $S_0^{n_a}T_0^{-m_a}$ to these. After k_{ar} steps we have only a coset of $\{i/\pi_a^{rn_a}\}_{i=0}^{d-1}$ and $S_0^{k_{ar}n_a}T_0^{-k_{ar}m_a}$ acts as a $\pi_{a+1}^{rn_{a+1}}\cdots\pi_h^{rn_h}$ to 1 map. Further application of $S_0^{n_a}T_0^{-m_a}$ acts in a 1–1 manner and moves $x_{ar}+i/\pi_a^{rn_a}$ to $q^{n_a}x_{ar}/p^{m_a}+m_{ar}^*i/\pi_a^{rn_a}$.

5. The Distribution δ_a For $a = 1$ to h

Recall that $(-r-k_{ar}m_a, k_{ar}n_a)$ is associated to the preimages $\{x_{ar}+i/\pi_a^{rn_a}\}_{i=0}^{d-1}$, where $x_{ar}=\varphi(S^{k_{ar}n_a}T^{-r-k_{ar}m_a}\hat{y})$.

Definition 5.1:

$$\delta_a(\hat{y},r)[\frac{i}{\pi_a^{rn_a}}]=E_{\hat{\mu}}\left[S^{-k_{ar}n_a}T^{r+k_{ar}m_a}\varphi^{-1}(x_{ar}+\frac{i}{\pi_a^{rn_a}})\middle|\mathcal{D}_{a-1}\right](\hat{y}).\quad\blacksquare$$

We want to compare $\delta_a(\hat{y},r)$ and $\delta_a(S^{n_a}T^{-m_a}\hat{y},r)$. Given $\varphi(\hat{y})=x$, the preimages associated to $(m_a,-n_a)$ are

$$\frac{p^{m_a}x}{q^{n_a}}+\frac{p^{m_a}j}{q^{n_a}}=\frac{p^{m_a}x}{q^{n_a}}+\frac{p_0^{m_a}\pi_1^{n_1m_a-n_am_1}\cdots\pi_{a-1}^{n_{a-1}m_a-n_am_{a-1}}j}{q_0^{n_a}\pi_{a+1}^{m_{a+1}n_a-m_an_{a+1}}\cdots\pi_h^{n_am_h-n_hm_a}}$$

which we will denote by $p^{m_a}x/q^{n_a}+p'j/q'$. Let $\hat{y}_j$ be that element in $\hat{Y}$ with

$$\varphi(\hat{y}_j)=\frac{p^{m_a}x}{q^{n_a}}+\frac{p'j}{q'}$$

and

$$\hat{y}_j(i-m_a,k+n_a)=\hat{y}(i,k)\quad\text{for } i\in\mathbb{Z},\quad k\geq 0.$$

$\delta_a(\hat{y},r)$ is the distribution put on $\{i/\pi_a^{rn_a}\}_{i=0}^{d-1}$ where $\delta_a(\hat{y},r)[i/\pi_a^{rn_a}]$ is the expectation of $x_{ar}+i/\pi_a^{rn_a}$ given $\mathcal{D}_{a-1}(\hat{y})$. Let

$$\bar{x}_{ar}=\varphi(S^{(k_{ar}-1)n_a}T^{-r-(k_{ar}-1)m_a}\hat{y}).$$

Then $\delta_a(\hat{y}_j,r)$ is the distribution put on $\{i'/\pi_a^{rn_a}\}_{i'=0}^{d-1}$ where $\delta_a(\hat{y}_j,r)[i'/\pi_a^{rn_a}]$ is the expectation of $\bar{x}_{ar}+i'/\pi_a^{rn_a}$, given $\mathcal{D}_{a-1}(\hat{y}_j)$. There is a 1–1 correspondence between $\{\bar{x}_{ar}+i'/\pi_a^{rn_a}\}_{i'=0}^{d-1}$ and $\{x_{ar}+i/\pi_a^{rn_a}\}_{i=0}^{d-1}$ by Lemma 4.5. Let $m_{ar}^*\delta_a(\hat{y}_j,r)$ be the distribution put on $\{i/\pi_a^{rn_a}\}_{i=0}^{d-1}$ by $m_{ar}^*\delta_a(\hat{y}_j,r)[i/\pi_a^{rn_a}]$

equals the expectation of $x_{ar} + i/\pi_a^{rn_a}$ given $\mathcal{D}_{a-1}(\hat{y}_j)$. This is just $\delta_a(\hat{y}_j, r)$ rearranged by the correspondence from Lemma 4.5. By $E(\hat{y}_j|\hat{y})$ we mean

$$E_{\hat{\mu}}[S_0^{n_a}T_0^{-m_a}\hat{y}_j|\bigvee_{i=0}^{\infty}T^{-i}(P)](\hat{y}).$$

LEMMA 5.2:

$$\delta_a(\hat{y}, r) = m_{ar}^*\delta_a(\hat{y}_j, r) \quad \text{for } j = 0 \text{ to } (q'-1).$$

Proof: We have $x_{ar} + i/\pi_a^{rn_a}$ at $S^{k_{ar}n_a}T_0^{-r-k_{ar}m_a}(x)$ iff

$$T_0^{m_a}S_0^{-n_a}(x) = \frac{p^{m_a}x}{q^{n_a}} + \frac{p'j}{q'}$$

and $S_0^{(k_{ar}+1)n_a}T_0^{-r-(k_{ar}+1)m_a}$ brings this to $x_{ar} + i/\pi_a^{rn_a}$. Write this as

$$\delta_a(\hat{y}, r) = \sum_{j=0}^{q'} E(\hat{y}_j|\hat{y})\; m_{ar}^*\delta_a(\hat{y}_j, r).$$

Then

$$h[\delta_a(\hat{y}, r)] \geq \sum_{j=0}^{q'} E(\hat{y}_j|\hat{y})h[m_{ar}^*\delta_a(\hat{y}_j, r)]$$

where the inequality is equality iff $m_{ar}^*\delta_a(\hat{y}_j, r)$ is the same for every j and thus is the same as $\delta_a(\hat{y}, r)$. But

$$h[\delta_a(\hat{y}_j, r)] = h[m_{ar}^*\delta_a(\hat{y}_j, r)] \leq h[\delta_a(\hat{y}, r)]$$

because the left-hand ones are conditioned on more, and

$$\int h[\delta_a(\hat{y}_j, r)]d\hat{\mu} = \int h[\delta_a(\hat{y}, r)]d\hat{\mu}$$

using the $T^{\pm 1}$ and $S^{\pm 1}$ invariance of $\hat{\mu}$. Thus $h[\delta_a(\hat{y}_j, r) \times q] = h[\delta_a(\hat{y}, r)]$ and we get the result. ∎

COROLLARY 5.3:

$$\delta_a(\hat{y}, r)[\frac{i}{\pi_a^{r_a}}] = \delta_a(S^{n_a}T^{-m_a}\hat{y}, r)[\frac{m_{ar}^* i}{\pi_a^{rn_a}}].$$

Proof: On the left is the expected value of $x_{ar} + i/\pi_a^{rn_a}$, given $\mathcal{D}_{a-1}(\hat{y})$. By Lemma 5.2 this is the same as the expected value of $S_0^{n_a}T_0^{-m_a}(x_{ar} + i/\pi_a^{rn_a})$ given $\mathcal{D}_{a-1}(S^{n_a}T^{-m_a}\hat{y})$. But

$$S_0^{n_a}T_0^{-m_a}x_{ar} = \varphi(S^{(k_{ar}+1)n_a}T^{-r-(k_{ar}+1)m_a}\hat{y}) \quad \text{and} \quad S_0^{n_a}T_0^{-m_a}(\frac{i}{\pi_a^{rn_a}}) = \frac{m_{ar}^* i}{\pi_a^{rn_a}}$$

as specified. ∎

This corollary says that the expectations for the preimages associated to $(-r-k_{ar}m_a, k_{ar}n_a)$ are independent of the possible preimages at $(m_a, -n_a)$.

LEMMA 5.4:
$$\delta_a(\hat{y}, r-1)\left[\frac{i}{\pi_a^{(r-1)n_a}}\right] = \Sigma\delta_a(\hat{y}, r)\left[\frac{j}{\pi_a^{rn_a}}\right]$$
where the sum is over all j such that
$$p_0\prod_{i\neq a}\pi_i^{n_i}j = i\times(m^*_{a(r-1)})^{k_{ar}-k_{a(r-1)}} \mod \pi_a^{(r-1)n_a}.$$

Proof: The expectation of $x_{ar}+j/\pi_a^{rn_a}$ equals the expectation of $x_{a(r-1)}+i/\pi_a^{(r-1)n_a}$ multiplied by the expectation of $x_{ar}+j/\pi_a^{rn_a}$ given $x_{a(r-1)}+i/\pi_a^{(r-1)n_a}$. Write this as

$$\delta_a(\hat{y}, r)\left[\frac{j}{\pi_a^{rn_a}}\right] =$$
$$\delta_a(\hat{y}, r-1)\left[\frac{i}{\pi_a^{(r-1)n_a}}\right]$$
$$\times E_\mu\left[S^{n_a}T^{-m_a}\varphi^{-1}\left(x_{ar}+\frac{j}{\pi_a^{rn_a}}\right)\middle|\varphi^{-1}\left(x_{a(r-1)}+\frac{i}{\pi_a^{(r-1)n_a}}\right)\vee\mathcal{D}_{a-1}(\hat{y})\right].$$

This is only for i and j such that if we move $x_{a(r-1)}+i/\pi_a^{(r-1)n_a}$ up the staircase $k_{ar}-k_{a(r-1)}$ times it equals $x_{ar}+j/\pi_a^{rn_a}$ moved horizontally to the right one step. In other words, i and j such that

$$px_{ar}+\frac{pj}{\pi_a^{rn_a}} = S_0^{n_a}T_0^{-m_a}(x_{a(r-1)})+\frac{(m^*_{a(r-1)})^{k_{ar}-k_{a(r-1)}}i}{\pi_a^{(r-1)n_a}}$$
$$\Leftrightarrow \frac{p_0\prod_{i\neq a}\pi_i^{n_i}j}{\pi_a^{(r-1)n_a}} = \frac{(m^*_{a(r-1)})^{k_{ar}-k_{a(r-1)}}i}{\pi_a^{(r-1)n_a}}$$

which gives us the j as specified. Fix i, sum over all such j. This gives the result. ∎

COROLLARY 5.5: *$\delta_a(T^r\hat{y}, 2r)$ determines $\delta_a(T^i\hat{y}, r+i)$ for $0\leq i\leq r$.*

Proof: By Corollary 5.3, $\delta_a(T^r\hat{y}, 2r)$ tells us $\delta_a(S^{-k_{a(r-i)}n_a}T^{k_{a(r-1)}m_a}T^r\hat{y}, 2r)$. Call this $\delta_a(\hat{u}, 2r)$. Let $w_{2r}=\varphi(S^{k_{a(2r)}n_a}T^{-2r-k_{a(2r)}m_a}\hat{u})$; this is used to compute $\delta_a(\hat{u}, 2r)$. Let

$$w_{r-i}=\varphi(S^{k_{a(r-i)}n_a}T^{-(r-i)-k_{a(r-i)}m_a}\hat{u})=\varphi(T^i\hat{y}).$$

By Definition 5.1,

$$\delta_a(\hat{u},2r)\left[\frac{j}{\pi_a^{n_a 2r}}\right] = E_{\hat{\mu}}\left[S^{-k_{a(2r)}n_a}T^{2r+k_{a(2r)}m_a}\varphi^{-1}\left(w_{2r}+\frac{j}{\pi_a^{n_a 2r}}\right)\middle|\mathcal{D}_{a-1}\right](\hat{u}).$$

Write as

$$E\left[w_{2r}+\frac{j}{\pi_a^{n_a 2r}}\middle|\mathcal{D}_{a-1}\right](\hat{u}).$$

If $w_{2r}+j/\pi_a^{n_a 2r}$ is consistent with w_{r-i} then the expectation of $w_{2r}+j/\pi_a^{n_a 2r}$ given $\mathcal{D}_{a-1}(\hat{u})$ equals the expectation of $w_{2r}+j/\pi_a^{n_a 2r}$ given w_{r-i} and $\mathcal{D}_{a-1}(\hat{u})$ multiplied by the expectation of w_{r-i} given $\mathcal{D}_{a-1}(\hat{u})$. Write this as

$$E\left[w_{2r}+\frac{j}{\pi_a^{n_a 2r}}\middle|\mathcal{D}_{a-1}\right](\hat{u}) = E\left[w_{2r}+\frac{j}{\pi_a^{n_a 2r}}\middle|w_{r-i}\right]\times E\left[w_{r-i}|\mathcal{D}_{a-1}\right](\hat{u}).$$

For the two to be consistent means i and j satisfy

$$\begin{aligned} p^{r+i}(w_{2r}+\frac{j}{\pi_a^{n_a 2r}}) &= \frac{q^{(k_{a(2r)}-k_{a(r-i)})n_a}}{p^{(k_{a(2r)}-k_{a(r-i)})m_a}}w_{r-i} \\ \Leftrightarrow \frac{p'j}{\pi_a^{n_a(r-i)}} &= 0 \mod 1 \\ \Leftrightarrow j = k\pi_a^{n_a(r-i)} &\quad \text{for } k=0 \text{ to } (\pi_a^{n_a(r+i)}-1). \end{aligned}$$

Thus

$$E\left[w_{2r}+\frac{j}{\pi_a^{n_a 2r}}\middle|w_{r-i}\right] \quad \text{can be written} \quad E\left[w_{2r}+\frac{k}{\pi_a^{n_a(r+i)}}\middle|w_{r-i}\right].$$

This differs from $\delta_a(T^i\hat{y},r+i)\left[k/\pi_a^{n_a(r+i)}\right]$ only by the locations these expectations are associated with. The first is at $(-r-[k_{a(2r)}-k_{a(r-i)}]m_a, [k_{a(2r)}-k_{a(r-i)}]n_a)$ and the second at $(-r-k_{a(r+i)}m_a, k_{a(r+i)}n_a)$. By Corollary 5.3 one is just a permutation of the other so in fact we can write

$$E\left[w_{2r}+\frac{j}{\pi_a^{n_a 2r}}\middle|\mathcal{D}_{a-1}\right](\hat{u}) = \delta_a(T^i\hat{y},r+i)\left[\frac{n^*k}{\pi_a^{n_a(r+i)}}\right]\times E\left[w_{r-i}|\mathcal{D}_{a-1}\right](\hat{u}).$$

Rewrite again as

$$\delta_a(\hat{u},2r)\left[\frac{k\pi_a^{n_a(r-i)}}{\pi_a^{n_a 2r}}\right] = \delta_a(T^i\hat{y},r+i)\left[\frac{n^*k}{\pi_a^{n_a(r+i)}}\right]\times\delta_a(\hat{u},r-i)[0].$$

We know the first and last terms by Lemma 5.4, and thus can find the term in the middle. ∎

LEMMA 5.6: *If $\hat{y}_1$ and $\hat{y}_2$ have $\mathcal{D}_{a-1}(\hat{y}_1) = \mathcal{D}_{a-1}(\hat{y}_2)$ then $\delta_a(\hat{y}_1, r)$ and $\delta_a(\hat{y}_2, r)$ differ by a translation (mod 1) of*

$$\varphi(S^{k_{ar}n_a}T^{-r-k_{ar}m_a}\hat{y}_2) - \varphi(S^{k_{ar}n_a}T^{-r-k_{ar}m_a}\hat{y}_2).$$

Proof: In Definition 5.1 the only difference would be in the ordering, which only depends on x_{ar} for both. ∎

6. Symmetric Points

Definition 6.1: A point $\hat{y} \in \hat{Y}$ is δ_0-symmetric if there exist two points $\hat{y}_1$ and $\hat{y}_2$ such that

(i) $\varphi(\hat{y}_1) = \varphi(\hat{y}_2) = \varphi(\hat{y})$.
(ii) There exists m such that $\hat{y}_1(-m, k_{0m}) \neq \hat{y}_2(-m, k_{0m})$.
(iii) For every n, m, $\delta_0(T^m\hat{y}_1, n) = \delta_0(T^m\hat{y}_2, n)$.

Definition 6.2: A point $\hat{y} \in \hat{Y}$ is δ_a-symmetric if there exist two points $\hat{y}_1$ and $\hat{y}_2$ such that:

(i) $\mathcal{D}_{a-1}(\hat{y}_1) = \mathcal{D}_{a-1}(\hat{y}_2) = \mathcal{D}_{a-1}(\hat{y})$.
(ii) There exists s such that

$$\hat{y}_1(-s - k_{as}m_a, k_{as}n_a) \neq \hat{y}_2(-s - k_{as}m_a, k_{as}n_a).$$

(iii) For every i, j, r, $\delta_a(T^iS^j\hat{y}_1, r) = \delta_a(T^iS^j\hat{y}_2, r)$.

We want to show in this section that the set of each type of symmetric point is T and S invariant.

LEMMA 6.3: *The set of δ_0-symmetric points is both T and S invariant. Since $\hat{\mu}$ is ergodic, this set has measure 0 or 1.*

Proof: Say $\hat{y}$ is δ_0-symmetric and let $\hat{y}_1$ and $\hat{y}_2$ be as specified.

T-Invariance

We want to show that $T\hat{y}_1$ and $T\hat{y}_2$ satisfy the definition. Note that $\hat{y}(-i, j) = T\hat{y}(-i-1, j)$.

(i) Obviously $\varphi(T\hat{y}_2) = \varphi(T\hat{y}_1) = \varphi(T\hat{y})$.
(ii) Since $\hat{y}_1(-i+1, k_{0i}) = \hat{y}_2(-i+1, k_{0i})$ for $i \le m$,

$$\varphi(S^{k_{0\,(i-1)}}T^{i+1}\hat{y}_1) = \varphi(S^{k_{0\,(i-1)}}T^{-i+1}\hat{y}_2).$$

Then the symbol at $(-i+1, k_{0i})$ is also the same since all cancellation is done here after $k_{0\,(i-1)}$ steps. So

$$T\hat{y}_1(-i, k_{0i}) = \hat{y}_1(-i+1, k_{0i}) = \hat{y}_2(-i+1, k_{0i}) = T\hat{y}_2(-i, k_{0i}) \quad \text{for} \quad i \le m.$$

Thus

$$\varphi(S^{k_{0m}}T^{-m+1}\hat{y}_1) = \varphi(S^{k_{0m}}T^{-m+1}\hat{y}_2)$$

and yet $\hat{y}_1(-m, k_{0m}) \neq \hat{y}_2(-m, k_{0m})$, showing that $\hat{y}_1$ and $\hat{y}_2$ correspond to different preimages and still will when moved up to $(-m, k_{0\,(m+1)})$. Then

$$\begin{aligned} T\hat{y}_1(-m-1, k_{0\,(m+1)}) &= \hat{y}_1(-m, k_{0(m+1)}) \\ &\neq \hat{y}_2(-m, k_{0(m+1)}) = T\hat{y}_2(-m-1, k_{0(m+1)}). \end{aligned}$$

(iii) $\delta_0(T^mT\hat{y}_1, n) = \delta_0(T^{m+1}\hat{y}_1, n) = \delta_0(T^{m+1}\hat{y}_2, n)$ by assumption, which equals $\delta_0(T^mT\hat{y}_2, n)$ as needed.

S-Invariance

We want to show that $S\hat{y}_1$ and $S\hat{y}_2$ satisfy the definition.

(i) Obviously $\varphi(S\hat{y}_1) = \varphi(S\hat{y}_2) = \varphi(S\hat{y})$.

(ii) $\hat{y}_1(-i, k_{0i}) = \hat{y}_2(-i, k_{0i})$ for $i < m$ says that $\varphi(S^{k_{0i}}T^{-i}\hat{y}_1) = \varphi(S^{k_{0i}}T^{-i}\hat{y}_2)$. All cancellation is done after moving up k_{0i} steps thus

$$\begin{aligned} \varphi(S^{k_{0i}+1}T^{-i}\hat{y}_1) = \varphi(S^{k_{0i}+1}T^{-i}\hat{y}_2) &\Rightarrow \hat{y}_1(-i, k_{0i}+1) = \hat{y}_2(-i, k_{0i}+1) \\ &\Rightarrow S\hat{y}_1(-i, k_{0i}) = S\hat{y}_2(-i, k_{0i}). \end{aligned}$$

But then $\hat{y}_1(-m, k_{0m}) \neq \hat{y}_2(-m, k_{0m})$ so

$$\begin{aligned} \varphi(S^{k_{0m}}T^{-m}\hat{y}_1) \neq \varphi(S^{k_{0m}}T^{-m}\hat{y}_2) &\Rightarrow \varphi(S^{k_{0m}+1}T^{-m}\hat{y}_1) \neq \varphi(S^{k_{0m}+1}T^{-m}\hat{y}_2) \\ &\Rightarrow \hat{y}_1(-m, k_{0m}+1) \neq \hat{y}_2(-m, k_{0m}+1) \\ &\Rightarrow S\hat{y}_1(-m, k_{0m}) \neq S\hat{y}_2(-m, k_{0m}). \end{aligned}$$

(iii)

$$\begin{aligned} \delta_0(T^mS\hat{y}_1, n) &= \delta_0(T^m\hat{y}_1, n) \quad \text{shifted by } q \\ &= \delta_0(T^m\hat{y}_2, n) \quad \text{shifted by } q \\ &= \delta_0(T^mS\hat{y}_2, n). \qquad \blacksquare \end{aligned}$$

LEMMA 6.4: *The set of δ_a-symmetric points, $a = 1$ to h, is both T and S invariant. Since $\hat{\mu}$ is ergodic, this set has measure 0 or 1.*

Proof: Recall that $k_{hi} = 0$ for all i, by Definition 4.1. Let $\hat{y}$ be a symmetric point and $\hat{y}_1, \hat{y}_2$ as specified.

T-Invariance

(i) Obviously $\mathcal{D}_{a-1}(T\hat{y}_1) = \mathcal{D}_{a-1}(T\hat{y}_2) = \mathcal{D}_{a-1}(T\hat{y})$.

(ii) We want to show that

$$T\hat{y}_1(-(s+1)-k_{a(s+1)}m_a, k_{a(s+1)}n_a) \neq T\hat{y}_2(-(s+1)-k_{a(s+1)}m_a, k_{a(s+1)}n_a)$$

which is equivalent to showing

$$\hat{y}_1(-s - k_{a(s+1)}m_a, k_{a(s+1)}n_a) \neq \hat{y}_2(-s - k_{a(s+1)}m_a, k_{a(s+1)}n_a).$$

For $a = h$ this will be true by assumption. For the rest, recall that

$$\hat{y}_1(-s - k_{as}m_a, k_{as}n_a) \neq \hat{y}_2(-s - k_{as}m_a, k_{as}n_a)$$

and this is the first such s. Given $\mathcal{D}_{a-1}(\hat{y})$ there are $\pi_a^{sn_a}$ possible preimages associated with $(-s - k_{as}m_a, k_{as}n_a)$ and further application of $S_0^{n_a}T_0^{-m_a}$ is 1-1. $\hat{y}_1$ and $\hat{y}_2$ correspond to different preimages and still do after we apply

$$S^{(k_{a(s+1)}-k_{as})n_a}T_0^{-(k_{a(s+1)}-k_{as})m_a}.$$

Thus they have different symbols at $(-s - k_{a(s+1)}m_a, k_{a(s+1)}n_a)$.

(iii)
$$\begin{aligned}\delta_a(T^iS^jT\hat{y}_1, r) &= \delta_a(T^{i+1}S^j\hat{y}_1, r)\\ &= \delta_a(T^{i+1}S^j\hat{y}_2, r) \quad \text{by assumption}\\ &= \delta_a(T^iS^jT\hat{y}_2, r).\end{aligned}$$

S-Invariance

(i) Obviously $\mathcal{D}_{a-1}(S\hat{y}_1) = \mathcal{D}_{a-1}(S\hat{y}_2) = \mathcal{D}_{a-1}(S\hat{y})$.

(ii) Let b be the first integer such that $m_a \leq bn_a$. Let s be the first integer such that $\hat{y}_1(-s - k_{as}m_a, k_{as}n_a) \neq \hat{y}_2(-s - k_{as}m_a, k_{as}n_a)$. Then we also get

$$\hat{y}_1(-s - k_{a(s+b)}m_a, k_{a(s+b)}n_a) \neq \hat{y}_2(-s - k_{a(s+b)}m_a, k_{a(s+b)}n_a)$$

and so of the possible preimages associated with $(-s - k_{a(s+b)}m_a, k_{a(s+b)}n_a)$, $\hat{y}_1$ must correspond to $\bar{x} + c_1/\pi_a^{sn_a}$ and $\hat{y}_2$ to $\bar{x} + c_2/\pi_a^{sn_a}$. But then the possibilities at $(-s - b - k_{a(s+b)}m_a, k_{a(s+b)}n_a + 1)$ are $q\bar{x}/p^b + qc_i/p^b\pi_a^{sn_a} + qj/p^b$ which we can write

$$\frac{q\bar{x}}{p^b} + \frac{q'c_i}{p'\pi_a^{n_ab-m_a}\pi_a^{sn_a}} + \frac{q'j}{p'\pi_a^{n_ab-m_a}},$$

where p', q', and π_a are all relatively prime. But for $c_1 \neq c_2$ these are disjoint sets, since they are translations of the set

$$\left\{ \frac{q\bar{x}}{p^b} + \frac{q'j}{p'\pi_a^{n_a b - m_a}} \right\}$$

by different amounts. Thus

$$\varphi(S^{k_{a(s+b)}n_a+1}T^{-s-b-k_{a(s+b)}m_a}\hat{y}_1) \neq \varphi(S^{k_{a(s+b)}n_a+1}T^{-s-b-k_{a(s+b)}m_a}\hat{y}_2)$$

$$\Leftrightarrow \varphi(S^{k_{a(s+b)}n_a}T^{-s-b-k_{a(s+b)}m_a}S\hat{y}_1) \neq \varphi(S^{k_{a(s+b)}n_a}T^{-s-b-k_{a(s+b)}m_a}S\hat{y}_2).$$

Now to show there exists $t \leq s+b$ such that

$$S\hat{y}_1(-t - k_{at}m_a, k_{at}n_a) \neq S\hat{y}_2(-t - k_{at}m_a, k_{at}n_a).$$

If we can find $t < s+b$ then we are done. Assume not; we will show that $t = s+b$ works. Since further movement is 1–1, the assumption also says that

$$S\hat{y}_1(-t - k_{a(s+b)}m_a, k_{a(s+b)}n_a) = S\hat{y}_2(-t - k_{a(s+b)}m_a, k_{a(s+b)}n_a) \quad \text{for } t < s+b.$$

Using that $\mathcal{D}_{a-1}(S\hat{y}_1) = \mathcal{D}_{a-1}(S\hat{y}_2)$, it is easy to see that

$$\varphi(S^{k_{a(s+b)}n_a}T^{-k_{a(s+b)}m_a}S\hat{y}_1) = \varphi(S^{k_{a(s+b)}n_a}T^{-k_{a(s+b)}m_a}S\hat{y}_2).$$

So $\hat{y}_1$ and $\hat{y}_2$ have the same first quadrant at $(-k_{a(s+b)}m_a, k_{a(s+b)}n_a)$ and the same symbols to the left up to $s+b$. There we know they correspond to different preimages so must have different symbols.

(iii)
$$\begin{aligned}\delta_a(T^iS^jS\hat{y}_1, r) &= \delta_a(T^iS^{j+1}\hat{y}_1, r)\\ &= \delta_a(T^iS^{j+1}\hat{y}_2, r) \quad \text{by assumption}\\ &= \delta_a(T^iS^jS\hat{y}_2, r). \quad \blacksquare\end{aligned}$$

LEMMA 6.5: *Let $B = \bigcup_{i=0}^{h}\{\hat{y}; \hat{y}$ is a δ_i-symmetric point $\}$. Then $\hat{\mu}(B) = 0$ or 1.*

Proof: B is just the union of T and S invariant sets. ∎

7. When the Set of Symmetric Points has Measure 1

THEOREM 7.1: *If $\hat{y}$ is a δ_0-symmetric point then the group of translations under which $\delta_0(\hat{y}, n)$ is invariant contains the group $< b_n >$ where b_n is a fraction whose denominator in least terms diverges to infinity.*

Proof: For $\hat{y}$ a δ_0-symmetric point, there exists $\hat{y}_1, \hat{y}_2$ such that at some first index m, $\hat{y}_1(-m, k_{0m}) \neq \hat{y}_2(-m, k_{0m})$ yet $\delta_0(\hat{y}_1, m) = \delta_0(\hat{y}_2, m)$. Thus $\delta_0(\hat{y}, m)$ is invariant under translation by $\varphi(T^{-m}S^{k_{0m}}\hat{y}_1) - \varphi(T^{-m}S^{k_{0m}}\hat{y}_2)$ by Lemma 3.6. Since the possibilities for $\varphi(T^{-m}S^{k_{0m}}\hat{y})$ are $\{\bar{x} + i/p_0^m\}_{i=0}^{p_0^m - 1}$, $\delta_0(\hat{y}, m)$ is invariant under translation by $(i_2 - i_1)/p_0^m$. Notice that in least terms, the denominator is ≥ 2.

We will prove the theorem by induction: if $\delta_0(\hat{y}, j)$ is invariant under translation by u/v with v nontrivial, then $\delta_0(\hat{y}, j+1)$ is invariant under translation by a fraction whose denominator, in least terms, is at least $2v$. The first paragraph gives us our beginning step.

The possible preimages for $\varphi(T^{-j}S^{k_{0j}}\hat{y})$ are $\{x_{0j} + i/p_0^j\}_{i=0}^{p_0^j - 1}$. For $j \geq m$, we have $\hat{y}_1$ associated to $x_{0j} + i_1/p_0^j$ and $\hat{y}_2$ to $x_{0j} + i_2/p_0^j, i_1 \neq i_2$. Thus $\delta_0(\hat{y}, j)$ is invariant under translation by $(i_2 - i_1)/p_0^j$ which we are assuming has the form u/v. We also know $\delta_0(\hat{y}_1, j+1) = \delta_0(\hat{y}_2, j+1)$ so $\delta_0(\hat{y}, j+1)$ is invariant under translation by $\varphi(T^{-j-1}S^{k_{0\,(j+1)}}\hat{y}_2) - \varphi(T^{-j-1}S^{k_{0\,(j+1)}}\hat{y}_1)$.

The possibilities for $\varphi(T^{-j-1}S^{k_{0\,(j+1)}}\hat{y}_2)$ are a subset of

$$\left\{x_{0\,(j+1)} + \frac{i}{p_0^{j+1}}\right\}_{i=0}^{p_0^{j+1}-1} = \left\{x_{0\,(j+1)} + \frac{i'}{p_0^{j+1}} + \frac{k}{p_0}\right\}_{i'=0}^{p_0^j - 1} {}_{k=0}^{p_0 - 1}$$

such that

$$p \times \left(x_{0\,(j+1)} + \frac{i'}{p_0^{j+1}} + \frac{k}{p_0}\right) = q^{k_{0\,(j+1)} - k_{0j}}\left(x_{0j} + \frac{i_2}{p_0^j}\right).$$

But the left side is exactly

$$q^{k_{0(j+1)} - k_{0j}} x_{0j} + \frac{\pi_1^{n_1} \cdots \pi^{n_h} i'}{p_0^j} \quad \text{mod } 1,$$

so the subset we want includes only i' such that

$$\pi_1^{n_1} \cdots \pi_h^{n_h} i' = q^{k_{0(j+1)} - k_{0j}} i_2 \quad \text{mod } p_0^j.$$

Since all the coefficients are relatively prime to p_0^j, such a unique i' exists; call it i_2'. Similarly the possibilities for

$$\varphi(T^{-j-1}S^{k_{0(j+1)}}\hat{y}_1) = \left\{x_{0(j+1)} + \frac{i_1'}{p_0^{j+1}} + \frac{k}{k_0}\right\}_{k=0}^{p_0-1}.$$

Thus $\delta_0(\hat{y}, j+1)$ is invariant under translation by a number of the form $(i_2' - i_1')/p_0^{j+1} + h/p_0$ where $0 \le h < p_0$.

Recall that $(i_2 - i_1)/p_0^j = u/v$ in least terms. Thus

$$\frac{q^{k_{0(j+1)}-k_{0j}}(i_2 - i_1)}{p_0^j} = \frac{q^{k_{0(j+1)}-k_{0j}}u}{v}.$$

But by the last paragraph the left side of this equation equals

$$\pi_1^{n_1}\cdots\pi_h^{n_h}(i_2' - i_1')/p_0^j$$

and thus

$$\frac{i_2' - i_1'}{p_0^j} = \frac{q^{k_{0(j+1)}-k_{0j}}u}{\pi_1^{n_1}\cdots\pi_h^{n_h}v}.$$

Denote this by $q'u/\pi'v$. We can now write the above translation as a number of the form

$$\frac{q'u}{\pi'vp_0} + \frac{h}{p_0} = \frac{q'u + \pi'vh}{\pi'vp_0}.$$

This may not be in least terms anymore; we will show that the reduced form has at least $2v$ in its denominator.

First consider v. We know v is relatively prime to u and since v is made up of components of p_0 it must also be relatively prime to q'. Thus v is relatively prime to $q'u + \pi'vh$.

Next consider p_0. There may be some part of p_0 that divides u so write $p_0 = p_1p_2$ where p_2 is relatively prime to u. Since v is made from p_0's and v is nontrivial, p_2 is nontrivial. But then there is some part of p_2 that divides v, call it $\bar{p}_2$. Once again it must be nontrivial since v is. Then $\bar{p}_2$ is relatively prime to $q'u$ and divides $\pi'vh$. Thus it is relatively prime to $q'u + \pi'vh$. So in least terms the denominator must be at least $\bar{p}_2v \ge 2v$. ∎

THEOREM 7.2: *If $\hat{y}$ is a δ_a-symmetric point (for some $a = 1, \ldots, h$) then the group of translations under which $\delta_a(\hat{y}, r)$ is invariant contains the group $< b_n >$ where b_n is a fraction whose denominator in least terms diverges to infinity.*

Proof: This is done similarly to Theorem 7.1 except now $\hat{y}_d$ will correspond to $x_{ar} + i_d/\pi_a^{tn_a}$ at the first step and to

$$\left\{ x_{a(t+1)} + \frac{i'_d}{\pi_a^{(t+1)n_a}} + \frac{k}{\pi_a^{n_a}} \right\}_{k=0}^{\pi^{n_a}-1_a}$$

at the next, where

$$p_0 \prod_{j \neq a} \pi_j i'_d = (m^*_{at})^{k_{a(t+1)} - k_{at}} i_d \bmod \pi_a^{tn_a}.$$

Again assume that $\delta_a(\hat{y}, t)$ is invariant under $(i_2 - i_1)/\pi_a^{tn_a}$ which equals u/v in least terms. Then $\delta_a(\hat{y}, t+1)$ is invariant under translation by a number of the form

$$\frac{i'_2 - i'_1}{\pi_a^{(t+1)n_a}} + \frac{h}{\pi_a^{n_a}} \quad \text{where } 0 \leq h < \pi_a^{n_a}.$$

But

$$\frac{p_0 \prod_{j \neq a} \pi_j^{n_j} (i'_2 - i'_1)}{\pi_a^{tn_a}} = \frac{(m^*_{at})^{k_{a(t+1)} - k_{at}} (i_2 - i_1)}{\pi_a^{tn_a}} = \frac{(m^*_{at})^{k_{a(t+1)} - k_{at}} u}{v}$$

and so

$$\frac{i'_2 - i'_1}{\pi_a^{tn_a}} = \frac{(m^*_{at})^{k_{a(t+1)} - k_{at}} u}{p_0 \prod_{j \neq a} \pi_j^{n_j} v}.$$

Denote this by $m'u/p'v$. Thus $\delta_a(\hat{y}, t+1)$ is invariant under translation by a number of the form

$$\frac{m'u}{p'v\pi_a^{n_a}} + \frac{h}{\pi_a^{n_a}} = \frac{m'u + p'vh}{p'v\pi_a^{n_a}}.$$

As in 7.1, we can show that, in least terms, v and a nontrivial part of $\pi_a^{n_a}$ will be in the denominator. ∎

LEMMA 7.3: *If $\hat{y}$ is a δ_a-symmetric point for some $a = 0$ to h, then $\delta_a(\hat{y}, n)$ converges weakly to Lebesgue measure on $[0, 1)$.*

Proof: Let $\mathcal{A}_n$ be the set of rotations under which $\delta_a(\hat{y}, n)$ is invariant. By Theorems 7.1 and 7.2 we know $\mathcal{A}_n \supset < b_n >$ where b_n is a fraction whose denominator, in least terms, diverges to infinity. For any $\alpha \in \mathcal{A}_n$, any $f \in C(S^1)$, $\int f\, d\delta_a(\hat{y},n) = \int R_\alpha(f) d\delta_a(\hat{y},n)$, where R_α is rotation by α. We can write

$$\int f d\delta_a(\hat{y},n) = \frac{1}{|\mathcal{A}_n|} \sum_{\alpha\in\mathcal{A}_n} \int R_\alpha(f) d\delta_a(\hat{y},n).$$

But $|\mathcal{A}_n|$, the cardinality of $\mathcal{A}_n$, diverges to infinity, so

$$\frac{1}{|\mathcal{A}_n|} \sum_{\alpha\in\mathcal{A}_n} R_\alpha(f) = \int f\, dm$$

by definition of Riemann Integral. Thus $\int f d\delta_a(\hat{y},n) \to \int f dm$. ∎

THEOREM 7.4: *If a.e. point is δ_a-symmetric for some a between 0 and h, then $\mu = m$ is Lebesgue measure.*

Proof: Using the definition of δ_0 and δ_a and translating back, we see that $R_{\varphi(S^{k_{0r}}T^{-r}\hat{y})}\delta_0(\hat{y},r)$ puts weights on

$$\{x_{0r}, x_{0r} + \frac{1}{p_0^r}, \ldots, x_{0r} + \frac{p_0^r - 1}{p_0^r}\}$$

and $R_{\varphi(S^{k_{ar}n_a}T^{-r-k_{ar}m_a}\hat{y})}\delta_a(\hat{y},r)$ puts weights on

$$\{x_{ar}, x_{ar} + \frac{1}{\pi_a^{rn_a}}, \ldots, x_{ar} + \frac{\pi_a^{rn_a} - 1}{\pi_a^{rn_a}}\}.$$

Generalize to any set C by

$$R_{\varphi(S^{k_{0r}}T^{-r}\hat{y})}\delta_0(\hat{y},r)[C] = E_{\hat{\mu}}\left[S^{-k_{0r}}T^r\varphi^{-1}(C)|\bigvee_{i=0}^{\infty}T^{-i}(P)\right](\hat{y})$$

and

$$R_{\varphi(S^{k_{ar}n_a}T^{-r-k_{ar}m_a}\hat{y})}\delta_a(\hat{y},r)[C] = E_{\hat{\mu}}\left[S^{-k_{ar}n_a}T^{r+k_{ar}m_a}\varphi^{-1}(C)|\mathcal{D}_{a-1}\right](\hat{y}).$$

Let $B_a = \{\hat{y} : \hat{y}$ a δ_a-symmetric point $\}$. We are assuming that

$$\mu\left(\bigcup_{a=0}^{h} B_a\right) = 1,$$

thus there is at least one B_a with full measure. Then $\hat{Y} = B_a$ a.e. If $a = 0$,

$$\mu(C) = \int_{B_0} 1_{\varphi^{-1}(C)} d\hat{\mu} = \int_{B_0} E_{\hat{\mu}}\left[\varphi^{-1}(C)|S^{k_{0r}}\bigvee_{i=r}^{\infty}T^{-i}(P)\right] d\hat{\mu}$$

$$= \int_{B_0} E_{\hat{\mu}} \left[S^{-k_{0r}} T^r \varphi^{-1}(C) | \bigvee_{i=0}^{\infty} T^{-i}(P) \right] d\hat{\mu} = \int_{B_0} R_{\varphi(S^{k_{0r}} T^{-r} \hat{y})} \delta_0(\hat{y}, r)[C] d\hat{\mu}.$$

By weak convergence of $\delta_0(\hat{y}, r)$ to Lebesgue measure m, this has the limit

$$\int_{B_0} m(C) d\hat{\mu} = m(C)\hat{\mu}(B_0) = m(C) \times 1.$$

If $a = 1, \ldots, h$,

$$\mu(C) = \int_{B_a} 1_{\varphi^{-1}(C)} d\hat{\mu} = \int_{B_a} E_{\hat{\mu}} \left[\varphi^{-1}(C) | S^{k_{ar} n_a} T^{-r-k_{ar} m_a} \mathcal{D}_{a-1} \right] d\hat{\mu}$$

$$= \int_{B_a} E_{\hat{\mu}} \left[S^{-k_{ar} n_a} T^{r+k_{ar} m_a} \varphi^{-1}(C) | \mathcal{D}_{a-1} \right] d\hat{\mu}$$

$$= \int_{B_a} R_{\varphi(S^{k_{ar} n_a} T^{-r-k_{ar} m_a} \hat{y})} \delta_a(\hat{y}, r)[C] d\hat{\mu}.$$

By weak convergence of $\delta_a(\hat{y}, r)$ to Lebesgue measure m, this has the limit

$$\int_{B_a} m(C) d\hat{\mu} = m(C)\hat{\mu}(B_a) = m(C) \times 1. \qquad \blacksquare$$

8. When the Set of Symmetric Points has Measure 0

Let $\mathcal{H}_a$ be the minimal T and S invariant σ-algebra for which the functions $\delta_a(\hat{y}, n)$ are measurable. $\mathcal{H}_0$ is trivially S-invariant because $\delta_0(\hat{y}, n)$ determines $\delta_0(S^k \hat{y}, n)$ for all $k \in \mathbb{N}$. Let $\mathcal{H}_{ar}$ be the minimal S-invariant σ-algebra for which $\delta_a(T^r \hat{y}, 2r)$ is measurable. By Corollaries 3.5 and 5.5 the $\mathcal{H}_{ar}$ are nested and refine to $\mathcal{H}_a$.

LEMMA 8.1: *The action of S on $\mathcal{H}_{0r}$ is periodic. Thus $h(S, \mathcal{H}_{0r}) = 0$ and using the refinement, $h(S, \mathcal{H}_0) = 0$.*

Proof: This follows from Corollary 3.3. Note that the period depends on r. $\blacksquare$

LEMMA 8.2: *$h_{\hat{\mu}}(S^{n_a}, \mathcal{H}_{ar}) = h_{\hat{\mu}}(T^{m_a}, \mathcal{H}_{ar})$ and thus*

$$h_{\hat{\mu}}(S^{n_a}, \mathcal{H}_a) = h_{\hat{\mu}}(T^{m_a}, \mathcal{H}_a)$$

from refinement. So

$$h_{\hat{\mu}}(T,\mathcal{H}_a)=\frac{n_a}{m_a}h_{\hat{\mu}}(S,\mathcal{H}_a).$$

Proof: By the construction of $\delta_a(T^r\hat{y},2r)$ we can find an integer d such that $S^{dn_a}T^{-dm_a}$ acts as the identity on $\mathcal{H}_{ar}$. Thus

$$h_{\hat{\mu}}(S^{dn_a},\mathcal{H}_{ar})=h_{\hat{\mu}}(T^{dm_a},\mathcal{H}_{ar})\Rightarrow dh_{\hat{\mu}}(S^{n_a}\mathcal{H}_{ar})=dh_{\hat{\mu}}(T^{m_a},\mathcal{H}_{ar}). \quad \blacksquare$$

LEMMA 8.3:

$$h_{\hat{\mu}}(T,\mathcal{H}_a)=\frac{\log p}{\log q}h_{\hat{\mu}}(S,\mathcal{H}_a),$$

so together with the last lemma we have $h_{\hat{\mu}}(T,\mathcal{H}_a)=0=h_{\hat{\mu}}(S,\mathcal{H}_a)$.

Proof: $\mathcal{H}_a$ is T and S invariant by construction, so we can use Remark 2.8. p and q are generators of a nonlacunary group so

$$\frac{\log p}{\log q}\neq\frac{n_a}{m_a}. \quad \blacksquare$$

LEMMA 8.4: *If $\hat{\mu}$ a.e. $\hat{y}\in\hat{Y}$ is not symmetric, then*

$$T(P)\subseteq\bigvee_{a=0}^{h}\mathcal{H}_a\vee\bigvee_{i=0}^{\infty}T^{-i}(P)\quad\text{and}\quad h_{\hat{\mu}}\left(P|\bigvee_{a=0}^{h}\mathcal{H}_a\vee\bigvee_{i=1}^{\infty}T^{-i}(P)\right)=0.$$

Proof: Suppose not. This means we can find points $\hat{y}_1$ and $\hat{y}_2$ with $\varphi(\hat{y}_1)=\varphi(\hat{y}_2),\delta_a(\hat{y}_1,r)=\delta_a(\hat{y}_2,r)\ \forall r,a=0,\ldots,h$, but $\hat{y}_1(-1,0)\neq\hat{y}_2(-1,0)$. If also $\hat{y}_1(-1,k_{01})\neq\hat{y}_2(-1,k_{01})$ then $\hat{y}_1$ is a δ_0-symmetric point and we have a contradiction. So assume $\hat{y}_1(-1,k_{01})=\hat{y}_2(-1,k_{01})$.

Now define $u_i(b,c)=\hat{y}_i(b,c)$ for $b\geq-1,c\geq0$. Extend so that $\mathcal{D}_0(\hat{u}_1)=\mathcal{D}_0(\hat{u}_2)$. Let $\delta_a(\hat{u}_i,r)=\delta_a(\hat{y}_i,r)$. By assumption $\hat{u}_1(-1,0)\neq\hat{u}_1(-1,0)$. If also $\hat{u}_1(-1-k_{11}m_1,k_{11}n_1)\neq\hat{u}_2(-1-k_{11}m_1,k_{11}n_1)$ then $\hat{u}_1$ is a δ_1-symmetric point and we have a contradiction. Thus they must be the same at $(-1-k_{11}m_1,k_{11}n_1)$.

Continue with this process until finally we will have built two points with the same $\mathcal{D}_{h-1}$, same distributions which differ at $(-1,0)$. But that makes them δ_h-symmetric points. Thus they must be equal at $(-1,0)$ and so $\hat{y}_1(-1,0)=\hat{y}_2(-1,0)$. $\blacksquare$

LEMMA 8.5: $h_{\hat{\mu}}(T,P) = h_{\hat{\mu}}\left(T, \bigvee_{a=0}^{h} \mathcal{H}_a\right)$.

Proof:

$$h_{\hat{\mu}}(T,P) = h_{\hat{\mu}}\left(T, \bigvee_{a=0}^{h} \mathcal{H}_a\right) + h_{\hat{\mu}}\left(T, P | \bigvee_{a=0}^{h} \mathcal{H}_a\right).$$

But

$$h_{\hat{\mu}}\left(T, P | \bigvee_{a=0}^{h} \mathcal{H}_a\right) = h_{\hat{\mu}}\left(P | \bigvee_{a=0}^{h} \mathcal{H}_a \vee \bigvee_{i=1}^{\infty} T^{-i}(P)\right) = 0$$

by the last lemma. ∎

LEMMA 8.6: *For $\hat{\mu} \in \hat{\mathcal{M}}_0$, but $\hat{\mu} \neq \hat{m}, h_{\hat{\mu}}(T,P) = h_{\hat{\mu}}(S,P) = 0$, and thus by Remark 2.7, $h_{\hat{\mu}}(T) = h_{\hat{\mu}}(S) = 0$.*

Proof: By Lemmas 8.1 and 8.3

$$h_{\hat{\mu}}\left(S, \bigvee_{a=0}^{h} \mathcal{H}_a\right) \leq \sum_{a=0}^{h} h_{\hat{\mu}}(S, \mathcal{H}_a) = 0.$$

Thus

$$0 = h_{\hat{\mu}}\left(S, \bigvee_{a=0}^{h} \mathcal{H}_a\right) = \frac{\log q}{\log p} h_{\hat{\mu}}\left(T, \bigvee_{a=0}^{h} \mathcal{H}_a\right) = \frac{\log q}{\log p} h_{\hat{\mu}}(T,P)$$

by the last lemma, which equals $h_{\hat{\mu}}(S,P)$ ∎

This completes the proof of Theorem A. In review, we have shown that multiplication by p and q on the circle gives rise to distributions on which combinations of these functions act periodically. The exact combination needed is given by the ratio of terms common to both integers. We then use these distributions to define symmetric points and show that the set of such points has measure 1 or 0. This is the dichotomy between Lebesgue measure and measures of entropy zero.

References

[F] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in diophantine approximation*, Math. Systems Theory **1** (1967), 1–49.

[L] R. Lyons, *On measures simultaneously 2- and 3-invariant*, Isr. J. Math. **61** (1988), 219–224.

[R] D. Rudolph, *×2 and ×3 invariant measures and entropy*, Ergodic Theory and Dynamical Systems **10** (1991), 395–406.

[Ro] V. A. Rohlin, *On the decomposition of dynamical systems into transitive components*, Mat. Sb. **67** (1949), 235–249.