Chapter 9
Commuting Automorphisms

Automorphisms or endomorphisms of (infinite) compact groups are soft in the following sense: there are many invariant probability measures, and many closed invariant subsets. We have already seen the symbolic coding of the map $x \mapsto 2x \pmod{1}$ and of the toral automorphism corresponding to the matrix $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$; in each case the symbolic description allows many invariant measures and closed invariant sets to be found.

Furstenberg [64] noted that the situation is very different for measures or closed invariant sets invariant under two genuinely distinct endomorphisms. We start by presenting his original topological result for closed subsets of the circle invariant under two endomorphisms, and go on to describe related measurable results.

We will only concern ourselves with these kind of rigidity questions and will not study the many other interesting dynamical or ergodic theoretic properties of these systems in detail, for which we refer to the monograph of Schmidt [182] and its references.

9.1 Closed Invariant Sets: Furstenberg's Theorem

Before addressing the measurable questions, we describe[(39)] the simple topological origin of the ergodic-theoretic questions considered in the next sections. The properties being dealt with in this chapter concern the semigroup $2^{\mathbb{N}}3^{\mathbb{N}} = \{2, 3, 4, 6, 9, 12, \dots\} \subseteq \mathbb{N}$ generated by 2 and 3, and we begin with some general observations about such semigroups.

Definition 9.1 (Lacunarity). A multiplicative semigroup $S \subseteq \mathbb{N}$ is *lacunary* if there is some $a \in \mathbb{N}$ with the property that any $s \in S$ is an integer power of a.

Two elements s_1, s_2 of a semigroup are said to be *multiplicatively independent* if $s_1^m = s_2^n$ for $m, n \in \mathbb{N}_0$ implies that $m = n = 0$.

Clearly the semigroup generated by a single element is lacunary; for example $\{2, 4, 8, \dots\}$ is lacunary. There are many others however; e.g. the semigroup generated by 4 and 8 is lacunary but not generated by a single element.

Each $k \in \mathbb{N}$ defines an endomorphism $S_k : \mathbb{T} \to \mathbb{T}$ defined by $S_k(t) = kt$ modulo 1. A set $A \subseteq \mathbb{T}$ is called S_k-invariant if $S_k x \in A$ whenever $x \in A$, and is called S-invariant for a subset $S \subseteq \mathbb{N}$ if $S_k x \in A$ whenever $x \in A$ and $k \in S$. Lacunary semigroups have many non-trivial closed invariant sets, as shown in the next example.

Example 9.2. The middle-third Cantor set

$$\left\{ x \in \mathbb{T} \mid x = \sum_{n=1}^{\infty} e_n 3^{-n} \text{ has } e_n \in \{0, 2\} \text{ for all } n \geqslant 1 \right\}$$

is invariant under S_3 (and hence under any semigroup in the lacunary semigroup $\{3, 9, 27, \dots\}$).

The next result, due to Furstenberg [64], shows that there are no 'fractal' or[†] 'non-algebraic' closed invariant sets under the action of a non-lacunary semigroup.

Theorem 9.3 (Furstenberg's '$\times 2, \times 3$' theorem). *Let S be a non-lacunary semigroup in $\mathbb{N}$ and let A be a closed subset of $\mathbb{T}$ invariant under S. Then either A is a finite set of rational points or $A = \mathbb{T}$.*

Lemma 9.4 (Non-lacunary subgroups). *The following properties of a semigroup $S \subseteq \mathbb{N}$ are equivalent.*

(1) *S is non-lacunary;*
(2) *S contains two multiplicatively independent elements;*
(3) *if $S = \{s_1, s_2, \dots\}$ with $s_1 < s_2 < s_3 < \cdots$ then $\frac{s_{n+1}}{s_n} \to 1$ as $n \to \infty$.*

PROOF. Suppose (3) holds and let $a, b \in S \smallsetminus \{1\}$ with $|\frac{a}{b} - 1| < \frac{1}{2}$. If we have $a^m = b^n$ then uniqueness of prime factorization implies that $a = c^k$ and $b = c^\ell$ for some integers c, k, ℓ. Therefore $|c^{k-\ell} - 1| < \frac{1}{2}$ which is impossible for an integer $c \geqslant 2$. This shows that (3) implies (2). That (2) implies (1) is clear.

Assume now that (1) holds. Since $\log S$ is an additive semigroup in $\mathbb{R}$, we have that $L = \log S - \log S$ is an additive subgroup of $\mathbb{R}$, and is therefore either dense or discrete. If L is discrete, then $L = \mathbb{Z}t$ for some $t > 0$, so $S \subseteq (a)^{\mathbb{N}}$ for $a = \exp t$. Note that $a \in \mathbb{Q}$ (since $t = \log s_1 - \log s_2$ for some $s_1, s_2 \in S$) and that $a = \exp(t)$ may not be in S but $a^k \in S$ for some $k \geqslant 1$, which gives that $a \in \mathbb{N}$ and that S is lacunary in contradiction to our assumption.

[†] We do not want to define fractal sets or algebraic sets precisely, but hope that the reader agrees with us that the Cantor set is fractal and non-algebraic while finite sets of rational numbers are not fractal but are somewhat algebraic.

Therefore L is dense. Since S is a semigroup there exists $t_1, t_2, \ldots \in \log S$ with

$$\log S = \Big\{ \sum_k n_k t_k \mid n_k \in \mathbb{N}_0 \text{ for all } k, n_k = 0 \text{ for all but finitely many } k \Big\}.$$

In particular, $L = \bigcup_{n=1}^{\infty} L_n$ with

$$L_n = \log S - n(t_1 + \cdots + t_n).$$

satisfying $L_n \subseteq L_{n+1}$. Fix some $\varepsilon > 0$. Since L is dense there exists some integer $n \geqslant 1$ such that $L_n \cap [0, t_1]$ is ε-dense in $[0, t_1]$. Since $L_n + t_1 \subseteq L_n$ for all $n \geqslant 1$, we see that L_n is ε-dense in $[0, \infty)$. This implies (3). □

Lemma 9.5 (Key observation for density). *Let S and $A \subseteq \mathbb{T}$ be as in Theorem 9.3 but assume that 0 is an accumulation point. Then $A = \mathbb{T}$.*

Proof. By Lemma 9.4(3) the elements $s_1 < s_2 < \cdots$ of S satisfy $\frac{s_{n+1}}{s_n} \to 1$ as $n \to \infty$. Fix $\varepsilon > 0$ and choose N so that $\frac{s_{n+1}}{s_n} < 1 + \varepsilon$ for $n > N$. Since 0 is a limit point of A, we may find $x_n \in A$ with $0 \neq |x_n| < \varepsilon / s_n$, then the finite set $\{s x_n \mid s \in S, s_n \leqslant s \leqslant 1/|x_n|\}$ is ε-dense in $\mathbb{T}$ and lies in A. As $\varepsilon > 0$ was arbitrary and A is closed the lemma follows. □

The following short proof is taken from a paper of Boshernitzan [20].

Proof of Theorem 9.3. Write A' for the set of limit points of A and assume that A is infinite, so A' is a non-empty closed invariant set. We claim that it must contain a rational point. Assume for the purposes of a contradiction that A' does not contain any rational point (which implies that A' is infinite as it contains an irrational point and its orbit), and fix $\varepsilon > 0$. Since S is non-lacunary, we may choose multiplicatively independent numbers $p, q \in S$. Find $t \geqslant 3$ with the properties that $t\varepsilon > 1$ and $\gcd(p, t) = \gcd(q, t) = 1$. It follows that

$$p^u \equiv q^u \equiv 1 \pmod{t} \tag{9.1}$$

for some $u \geqslant 1$ (e.g. $u = \phi(t)$ where ϕ is the Euler function). Define a sequence of sets

$$B_t \subseteq B_{t-1} \subseteq \cdots \subseteq B_1 = A'$$

by

$$B_{j+1} = \left\{ x \in B_j \mid x + \tfrac{1}{t} \in B_j \pmod{1} \right\} \tag{9.2}$$

for each j, $1 \leqslant j \leqslant t-1$. We prove the following statements by induction:

- B_j is invariant under S_{p^u} and S_{q^u}.
- B_j is a closed infinite set of irrational numbers.

For $j = 1$ both properties hold by assumption; assume they hold for some j, $1 \leqslant j \leqslant t-1$. Define a set $D_j = B_j - B_j$. Since B_j is compact by assumption, D_j is closed; since B_j is invariant under both S_{p^u} and S_{q^u}, so is D_j; finally 0 must be a limit point of D_j since B_j is infinite. By assumption, the semigroup S' generated by p^u and q^u is non-lacunary. By Lemma 9.5 it follows that $D_j = \mathbb{T}$. We deduce that B_{j+1} is non-empty. By the choice of u in (9.1), the set B_{j+1} is invariant under S_{p^u} and S_{q^u}, contains no rational point by assumption, and is therefore infinite. Finally, B_{j+1} is a closed set because B_j is a closed set and the condition in (9.2) is closed.

We deduce by induction that each of the sets B_j is non-empty, and in particular B_t is non-empty. Pick any point $x_1 \in B_t$ and write

$$x_i = x_1 + \tfrac{i-1}{t} \in B_{t+1-i}$$

for $1 \leqslant i \leqslant t$. By choice of t the set $C = \{x_i \mid 1 \leqslant i \leqslant t\}$ is ε-dense in $\mathbb{T}$ and $C \subseteq B_1 = A'$. Since ε was arbitrary, it follows that A' is dense in $\mathbb{T}$, contradicting the assumption that A' does not contain any rationals.

Thus A' contains some rational $r = n/t$ say. Recall that p and q are multiplicatively independent elements of S. We may assume (replacing r by $p^a q^b r$ for suitable a, b if need be) that

$$\gcd(n,t) = \gcd(p,t) = \gcd(q,t) = 1.$$

As before, choose $u \geqslant 1$ so that $p^u \equiv q^u \equiv 1 \pmod{t}$. The sets A and A' are both invariant under S_{p^u} and S_{q^u}, and by choice of u so are their translates $A' - r$ and $A - r$. Now 0 lies in $A' - r$ and Lemma 9.5 implies $A - r = \mathbb{T}$ and so $A = \mathbb{T}$. □

Exercises for Section 9.1

Exercise 9.1.1. Extend Example 9.2 to show that a lacunary semigroup has many closed invariant subsets for its natural action on the circle by the following steps.

(1) For any $k > 1$, finite set $F \subseteq \mathbb{N}$, and set $A \subseteq \{0, \dots, k-1\}^F$, show that the set

$$\left\{ x \in \mathbb{T} \mid x = \sum_{n=1}^{\infty} e_n k^{-n}, e|_{F+n} \in A \text{ for all } n \geqslant 1 \right\}$$

(where $e|_{F+n}$ means the projection of $e = (e_n)$ onto the set of coordinates $F + n \subseteq \mathbb{N}$) is invariant under S_k.

(2) More generally, show that for any $k > 1$ the map S_k has uncountably many closed invariant sets.

9.2 Joinings

Recall from Section 2.5 (see also [52, Def. 6.7]) that a *joining* of two measure-preserving systems X and Y is a Borel probability measure on $X \times Y$ that is invariant under $T \times S$, defined on $\mathscr{B}_X \otimes \mathscr{B}_Y$, and projects to the measures μ and ν on the X and Y coordinates respectively, where $\mathsf{X} = (X, \mathscr{B}_X, \mu, T)$ and $\mathsf{Y} = (Y, \mathscr{B}_Y, \nu, S)$. This definition extends in a natural way to two measure-preserving actions of a group: the only change is that the joining measure is required to be invariant under the product group action. As in [52, Def. 6.14], we say that two group actions are disjoint if the only joining is the product of the two measures.

The space of joinings between two ergodic group automorphisms is a vast and unmanageable collection in general, whereas the space of joinings between two ergodic circle rotations is much smaller and easy to understand. One of the manifestations of rigidity for mixing $\mathbb{Z}^2$-actions by automorphisms with finite positive entropy is that non-trivial joinings(40) only exist when the two systems are algebraically related. In this section we record a particularly simple instance of this phenomena on disconnected groups [50], which also helps to motivate some arguments that will appear in the proof of Rudolph's theorem (Theorem 9.9).

Notice that the construction of the relatively independent joining (see the monograph of Furstenberg [66, Ch. 5, Sect. 4] or [52, Def. 6.15]) shows that disjointness between two systems implies that they have no non-trivial common factors.

Recall Ledrappier's example, which is the $\mathbb{Z}^2$-action defined by the shift on the compact group

$$X_{\text{⠓}} = \{x \in \mathbb{F}_2^{\mathbb{Z}^2} \mid x_{\mathbf{n}+\mathbf{e}_1} + x_{\mathbf{n}+\mathbf{e}_2} + x_{\mathbf{n}} = 0 \text{ for all } \mathbf{n} \in \mathbb{Z}^2\},$$

where $\mathbb{F}_2 = \{0, 1\}$ denotes the field with two elements (see [116] for Ledrappier's original paper or [52, Sect. 8.2]). As the notation suggests, having fixed the binary alphabet $\{0, 1\}$, this system is determined by its defining shape ⠓. In this section we will consider a simple instance of how the measurable structure of such a system varies as the defining shape is changed.

Example 9.6 (Reverse Ledrappier's Example). The reverse Ledrappier example is the $\mathbb{Z}^2$-action by shifts on the compact group

$$X_{\text{⠚}} = \{x \in \mathbb{F}_2^{\mathbb{Z}^2} \mid x_{\mathbf{n}-\mathbf{e}_1} + x_{\mathbf{n}+\mathbf{e}_2} + x_{\mathbf{n}} = 0 \text{ for all } \mathbf{n} \in \mathbb{Z}^2\}.$$

Write $\mathsf{X}_{\text{⠓}}$ for the measure-preserving $\mathbb{Z}^2$ system defined by the shift σ on $X_{\text{⠓}}$, preserving Haar measure $m_{X_{\text{⠓}}}$ defined on the Borel σ-algebra $\mathscr{B}_{X_{\text{⠓}}}$, and similarly for the reverse shape ⠚. Notice that we write σ for the natural shift action on any group of the form $F^{\mathbb{Z}^2}$ or on any of its invariant subgroups, where F denotes any finite group.

Theorem 9.7. *The systems* $\mathsf{X}_{\text{⠓}}$ *and* $\mathsf{X}_{\text{⠚}}$ *are disjoint.*

We will prove this by showing that any joining on the group $X_{\text{⠓}} \times X_{\text{⠚}}$ is invariant under translation by $X_{\text{⠓}}$ (acting canonically on the first factor).

The geometric arguments used in the proof are related to the notion of expansive subdynamics introduced by Boyle and Lind [27] to study geometrical properties of expansive topological $\mathbb{Z}^d$-actions.

Before embarking on the proof we assemble some properties of the two systems.

Lemma 9.8. *Write* $Y = \{x \in X_{\bullet\bullet}^{\bullet} \mid x_{\mathbf{m}} = 0 \text{ for } m_1 > 0\}$*. Then the subgroup*

$$Z = \bigcup_{n=1}^{\infty} \sigma^{(-n,0)}(Y)$$

is dense in $X_{\bullet\bullet}^{\bullet}$*. It follows that the Haar measure* $m_{X_{\bullet\bullet}^{\bullet}}$ *is the only Borel probability measure on* $X_{\bullet\bullet}^{\bullet}$ *invariant under translation by all elements of* Z*, and is the only* σ*-invariant probability measure invariant under translation by all elements of* Y*.*

PROOF. It is enough to prove that Z is a dense subgroup of $X_{\bullet\bullet}^{\bullet}$; the claim about invariant measures follows since a Borel measure is determined by how it integrates continuous functions (as in the proof that (3) $\implies$ (1) in [52, Th. 4.14]). Now $\sigma^{(0,\pm 1)}(Y) = Y$, so the subgroup Z is invariant under the whole shift σ. Thus in order to show that Z is dense, it is enough to show that for any non-empty cylinder set $C \subseteq X_{\bullet\bullet}^{\bullet}$ defined by specifying the coordinates in some square $\{(a_1, a_2) \mid 0 \leqslant a_1, a_2 \leqslant N\}$ contains an element of Z. Choose a sequence

$$(y_{(0,0)}, y_{(1,0)}, \dots, y_{(2N,0)}) \tag{9.3}$$

(see Figure 9.1) chosen to ensure† that the only way to extend y to the coordinates in

$$\{(a_1, a_2) \mid 0 \leqslant a_1, a_2 \leqslant N\}$$

agrees with the condition defining the cylinder set C (e.g. fix some $x \in C$ and define $y_{(j,0)} = x_{(j,0)}$ for all $j = 0, \dots, 2N$).

We extend the finite sequence in (9.3) by 0s on the right and the left to define an element of $\{0,1\}^{\mathbb{Z}}$. Using these coordinates and the defining relation $:\!\bullet$ we may define $y_{(n_1,n_2)}$ for $n_2 \geqslant 0$ uniquely. The coordinates of $y_{(n_1,n_2)}$ for $n_2 < 0$ are not determined by these choices. However, by choosing $y_{(2N,n_2)} = y_{(2N,0)}$ for all $n_2 \leqslant 0$ and applying the defining relation once more, this defines a point y in $\sigma^{(-2N,0)}Y \cap C$ as required. $\square$

We now turn to the main argument, which may be described as studying the *entropy geometry* of $\mathsf{X}_{\bullet\bullet}^{\bullet}$ and $\mathsf{X}_{\bullet\bullet}^{\;\bullet}$. Less cryptically, we calculate the entropy $h_\rho(\sigma^{(1,0)})$ in two different ways and will obtain information about ρ as a

† By definition of $X_{\bullet\bullet}^{\bullet}$: for example, $y_{(0,0)}$ and $y_{(0,1)}$ together determine the coordinate $y_{(1,0)} = y_{(0,0)} + y_{(0,1)}$, while $y_{(0,0)}, y_{(0,1)}, y_{(0,2)}$ together also determine $y_{(1,1)}$, and hence $y_{(2,0)} = y_{(1,0)} + y_{(1,1)}$.

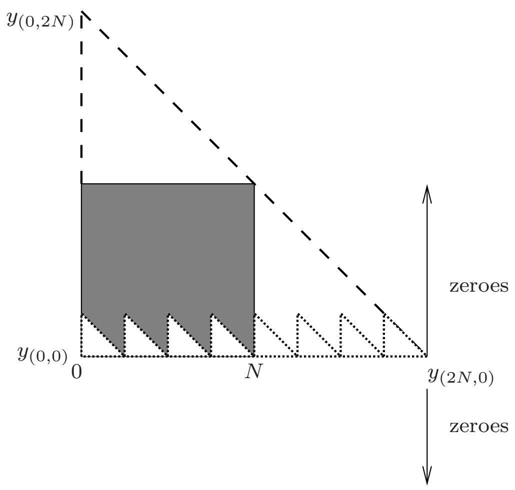

Fig. 9.1: Constructing a point in Z that meets the cylinder set C. The zeroes to the right of $(2N,0)$ force zeroes in the whole upper right corner.

result. Notice that for any finite alphabet group F and closed shift-invariant subgroup of $F^{\mathbb{Z}^2}$, the *state partition*

$$\xi = \Big\{ \big\{ x \in F^{\mathbb{Z}^2} \mid x_{(0,0)} = f \big\} \mid f \in F \Big\} \tag{9.4}$$

is a generator for the whole action in the sense that $\bigvee_{\mathbf{n}\in\mathbb{Z}^2} \sigma^{-\mathbf{n}}(\xi)$ is the Borel σ-algebra in $F^{\mathbb{Z}^2}$. Of course the state partition may or may not generate under the action of some subgroup $L \leqslant \mathbb{Z}^2$, depending on the exact rules defining the closed shift-invariant subgroup and its relation to the subgroup L.[41]

More generally, to any finite set of coordinates in $\mathbb{Z}^2$ there is a naturally associated partition defined by all the possible cylinder sets obtained by specifying those coordinates. Depending on the rules defining the closed shift-invariant subgroup the same partition may be defined by different sets of coordinates. Similarly, infinite sets of coordinates define sub σ-algebras and once more one may obtain the same σ-algebra possibly in different ways.

Proof of Theorem 9.7. Let $X = X_{\substack{\bullet\\ \bullet\bullet}} \times X_{\substack{\bullet\\ \bullet\bullet}} \subseteq \left(\mathbb{F}_2^2\right)^{\mathbb{Z}^2}$, write σ for the usual shift $\mathbb{Z}^2$-action on X, and let $\rho \in \mathsf{J}(\mathsf{X}_{\substack{\bullet\\ \bullet\bullet}}, \mathsf{X}_{\substack{\bullet\\ \bullet\bullet}})$ be a joining of the two systems (so ρ is a shift-invariant Borel probability measure on X that projects to Haar measure on each of the two coordinates). It will be helpful to think of $X_{\substack{\bullet\\ \bullet\bullet}}$ as being a 'top' layer and $X_{\substack{\bullet\\ \bullet\bullet}}$ as a 'bottom' layer. Notice that points in X obey the rule $\substack{\bullet\\ \bullet\bullet}$ in the top layer, obey the rule $\substack{\bullet\\ \bullet\bullet}$ in the bottom layer and, as may be checked directly, obey the rule

(9.5)

in both layers at once.†

By the future formula for entropy (Proposition 2.19(1)) we have

$$h_\rho\left(\sigma^{(1,0)}\right) = \sup_\xi H_\rho\Big(\xi \big| \bigvee_{i=1}^{\infty} \sigma^{-(i,0)}(\xi)\Big),$$

where the supremum is taken over all finite partitions. As usual (by Theorem 1.21 or 2.20) the supremum is attained by a partition that generates under $\sigma^{(1,0)}$, or by taking a limit along a sequence of partitions that generate.

Define‡

$$\xi_\ell = \bigvee_{j=0}^{2\ell} \sigma^{(0,j)}\xi \vee \bigvee_{j=1}^{2\ell} \sigma^{(-j,-j)}\xi$$

to be the partition defined by specifying the $(4\ell+1)$ coordinates illustrated in Figure 9.2. The sequence of partitions (ξ_ℓ) satisfies the hypothesis of the

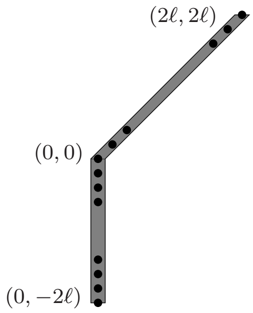

Fig. 9.2: The coordinates defining the partition ξ_ℓ.

† Alternatively, one can argue in terms of polynomials as follows. The relation corresponds to annihilation by the polynomial $1+u_1+u_2$ in $\mathbb{F}_2[u_1^{\pm 1}, u_2^{\pm 1}]$, the relation corresponds to annihilation by $1+u_1^{-1}+u_2$, and the relation (9.5) corresponds to annihilation by the product

$$(1+u_2+u_2)(1+u_1^{-1}+u_2) = u_1^{-1} + u_1 + u_1u_2 + u_1^{-1}u_2 + u_2^2,$$

whose support $\{(-1,0),(1,0),(-1,1),(1,1),(0,2)\}$ is illustrated in (9.5).

‡ We defined ξ_ℓ formally for completeness, but it is less confusing to simply argue in the following directly with the coordinates that define ξ_ℓ as in Figure 9.2. Linking these two view points is the formula

$$\sigma^{-\mathbf{n}}\{x \in F^{\mathbb{Z}^2} \mid x_{\mathbf{m}} = f\} = \{x \in F^{\mathbb{Z}^2} \mid x_{\mathbf{m}+\mathbf{n}} = f\}$$

for all $\mathbf{m}, \mathbf{n} \in \mathbb{Z}^2$ and $f \in F$.

sequence formulation of the Kolmogorov–Sinaĭ theorem (Theorem 2.20), so

$$h_\rho\left(\sigma^{(1,0)}\right) = \lim_{\ell\to\infty} H_\rho\Big(\xi_\ell \big| \bigvee_{i=1}^{\infty} \sigma^{-(i,0)}\xi_\ell\Big) = \lim_{\ell\to\infty} H_\rho\big(\xi_\ell \big| \mathscr{C}_\ell\big), \tag{9.6}$$

where $\mathscr{C}_\ell = \bigvee_{i=1}^{\infty} \sigma^{-(i,0)}\xi_\ell$ is the σ-algebra corresponding to the known coordinates in Figure 9.3. In that sense $H_\rho\left(\xi_\ell \big| \mathscr{C}_\ell\right)$ is the entropy corresponding to learning the shaded coordinates given complete knowledge of the coordinates to the right, as illustrated in Figure 9.3.

We also note that in the top layer $X_{:\cdot}$ the coordinates at $(1,1),\dots,(2\ell,2\ell)$ are determined (by the rule $:\cdot$ defining $X_{:\cdot}$), while in the bottom layer $X_{\cdot:}$ the coordinates at $(0,-1),\dots,(0,-2\ell)$ are determined (by the rule $\cdot:$). The

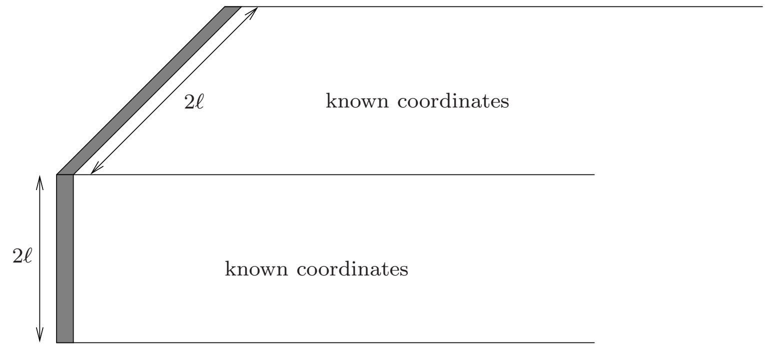

Fig. 9.3: The entropy calculation $H_\rho\left(\xi_\ell|\mathscr{C}_\ell\right)$.

remaining coordinates are determined once a single choice is made in the top layer and a single choice is made in the bottom layer at one (arbitrary) undetermined coordinate. For example, we may choose the top layer coordinate at $(0,-\ell)$ (that is, in the middle of the left vertical front) and the bottom layer coordinate at (ℓ,ℓ) (that is, in the middle of the left roof) as in Figure 9.4.

Once these choices have been made, the defining relation $:\cdot$ in the top layer propagates the choice to all of the top layer shaded coordinates, and the defining relation $\cdot:$ in the bottom layer propagates the choice to all of the bottom layer shaded coordinates. Thus

$$H_\rho\left(\xi_\ell \big| \mathscr{C}_\ell\right) = H_\rho\left(\sigma^{(0,\ell)}\xi \vee \sigma^{(-\ell,-\ell)}\xi \big| \mathscr{C}_\ell\right) \tag{9.7}$$

where, as in (9.4), we write ξ for the state partition. In particular, (9.6) shows that $h_\rho(\sigma^{(1,0)}) \leqslant \log 4$.

We now show how this combinatorial argument can be used to give a geometrical decomposition of the entropy. Write

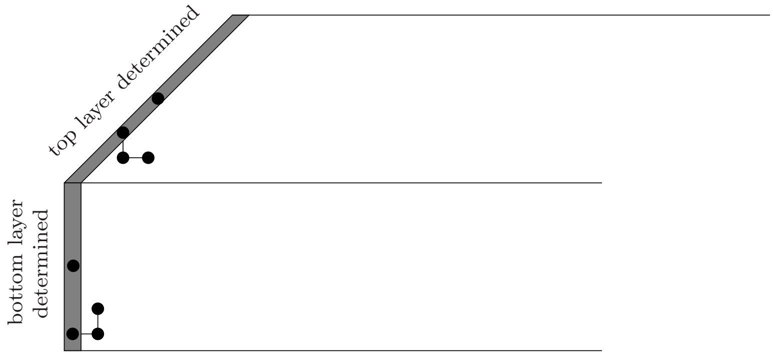

Fig. 9.4: The choice required to determine the shaded region.

$$\mathsf{H}_1 = \{\mathbf{m} \in \mathbb{Z}^2 \mid \mathbf{m} \cdot (1,0) > 0\}$$

for the right half-plane in $\mathbb{Z}^2$, and

$$\mathsf{H}_2 = \{\mathbf{m} \in \mathbb{Z}^2 \mid \mathbf{m} \cdot (1,-1) > 0\},$$

for the lower right diagonal half-plane in $\mathbb{Z}^2$. Notice that $\mathsf{H}_1 \cap \mathsf{H}_2$ is approximated by the region of known coordinates in Figure 9.3 for large ℓ. By (9.7) and additivity of entropy (Proposition 1.7) we have

$$\begin{aligned} H_\rho\left(\xi_\ell \big| \mathscr{C}_\ell\right) &= H_\rho\left(\sigma^{(0,\ell)}\xi \big| \mathscr{C}_\ell\right) + H_\rho\left(\sigma^{-(\ell,\ell)}\xi \big| \mathscr{C}_\ell \vee \sigma^{(0,\ell)}\xi\right) \\ &= H_\rho\left(\xi \big| \sigma^{(0,-\ell)}\mathscr{C}_\ell\right) + H_\rho\left(\xi \big| \sigma^{(\ell,\ell)}\mathscr{C}_\ell \vee \sigma^{(\ell,2\ell)}\xi\right) \end{aligned} \tag{9.8}$$

for any $\ell \geqslant 1$. Studying the coordinates that are used to define the σ-algebra $\sigma^{(0,-\ell)}\mathscr{C}_\ell$, we see that this defines an increasing sequence of σ-algebras converging to $\bigvee_{\mathbf{n}\in\mathsf{H}_1} \sigma^{-\mathbf{n}}\xi$. Noting that $\sigma^{(-1,\ell-1)}\xi \subseteq \mathscr{C}_{\ell+1}$ or equivalently $\sigma^{(\ell,2\ell)}\xi \subseteq \sigma^{(\ell+1,\ell+1)}\mathscr{C}_{\ell+1}$, we see in the same way that $\sigma^{(\ell,\ell)}\mathscr{C}_\ell \vee \sigma^{(\ell,2\ell)}\xi$ also increases with ℓ and has the limit $\bigvee_{\mathbf{n}\in\mathsf{H}_2} \sigma^{-\mathbf{n}}\xi$. Hence we may apply continuity of entropy (Proposition 2.12), combine (9.6) and (9.8) to obtain that

$$\begin{aligned} h_\rho(\sigma^{(1,0)}) &= H_\rho\Big(\xi \big| \bigvee_{\mathbf{n}\in\mathsf{H}_1} \sigma^{-\mathbf{n}}\xi\Big) + H_\rho\Big(\xi \big| \bigvee_{\mathbf{n}\in\mathsf{H}_2} \sigma^{-\mathbf{n}}\xi\Big) \\ &= H_\rho\big(\xi \big| \mathscr{B}_{\mathsf{H}_1}\big) + H_\rho\big(\xi \big| \mathscr{B}_{\mathsf{H}_2}\big) \end{aligned} \tag{9.9}$$

where $\mathscr{B}_{\mathsf{H}_j}$ is the σ-algebra defined by the coordinates in H_j for $j = 1, 2$. We note that each of the terms in (9.9) is no larger than $\log 2$ by the argument concerning Figure 9.4.

We will now use the joining assumption to obtain a second formula for the entropy $h_\rho(\sigma^{(1,0)})$. In fact, $\mathscr{B}_1 = \mathscr{B}_{X_{\therefore}} \times \mathscr{N}_{X_{\therefore}}$ is the σ-algebra corresponding to the factor $\mathsf{X}_{\substack{\bullet\,\\ \bullet\,\bullet}}$ of the dynamical system

$$(X, \mathscr{B}_{X_{\substack{\bullet\,\\ \bullet\,\bullet}}} \otimes \mathscr{B}_{X_{\substack{\,\bullet\\ \bullet\,\bullet}}}, \rho, \sigma^{(1,0)}),$$

so by the Abramov–Rokhlin formula (Corollary 2.21) we can also decompose the entropy as

$$h_\rho(\sigma^{(1,0)}) = h_{m_{\substack{\bullet\,\\ \bullet\,\bullet}}}(\sigma^{(1,0)}) + h_\rho(\sigma^{(1,0)} \big| \mathscr{B}_1) \tag{9.10}$$

where we write $m_{\substack{\bullet\,\\ \bullet\,\bullet}}, m_{\substack{\,\bullet\\ \bullet\,\bullet}}$ for the Haar measures on $X_{\substack{\bullet\,\\ \bullet\,\bullet}}, X_{\substack{\,\bullet\\ \bullet\,\bullet}}$ respectively. We claim that the right-hand side of (9.9) and of (9.10) agree with each other term by term.

Going through the arguments from equations (9.6) to (9.9) again, but for $h_\rho(\sigma^{(1,0)}|\mathscr{B}_1)$ and the only change throughout that every entropy expression is in addition conditioned on the invariant σ-algebra $\mathscr{B}_1$, we obtain that

$$h_\rho(\sigma^{(1,0)} \big| \mathscr{B}_1) = H_\rho\left(\xi \big| \mathscr{B}_{\mathsf{H}_1} \vee \mathscr{B}_1\right) + H_\rho\left(\xi \big| \mathscr{B}_{\mathsf{H}_2} \vee \mathscr{B}_1\right). \tag{9.11}$$

However,

$$\mathscr{B}_{\mathsf{H}_1} \vee \mathscr{B}_1 = \bigvee_{\mathbf{n} \in \mathsf{H}_1} \sigma^{-\mathbf{n}}\xi \vee \mathscr{B}_1 = \mathscr{B}_{X_{\substack{\bullet\,\\ \bullet\,\bullet}}} \otimes \mathscr{B}_{X_{\substack{\,\bullet\\ \bullet\,\bullet}}}$$

since the top layer $\mathsf{X}_{\substack{\bullet\,\\ \bullet\,\bullet}}$ corresponds to $\mathscr{B}_1$ and for the bottom layer $\mathsf{X}_{\substack{\,\bullet\\ \bullet\,\bullet}}$ all coordinates are determined by the coordinates in H_1. Therefore the first term on the right-hand side of (9.11) vanishes.

For the second term in (9.11) we note that

$$\mathscr{B}_{\mathsf{H}_2} = \bigvee_{\mathbf{n} \in \mathsf{H}_2} \sigma^{-\mathbf{n}}\xi \supseteq \mathscr{B}_1 = \mathscr{B}_{X_{\substack{\bullet\,\\ \bullet\,\bullet}}} \times \mathscr{N}_{X_{\substack{\,\bullet\\ \bullet\,\bullet}}} \tag{9.12}$$

since in the top layer (a copy of $\mathsf{X}_{\substack{\bullet\,\\ \bullet\,\bullet}}$) the coordinates in H_2 determine everything. Therefore, (9.11) becomes

$$h_\rho(\sigma^{(1,0)} \big| \mathscr{B}_1) = H_\rho\left(\xi \big| \mathscr{B}_{\mathsf{H}_2}\right).$$

Together with (9.9) and (9.10) we therefore have

$$\log 2 = h_{m_{\substack{\bullet\,\\ \bullet\,\bullet}}}(\sigma^{(1,0)}) = H_\rho\left(\xi \big| \mathscr{B}_{\mathsf{H}_1}\right).$$

Note that this also give

$$\log 2 = H_\rho\left(\sigma^{(n,0)}\xi \big| \sigma^{(n,0)}\mathscr{B}_{\mathsf{H}_1}\right)$$

for any $n \in \mathbb{Z}$. Adding these just as in the proof of the future formula (Proposition 2.19(1)) and using $\xi \vee \bigvee_{\mathbf{n} \in \mathsf{H}_1} \sigma^{-\mathbf{n}}(\xi) = \sigma^{(1,0)} \bigvee_{\mathbf{n} \in \mathsf{H}_1} \sigma^{-\mathbf{n}}(\xi)$ we get for any $N \geqslant 1$ that

$$N \log 2 = H_\rho \left(\bigvee_{n=0}^{N-1} \sigma^{(n,0)} \xi \big| \mathscr{B}_{\mathsf{H}_1} \right).$$

Now use the same argument as in (9.12) to see that

$$\mathscr{B}_{\mathsf{H}_1} = \bigvee_{\mathbf{n} \in \mathsf{H}_1} \sigma^{-\mathbf{n}} \xi \supseteq \mathscr{B}_2 = \mathscr{N}_{X_{\therefore}} \times \mathscr{B}_{X_{\therefore}}. \tag{9.13}$$

Let $\xi_{(1)} \subseteq \mathscr{B}_1$ be the pre-image in $X_{\therefore} \times X_{\therefore}$ of the state partition in $X_{\therefore}$ so that $\xi \vee \mathscr{B}_2 = \xi_{(1)} \vee \mathscr{B}_2$ and therefore

$$N \log 2 = H_\rho \left(\bigvee_{n=0}^{N-1} \sigma^{(n,0)} \xi_{(1)} \big| \mathscr{B}_{\mathsf{H}_1} \right).$$

Since $\bigvee_{n=0}^{N-1} \sigma^{-(n,0)} \xi_{(1)}$ contains 2^N elements, and by the characterization of maximal entropy (Proposition 1.5), we deduce each atom of $\bigvee_{n=0}^{N-1} \sigma^{-(n,0)} \xi_{(1)}$ almost surely has $\rho_x^{\mathscr{B}_{\mathsf{H}_1}}$-measure $\frac{1}{2^N}$. However, as these partitions together with $\mathscr{B}_{\mathsf{H}_1}$ generate the Borel σ-algebra by (9.13) this forces $\rho_x^{\mathscr{B}_{\mathsf{H}_1}}$ to be the Haar measure on a coset of the group $Y \subseteq X_{\therefore} \times \{0\}$ (since the Haar measure gives the same weights to the partition elements).

It follows that ρ is also invariant under Y and under σ, which implies that ρ is also invariant under the group $Z \subseteq X_{\therefore} \times \{0\}$ as in Lemma 9.8. Now notice that Z leaves invariant the $\mathscr{B}_2$-atoms $X_{\therefore} \times \{x_2\}$ for all points $x = (x_1, x_2)$ lying in $X_{\therefore} \times X_{\therefore}$. Therefore, the conditional measure $\rho_x^{\mathscr{B}_2}$ are invariant under Z and Lemma 9.8 implies that (almost surely the conditional measures $\rho_x^{\mathscr{B}_2}$ and so also) ρ is invariant under translation by $X_{\therefore} \times \{0\}$. This forces ρ to be $m_{\therefore} \times \nu$ for some measure ν on $X_{\therefore}$, and ν must be $m_{\therefore}$ since it is the projection of ρ onto the bottom layer. This proves Theorem 9.7. $\square$

Exercises for Section 9.2

Exercise 9.2.1. [42] Show that the system $\mathsf{X}_{\therefore}$ has many invariant measures.
(a) Find a non-trivial invariant closed subset in $\mathsf{X}_{\therefore}$ (e.g. by constructing closed subgroup that is invariant under $\sigma^{(2,0)}$ and $\sigma^{(0,2)}$.)
(a) Find an invariant probability measure other than $m_{\therefore}$ for which $\sigma^{(1,0)}$ has positive entropy.

Exercise 9.2.2. [43] Show that $X_{\therefore}$ does exhibit a certain weak form of measure rigidity by showing that if ρ is an invariant Borel probability measure on $X_{\therefore}$ and $\sigma^{(1,0)}$ is mixing with respect to ρ, then $\rho = m_{\therefore}$.

9.3 Rigidity of Positive Entropy Measures for ×2, ×3: Rudolph's Theorem

It is clear that the maps $S_2 : x \mapsto 2x$ modulo 1 and $S_3 : x \mapsto 3x$ modulo 1 on the circle $\mathbb{T}$ each have many invariant probability measures and many closed invariant sets, both infinite and finite.

Furstenberg's topological result described in Section 9.1 motivated the question about how probability measures invariant under both of the maps $\times 2$ and $\times 3$ might look like. That is, could it be that an ergodic probability measure which is invariant under both maps, must be the Lebesgue measure or an atomic measure? A measure invariant under both maps will be called a $\times 2, \times 3$-invariant measure, and the result sought says that under suitable hypotheses (optimally with the only assumption that the measure has no atoms) the only $\times 2, \times 3$-invariant probability measure is Lebesgue measure.

Progress was made on this question by Lyons [129], who proved the result for measures with completely positive entropy with respect to one of the maps (see Definition 2.25). The assumption of completely positive entropy was reduced to just positive entropy by Rudolph [181], and this is the result we present here. Johnson [91] later removed the coprimality assumption, extending(44) Rudolph's theorem to measures invariant under multiplication by 5 and 10 for example.

Theorem 9.9 (Rudolph). *Let S_2 denote $x \mapsto 2x \pmod 1$ and S_3 denote $x \mapsto 3x \pmod 1$ on $\mathbb{T}$. Let μ be an S_2, S_3-invariant probability measure on $\mathbb{T}$ which is ergodic with the property that $h_\mu(S_2^m S_3^n) > 0$ for some $m, n \in \mathbb{N}$. Then μ is the Lebesgue measure $m_{\mathbb{T}}$ on $\mathbb{T}$.*

Here ergodicity is meant with respect to the joint action, that is if $A \subseteq \mathbb{T}$ is measurable and $S_2^{-1}A = S_3^{-1}A = A \pmod \mu$ then $\mu(A) = 0$ or 1.

In order to prove this, it is convenient to work with the invertible extension of the $\mathbb{N}^2$-action in which the pair (m, n) is sent to the map $S_2^m S_3^n$ (see Section A.5 for the construction of the invertible extension of a single map which formally may be used here for each of the generators in turn). To construct the invertible extension, define

$$X = \{x \in \mathbb{T}^{\mathbb{Z}^2} \mid x_{\mathbf{n}+\mathbf{e}_1} = 2x_{\mathbf{n}} \text{ and } x_{\mathbf{n}+\mathbf{e}_2} = 3x_{\mathbf{n}} \text{ for all } \mathbf{n} \in \mathbb{Z}^2\}.$$

The conditions defining X as a subset of $\mathbb{T}^{\mathbb{Z}^2}$ are closed and homogeneous, so X is a compact abelian group. The group X is invariant under the shift action of $\mathbb{Z}^2$ defined by

$$(x_{\mathbf{n}})_{\mathbf{n}} \overset{\text{action of } \mathbf{m}}{\longrightarrow} (x_{\mathbf{n}+\mathbf{m}})_{\mathbf{n}}.$$

Write T_2 for the left shift by $\mathbf{e}_1$ and T_3 for the down shift by $\mathbf{e}_2$. Thus

$$(T_2^m T_3^n x)_{(a,b)} = x_{(a+m,b+n)}$$

for all $a, b, m, n \in \mathbb{Z}$, so T_2, T_3 are commuting automorphisms of X. Together they define a $\mathbb{Z}^2$-action that is the invertible extension of the $\mathbb{N}^2$-action defined by the commuting endomorphisms S_2 and S_3.

Write $\pi_{\mathbf{0}} : X \to \mathbb{T}$ for the surjective homomorphism defined by

$$\pi_{\mathbf{0}}(x) = x_{(0,0)}.$$

By the definition of X, for any $m, n \geqslant 0$ the diagram

$$\begin{tikzcd} X \arrow[r, "T_2^m T_3^n"] \arrow[d, "\pi_{\mathbf{0}}"'] & X \arrow[d, "\pi_{\mathbf{0}}"] \\ \mathbb{T} \arrow[r, "S_2^m S_3^n"'] & \mathbb{T} \end{tikzcd}$$

commutes. Every T_2, T_3-invariant measure μ_X on X defines a S_2, S_3-invariant measure $\mu_{\mathbb{T}} = (\pi_{\mathbf{0}})_* \mu_X$ on $\mathbb{T}$ and (by the properties of the invertible extension) *vice versa*. Moreover, ergodicity is preserved: μ_X is ergodic if and only if $\mu_{\mathbb{T}}$ is (see Section A.5).

It will be useful to visualize the group X as a lattice of circles in which a step to the right is an application of S_2, a step up is an application of S_3 and therefore each point in $\mathbb{T}$ at a given lattice site has two choices for the point to its left and three choices for the point below it. This is illustrated in Figure 9.5.

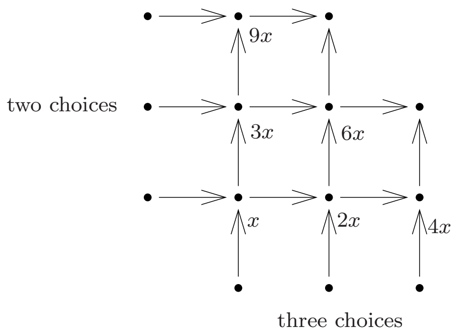

Fig. 9.5: The invertible extension of $\times 2, \times 3$.

Define a partition on $\mathbb{T}$ by

$$\xi_{\mathbb{T}} = \{[0, \tfrac{1}{6}), [\tfrac{1}{6}, \tfrac{2}{6}), \dots, [\tfrac{5}{6}, 1)\}$$

and induce a partition on X by

$$\xi_X = \pi_{\mathbf{0}}^{-1} \xi_{\mathbb{T}}.$$

Lemma 9.10 (The generator). *The partition $\xi_{\mathbb{T}}$ is a generator for S_2 and is a generator for S_3. In general the partition ξ_X is not a generator for T_2, nor is it a generator for T_3. Nonetheless,*

$$h_{\mu_X}(T_2, \xi_X) = h_{\mu_X}(T_2) = h_{\mu_{\mathbb{T}}}(S_2) = h_{\mu_{\mathbb{T}}}(S_2, \xi_{\mathbb{T}})$$

for any probability measure μ_X on X that is invariant under the $\mathbb{Z}^2$-action,

Proof. The only non-trivial part of this is to show that

$$h_{\mu_X}(T_2, \xi_X) = h_{\mu_X}(T_2).$$

Let $\mathscr{C} = \bigvee_{i=-\infty}^{\infty} T_2^{-i}\xi_X$. In Figure 9.6, $\mathscr{C}$ is the σ-algebra generated by the coordinates in the upper half-plane.

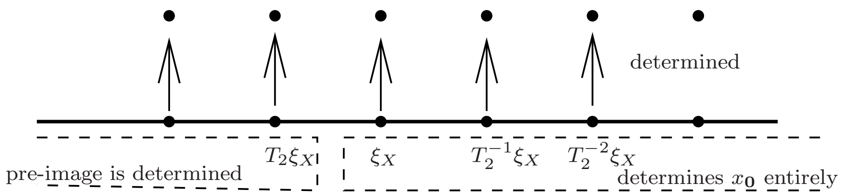

Fig. 9.6: How much is determined about x if $[x]_{\mathscr{C}}$ is known.

Now $T_3\xi_X$ (and the associated factor $T_3\mathscr{C}$) gives the same entropy to T_2. Since $T_3^n\mathscr{C} \nearrow \mathscr{B}_X$, the Kolmogorov–Sinaĭ theorem (Theorem 2.20) shows that

$$h_{\mu_X}(T_2) = \lim_{n\to\infty} h_{\mu_X}(T_2, T_3^n\xi) = h_{\mu_X}(T_2, \xi_X).$$

□

Due to the above, and also because we will work mostly on X, we will not distinguish between $\mu_{\mathbb{T}}$ and μ_X any longer, and simply write μ in the following. An important step in the proof is a relationship between the entropy of T_2 and T_3 which holds for the ×2, ×3-invariant measure μ on X.

Proposition 9.11 (Key entropy formula). *For any σ-algebra $\mathscr{S}$ invariant under T_2 and T_3,*

$$h_\mu(T_3 \big| \mathscr{S}) = \frac{\log 3}{\log 2} h_\mu(T_2 \big| \mathscr{S}).$$

In particular,

$$h_\mu(T_3) = \frac{\log 3}{\log 2} h_\mu(T_2).$$

PROOF. Let $\kappa = \frac{\log 3}{\log 2}$. For any $n \geqslant 1$ the partition $\bigvee_{i=0}^{n-1} S_3^{-i}\xi_{\mathbb{T}}$ comprises intervals of length $\frac{1}{2\cdot 3^n}$, while the partition $\bigvee_{i=0}^{m-1} S_2^{-i}\xi_{\mathbb{T}}$ comprises intervals of length $\frac{1}{2^m\cdot 3}$. Write $\ell_3(n) = \frac{1}{2\cdot 3^n}$ and $\ell_2(m) = \frac{1}{2^m\cdot 3}$. If $m = \lfloor \kappa n \rfloor$ then

$$2^m \leqslant 3^n \leqslant 2^{m+1}$$

and

$$\frac{1}{2}\ell_2(m) = \frac{1}{2^{m+1}\cdot 3} \leqslant \frac{1}{3\cdot 3^n} \leqslant \ell_3(n) \leqslant \frac{1}{2\cdot 2^m} = \frac{3}{2}\ell_2(m).$$

This implies that each interval in $\bigvee_{i=0}^{n-1} S_3^{-i}\xi_{\mathbb{T}}$ intersects at most three intervals in $\bigvee_{i=0}^{m-1} S_2^{-i}\xi_{\mathbb{T}}$ and *vice versa*. It follows that

$$H_\mu\left(\bigvee_{i=0}^{n-1} T_3^{-i}\xi_X \Big| \bigvee_{j=0}^{m-1} T_2^{-j}\xi_X\right) \leqslant \log 3,$$

and

$$H_\mu\left(\bigvee_{j=0}^{m-1} T_2^{-j}\xi_X \Big| \bigvee_{i=0}^{n-1} T_3^{-i}\xi_X\right) \leqslant \log 3,$$

so

$$\begin{aligned} H_\mu\left(\bigvee_{i=0}^{n-1} T_3^{-i}\xi_X \Big| \mathscr{S}\right) &\leqslant H_\mu\left(\bigvee_{i=0}^{n-1} T_3^{-i}\xi_X \vee \bigvee_{j=0}^{m-1} T_2^{-j}\xi_X \Big| \mathscr{S}\right) \\ &\leqslant H_\mu\left(\bigvee_{j=0}^{m-1} T_2^{-j}\xi_X \Big| \mathscr{S}\right) + \underbrace{H_\mu\left(\bigvee_{i=0}^{n-1} T_3^{-i}\xi_X \Big| \bigvee_{j=0}^{m-1} T_2^{-j}\xi_X\right)}_{\leqslant \log 3}. \end{aligned}$$

Now divide by n and let $n \to \infty$ (with $m = \lfloor \kappa n \rfloor$) to deduce that

$$h_\mu(T_3 \big| \mathscr{S}) \leqslant \underbrace{\lim_{n\to\infty} \frac{m}{n}}_{\to\kappa} h_\mu(T_2 \big| \mathscr{S}) + 0.$$

The same argument with the role of S_2 and S_3 reversed gives the reverse inequality. $\square$

A similar argument applies to any map $T_2^m T_3^n$ with $m \geqslant 0$ and $n \geqslant 0$, so we deduce that the following properties of a measure μ are equivalent:

- $h_\mu(T_2) > 0$;
- $h_\mu(T_3) > 0$;
- $h_\mu(T_2^m T_3^n) > 0$ for some $m \geqslant 0$, $n \geqslant 0$;
- $h_\mu(T_2^m T_3^n) > 0$ for all $(m, n) \neq (0, 0)$ with $m \geqslant 0$, $n \geqslant 0$.

Corollary 9.12 (Identical Pinsker σ-algebra). *The maps T_2 and T_3 of (X,μ) have the same Pinsker algebra.*

Proof. Let $\mathscr{P}(T_2)$ be the Pinsker algebra of T_2. Then it is easy to check that $T_3\mathscr{P}(T_2)=\mathscr{P}(T_2)$. Proposition 9.11 then implies that

$$\begin{aligned} h_\mu\left(T_3\middle|\mathscr{P}(T_2)\right) &= \kappa h_\mu\left(T_2\middle|\mathscr{P}(T_2)\right)\\ &= \kappa h_\mu(T_2)\\ &= h_\mu(T_3). \end{aligned}$$

This in turn gives $h_\mu(T_3\big|\mathscr{P}(T_2))=0$ by the Abramov–Rokhlin formula (Corollary 2.21), so $\mathscr{P}(T_2)\subseteq\mathscr{P}(T_3)$. The corollary follows by repeating the argument with T_2 and T_3 switched. □

Proving Theorem 9.9 requires a more geometric understanding of entropy (using similar pictures and ideas as those already used in the last section). Let

$$\mathscr{A}_1=\sigma\left(\bigvee_{i=1}^{\infty}T_2^{-i}\xi_X\right)=T_2^{-1}\pi_{\mathbf{0}}^{-1}\mathscr{B}_{\mathbb{T}},$$

illustrated in Figure 9.7. Also note that $T_3^{-1}\mathscr{A}_1$ equals the σ-algebra generated by the partition at the coordinates

$$\{(1,1),(2,1),(3,1)\dots\}$$

and that $\mathscr{A}_1\supseteq T_3^{-1}\mathscr{A}_1$.

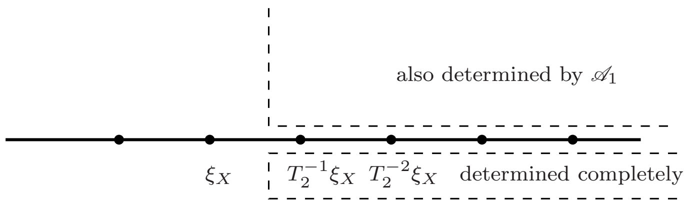

Fig. 9.7: The σ-algebra $\mathscr{A}_1$.

Knowledge of $S_2(y)$ and $S_3(y)$ in $\mathbb{T}$ determines $y\in\mathbb{T}$ uniquely, so

$$T_3^{-1}\xi_X\vee\mathscr{A}_1=\pi_{\mathbf{0}}^{-1}\mathscr{B}_{\mathbb{T}}=\xi_X\vee\mathscr{A}_1.$$

Using this repeatedly and applying T_3^n we also obtain

$$\xi_X\vee T_3^n\mathscr{A}_1=T_3^n\xi_X\vee T_3^n\mathscr{A}_1 \tag{9.14}$$

for any $n\geqslant 0$, see also Figure 9.8.

Fig. 9.8: Proof that $T_3^n \xi_X \vee \mathscr{A}_1 = \xi_X \vee \mathscr{A}_1$.

Lemma 9.10 shows that

$$\begin{aligned} h_\mu(T_2) &= H_\mu(\xi_X \big| \mathscr{A}_1) \\ &= H_\mu(T_3^n \xi_X \big| T_3^n \mathscr{A}_1) && \text{(since } \mu \text{ is invariant under } T_3) \\ &= H_\mu(\xi_X \big| T_3^n \mathscr{A}_1). && \text{(by (9.14))} \end{aligned}$$

It follows that if

$$\mathscr{A} = \bigvee_{i=1}^{\infty} \bigvee_{j=-\infty}^{\infty} T_2^{-i} T_3^{-j} \xi_X$$

denotes the σ-algebra determined by the coordinates in the half-plane

$$\{(m,n) \in \mathbb{Z}^2 \mid m > 0\},$$

then by continuity of entropy (Proposition 2.14),

$$h_\mu(T_2) = H_\mu(\xi_X \big| \mathscr{A}). \tag{9.15}$$

We refer to Exercise 9.3.1 for a tempting but wrong argument to prove Theorem 9.9 using this formula (see also the solution on page 332).

The next step is to identify the $\mathscr{A}$-atoms more closely, for which we will use some material from Appendix C.

Lemma 9.13 (Atoms are cosets). *$G = [0]_{\mathscr{A}}$ is a closed subgroup of X. Moreover, $[x]_{\mathscr{A}} = G + x$ for $x \in X$.*

Proof. The projection onto the coordinates in the half-space

$$\{(m,n) \in \mathbb{Z}^2 \mid m > 0\}$$

is a continuous homomorphism, so its kernel is a closed subgroup of X, and the pre-images of other points are the cosets of this kernel. □

The following provides a better understanding of the group G.

Lemma 9.14 (2-adic integers). *G is isomorphic to $\mathbb{Z}_2$, the 2-adic integers, and the isomorphism conjugates T_3 to multiplication by 3 on $\mathbb{Z}_2$.*

Proof. The isomorphism between G and $\mathbb{Z}_2$ may be described concretely as follows. For this, first recall that the invertible extension of $S_2 : \mathbb{T} \to \mathbb{T}$ is defined by

$$X_2 = \left\{x \in \mathbb{T}^{\mathbb{Z}} \mid x_{k+1} = 2x_k \text{ for all } k \in \mathbb{Z}\right\}$$

and define the subgroup

$$G_2 = \{x \in X_2 \mid x_k = 0 \text{ for all } k \geqslant 1\}.$$

A 2-adic integer

$$a = \sum_{m=0}^{\infty} a_m 2^m$$

with $a_m \in \{0,1\}$ for all $m \geqslant 0$, determines a point $g(a) \in G_2$ by setting

$$g(a)_k = 2^{k-1} \sum_{m=0}^{|k|} a_m 2^m \pmod 1$$

for all $k \leqslant 0$ and $g(a)_{(k,0)} = 0$ for all $k \geqslant 1$. Identify $\mathbb{Q}_2/\mathbb{Z}_2 \cong \mathbb{Z}[\frac{1}{2}]/\mathbb{Z} < \mathbb{R}/\mathbb{Z}$ with its image in $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Setting $g(a)_k = 2^{k-1}a$ modulo $\mathbb{Z}_2$ for all $k \in \mathbb{Z}$ gives an equivalent definition which allows us to see the properties of the map more clearly. In particular, $\mathbb{Z}_2 \ni a \longmapsto g(a) \in G_2$ is a group homomorphism, and so multiplication by 3 has the property $g(3a) = 3g(a)$. Using this we extend the map $g(a)$ to a map from $\mathbb{Z}_2$ to $G \subseteq X$ by setting

$$g(a)_{(k,\ell)} = 2^{k-1}3^{\ell}a$$

for all $a \in \mathbb{Z}_2$ and $(k,\ell) \in \mathbb{Z}^2$ (which is again to be understood using the inclusion $\mathbb{Q}_2/\mathbb{Z}_2 \cong \mathbb{Z}[\frac{1}{2}]/\mathbb{Z} \subseteq \mathbb{T}$).

We note that $g(a) = 0 \in G$ forces $a_k = 0$ for all $k \geqslant 0$ (for example, by looking at the coordinate $g(a)_{(-k,0)}$), so that the map is injective. Moreover, $g(\cdot) : \mathbb{Z}_2 \to G$ is also surjective. In fact if $g \in G$ we find $a_0 \in \{0,1\}$ with $g_{(0,0)} = \frac{a_0}{2}$ (since $2g_{(0,0)} = 0$ in $\mathbb{T}$). Having already found a_k for $k = 0, \dots, m-1$ we find $a_m \in \{0,1\}$ with

$$g_{(-m,0)} = \frac{a_m}{2} + 2^{-(m+1)} \sum_{k=0}^{m-1} a_k 2^k$$

(using $2g_{(-m,0)} = g_{(-m+1,0)}$ and the same formula for $g_{(-m+1,0)}$). The sequence $a_m \in \{0,1\}$ for $m \geqslant 0$ then defines an element $a \in \mathbb{Z}_2$ with $g = g(a)$. □

Since the atom $[x]_{\mathscr{A}}$ supports the measure $\mu_x^{\mathscr{A}}$ and $[x]_{\mathscr{A}}$ is a coset of G, we can define for almost every $x \in X$ a measure ν_x on G by

$$\nu_x(B) = \mu_x^{\mathscr{A}}(x + B)$$

for all measurable $B \subseteq G$, or equivalently

$$\nu_x = \mu_x^{\mathscr{A}} - x,$$

where we simply write translation by $-x$ instead of the push forward by translation by $-x$. This definition of ν_x is a precursor of what we will call a *leafwise measure* and the reader should interpret ν_x as a description of the atom of x with respect to μ and the σ-algebra $\mathscr{A}$ viewed from the point x, so that $0 \in G$ corresponds to x itself.

Proposition 9.15 (Pinsker measurability — the key to invariance). *The map $x \mapsto \nu_x$ is measurable with respect to the Pinsker σ-algebra of T_2 (or equivalently of T_3).*

PROOF. Fix some $n \geqslant 1$. Using the isomorphism between G and $\mathbb{Z}_2$ we define the subset $B_n(k) \subseteq G$ as the image of the ball $\{a \in \mathbb{Z}_2 \mid \|a - k\| \leqslant 2^{-n}\}$, where $k \in \mathbb{Z}$ or equivalently just $k \in \{0, \dots, 2^n - 1\}$.

It is enough to show that $g_k(x) = \nu_x(B_n(k))$ is Pinsker measurable for any n and k as above. We claim that g_k is in fact periodic with respect to T_3. For this, note that multiplication by 3 is invertible modulo 2^n and recall that $T_3^{-1}\mathscr{A} = \mathscr{A}$. By Lemma 2.5 applied to $\phi = T_3^{-1}$

$$\begin{aligned} g_k(T_3^{-1}x) &= \mu_{T_3^{-1}x}^{\mathscr{A}}\left(T_3^{-1}x + B_n(k)\right) && \text{(by definition of } g_k \text{ and } \nu_x) \\ &= \mu_x^{T_3\mathscr{A}}\left(x + T_3 B_n(k)\right) && \text{(by Lemma 2.5)} \\ &= \mu_x^{\mathscr{A}}\left(x + B_n(3k)\right) = g_{3k}(x). \end{aligned}$$

There exists some[†] m with $3^m \equiv 1$ modulo 2^n and in particular

$$B_n(k) = B_n(3^m k).$$

With this value of m we then also have

$$g_k(T_3^{-m}x) = g_{3^m k}(x) = g_k(x),$$

which gives the claimed periodicity. By Exercise 2.4.2 (see also the hint on page 329) and Corollary 9.12, ν_x is Pinsker measurable with respect to T_3 and T_2. □

Proposition 9.16 (Invariance). *For almost every $x \in X$, the measure ν_x is the Haar measure m_{G_x} of a closed subgroup $G_x \leqslant G$.*

PROOF. Since $T_3^k \bigvee_{i=-\infty}^{\infty} T_2^{-i}\xi_X \nearrow \mathscr{B}_X$ as $k \to \infty$, Theorem 2.29 implies that

[†] Here we are indeed using the fact that $\gcd(2,3) = 1$; for natural numbers with a common factor a different argument would be needed, and this may be found in the work of Johnson [91].

$$\mathscr{S} = \bigvee_{k \geqslant 0} T_3^k \bigcap_{n=0}^{\infty} \bigvee_{i=n}^{\infty} T_2^{-i}(\xi_X)$$

is the Pinsker σ-algebra modulo μ. Since $\mathscr{S} \subseteq \mathscr{A}$ we see from Proposition 9.15 that ν_x is $\mathscr{A}$-measurable, possibly after removing a μ-null set N from the space. Certainly $\mu_x^{\mathscr{A}}$ is $\mathscr{A}$-measurable (by Theorem 2.2 resp. [52, Th. 5.14]). This gives

$$\mu_x^{\mathscr{A}} = \mu_y^{\mathscr{A}},$$
$$\nu_x = \nu_y,$$

and

$$\mu_x^{\mathscr{A}}(N) = 0.$$

for $x, y \in [x]_{\mathscr{A}} \smallsetminus N$, where we possibly have to enlarge the null set N for the last property.

By definition of ν_x, this then implies

$$\mu_x^{\mathscr{A}} - x = \nu_x = \nu_y = \mu_x^{\mathscr{A}} - y = \nu_x - x + y,$$

for all $x, y \in [x]_{\mathscr{A}} \smallsetminus N$. We also have $\nu_x(N - x) = 0$, and so for ν_x-almost every $g \in G$ we have $y = x + g \in [x]_{\mathscr{A}} \smallsetminus N$ and so

$$\nu_x = \nu_x + g$$

for ν_x-almost every $g \in G$. This implies that $G_x = \operatorname{Supp}(\nu_x)$ is a closed subgroup of G and ν_x is the Haar measure on G_x. □

Proof of Theorem 9.9. By assumption $h_\mu(T_2) > 0$ and using (9.15) we see that

$$B = \{x \mid I_\mu(\xi_X \big| \mathscr{A})(x) > 0\}$$

has positive measure. Using the notation of the proof of Lemma 9.14 we have $[x]_{\xi_X \vee \mathscr{A}} = x + B_1(0)$ and hence

$$I_\mu(\xi_X \big| \mathscr{A})(x) = -\log \mu_x^{\mathscr{A}}([x]_{\xi_X \vee \mathscr{A}}) = -\log \nu_x(B_1(0)) = -\log g_0(x).$$

Hence $g_0(x) < 1$ for $x \in B$ and we see that the closed subgroup G_x in Proposition 9.16 contains (where defined) an element $g = g(a)$ with

$$a = \sum_{n=0}^{\infty} a_n 2^n \in \mathbb{Z}_2$$

and $a_0 = 1$. However, $\mathbb{Z}_2$ (and hence G) is topologically generated by any such element (since $\mathbb{Z}/2^{n+1}\mathbb{Z}$ is generated by any odd number). Thus $G_x = G$

for almost every $x \in B$. Since $g_0(x) = g_0(T_3 x)$ we know that the set B is T_3-invariant.

We claim that $G_x = G$ for almost every x. To prove this, it is enough to show that $B = X$. For almost every $x \in B$ we have $\mu_x^{\mathscr{A}} = x + m_G$ (and note that this property characterizes membership of B). By the double conditioning formula (Proposition 2.4, which in our case is somewhat easier as ξ_X is a finite partition) we have

$$\mu_x^{T_2 \mathscr{A}} = \mu_x^{\mathscr{A} \vee \xi_X} = \left(\mu_x^{\mathscr{A}}\right)_x^{\xi_X} = x + m_{T_2 G}$$

for almost every $x \in B$. Using the push-forward formula for conditional measures (Lemma 2.5 for $\phi = T_2^{-1}$) we obtain

$$\mu_{T_2^{-1} x}^{\mathscr{A}} = \left(T_2^{-1}\right)_* \mu_x^{T_2 \mathscr{A}} = T_2^{-1} x + m_G$$

for almost every $x \in B$, but this shows $T_2^{-1} x \in B$. Thus B is T_2-invariant, T_3-invariant, and has positive measure. By the assumption of ergodicity, this implies that $B \underset{\mu}{=} X$.

Now there are many ways to finish the argument. For example, the argument above gives $\mu_x^{\mathscr{A}} = x + m_G$ and hence

$$I_\mu(\xi_X \big| \mathscr{A})(x) = -\log(x + B_1(x)) = \log 2$$

almost surely. From this we obtain $h_\mu(T_2) = \log 2$, and so the original measure on $\mathbb{T}$ must have been $m_{\mathbb{T}}$. □

9.3.1 Semigroups of Polynomial Density and Positive Entropy

In this section we apply the powerful local entropy methods from Chapter 3 to extend Theorem 9.9 to classify measures on $\mathbb{T}$ that are invariant under the action of a large multiplicative sub-semigroup of $\mathbb{N}$, by explaining the method of Bourgain and Lindenstrauss [22] for establishing positive entropy in this much simpler (in comparison, a toy) situation. This exposition of the idea is taken from work of Einsiedler and Fish [45].

Let $S \subseteq \mathbb{N}$ be a multiplicative semigroup. Recall that if $S \subseteq a^{\mathbb{N}}$ for some $a > 1$, then S is called *lacunary* and in this case there is a multitude of S-invariant probability measures (under the natural action of multiplication by elements of S on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$) and there are many S-invariant closed infinite subsets of $X = \mathbb{R}/\mathbb{Z}$. On the other hand, if $S \not\subseteq a^{\mathbb{N}}$ for any $a > 1$, then S is called *non-lacunary* and as we have seen in Theorem 9.3, there are very few closed invariant subsets under the action of a non-lacunary semigroup.

Let us now ask Furstenberg's question regarding $\times 2$, $\times 3$-invariant measures in the right generality: Must a probability measure that is invariant and ergodic for a non-lacunary semigroup be of finite support or be Lebesgue measure? It is straightforward (see Exercise 9.3.4) to see that for $S = \mathbb{N}$ there are only two possible S-invariant and ergodic probability measures, namely the Dirac measure δ_0 and the Lebesgue measure $m_{\mathbb{T}}$. However, as with $S = \langle 2, 3\rangle$ for general non-lacunary semigroups the conjecture is still open. Let us state Johnson's generalization of Rudolph's Theorem† from [91].

Theorem 9.17 (Johnson). *Let S be a non-lacunary semigroup in $\mathbb{N}$, and let μ be an S-invariant ergodic probability measure on $\mathbb{T}$. If $h_\mu(s) > 0$ for some $s \in S$, then $\mu = m_{\mathbb{T}}$ is Lebesgue measure.*

One may ask whether it is possible to give stronger conditions on S which would allow a complete classification of S-invariant ergodic probability measures without the entropy hypothesis (as in the trivial case $S = \mathbb{N}$ considered in Exercise 9.3.4). This can indeed be done for the following class of semigroups.

Definition 9.18. A semigroup $S \subseteq \mathbb{N}$ has *polynomial density with exponent* $\alpha > 0$ if

$$|S \cap [1, M]| \geqslant M^{\alpha} \tag{9.16}$$

for all sufficiently large M.

Theorem 9.19. *Let S be a semigroup of polynomial density with positive exponent. Then any S-invariant and ergodic probability measure on $\mathbb{T}$ is either supported on a finite set of rational points, or is the Lebesgue measure.*

We will prove Theorem 9.19 by establishing, under this stronger assumption on S, the entropy hypothesis in Theorem 9.17.

Proof of Theorem 9.19. Let μ be an S-invariant ergodic probability measure. If $\mu(\mathbb{Q}/\mathbb{Z}) > 0$ then by ergodicity we must have $\mu(\mathbb{Q}/\mathbb{Z}) = 1$, and so μ must be supported on a finite set since each point in $\mathbb{Q}/\mathbb{Z}$ has a finite orbit under S. Suppose therefore that μ is an S-invariant ergodic probability measure on $\mathbb{T}$ with $\mu(\mathbb{Q}/\mathbb{Z}) = 0$.

Fix some $s_0 \in S \smallsetminus \{1\}$, write $T(x) = s_0 x$. Then we will show below (under the assumption $\mu(\mathbb{Q}/\mathbb{Z}) = 0$) that $h_\mu(s_0) \geqslant \delta = \frac{\alpha \log s_0}{5}$. This will imply the theorem by the work of Rudolph and Johnson. Let

$$\xi = \left\{[0, \tfrac{1}{s_0}), [\tfrac{1}{s_0}), \dots, [\tfrac{s_0-1}{s_0}, 1)\right\}$$

be the partition corresponding to fixing the first digit in the s_0-ary expansion of real numbers $x \in [0, 1) \cong \mathbb{T}$. Notice that $\bigvee_{i=0}^{n-1} T^{-i}\xi$ is the partition

† We note that if $2, 3 \in S$ then it is very easy to prove Theorem 9.17 as a corollary of Rudolph's theorem (Theorem 9.9).

corresponding to the first n digits in the s_0-ary expansion, and so comprises intervals of length $\frac{1}{s_0^n}$.

Now recall that the Shannon–MacMillan–Breiman theorem (Theorem 3.1) states that

$$-\frac{1}{n}\log\mu\left([x]_{\bigvee_{i=0}^{n-1}T^{-i}\xi}\right)\longrightarrow h_{\mu_x^{\mathscr{E}}}(T,\xi)$$

almost surely, where $\mu_x^{\mathscr{E}}$ is the ergodic component of μ with respect to T.

We assume now that the convergence takes place for $x\in\mathbb{T}$ and that

$$h_{\mu_x^{\mathscr{E}}}(T,\xi)<\delta.$$

Then for large enough $n\geqslant n_0$ we would have

$$-\frac{1}{n}\log\mu\left([x]_{\bigvee_{i=0}^{n-1}T^{-i}\xi}\right)<\delta,$$

and so

$$\mu\left(B(x,s_0^{-n})\right)\geqslant \mathrm{e}^{-n\delta},$$

where

$$B(x,s_0^{-n})=(x-s_0^{-n},x+s_0^{-n})\supseteq[x]_{\bigvee_{i=0}^{n-1}T^{-i}\xi}.$$

We define $M=M(n)$ by

$$M=2\mathrm{e}^{n\delta/\alpha},\tag{9.17}$$

and we may assume that (9.16) holds whenever $n\geqslant n_0$.

Recall that every $s\in S$ preserves the measure μ, so $\mu\left(T_s^{-1}B\right)=\mu(B)$ for any Borel set $B\subseteq\mathbb{T}$ and where $T_s(x)=sx$ for all $x\in\mathbb{T}$. This clearly implies (using $B=T_s(I)=sI$) that $\mu(sI)\geqslant\mu(I)$ for any Borel set I and $s\in S$. We now apply this to the interval

$$I=B(x,s_0^{-n})$$

to obtain

$$\mu\left(sB(x,s_0^{-n})\right)\geqslant \mathrm{e}^{-n\delta}.$$

Using this together with (9.17) and the assumption that S has polynomial density with exponent α we see that

$$\sum_{s\in S\cap[1,M]}\mu\left(sB(x,s_0^{-n})\right)\geqslant M^{\alpha}\mathrm{e}^{-n\delta}>1,$$

at least once n (and hence M) are sufficiently large, so let us say this holds also for $n\geqslant n_0$. Therefore, there must be distinct elements $s,s'\in S\cap[1,M]$ with

$$sB(x,s_0^{-n})\cap s'B(x,s_0^{-n})\neq\varnothing.$$

Of course, this overlapping must be understood in $\mathbb{T}=\mathbb{R}/\mathbb{Z}$, so if we identify $x\in\mathbb{T}$ with the corresponding element $x\in[0,1)\subseteq\mathbb{R}$, then we have

$$s(x+v) = s'(x+v') + k,$$

where $|v|, |v'| < s_0^{-n}$, and $k \in \mathbb{Z}$. Thus

$$x = \frac{s'v' - sv}{s - s'} + \frac{k}{s - s'},$$

so by (9.17) and the definition of $\delta = \frac{\alpha \log s_0}{5}$,

$$\left| \frac{s'v' - sv}{s - s'} \right| \leqslant 2Ms_0^{-n} = 2M\mathrm{e}^{-n \log s_0} \ll MM^{-\frac{\alpha}{\delta} \log s_0} = M^{1-5} = M^{-4}$$

and

$$|s - s'| < M.$$

This already should be surprising — the real number x (about which we only assumed $h_{\mu_x^{\mathscr{E}}}(T, \xi) < \delta$) has a rational approximation of the shape

$$\left| x - \frac{p_n}{q_n} \right| \ll M^{-4}$$

with denominator $q_n < M_n = 2\mathrm{e}^{n\delta/\alpha}$.

We now apply the argument used above with $n+1$ in place of n; note that $M_n \leqslant M_{n+1} \ll M_n$ and so we find a rational approximation

$$\left| x - \frac{p_{n+1}}{q_{n+1}} \right| \ll M_n^{-4}$$

with $q_{n+1} < M_{n+1} \ll M_n$. Together these two approximations give

$$\left| \frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}} \right| \ll M^{-4}. \tag{9.18}$$

On the other hand, the left-hand side of (9.18) is either zero, or has

$$\left| \frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}} \right| \geqslant \frac{1}{q_n q_{n+1}} \geqslant M^{-2}.$$

Assuming n is large enough (increasing n_0 further if necessary we again suppose $n \geqslant n_0$ is sufficient), the latter inequality gives a contradiction to (9.18). Hence we must have $\frac{p_n}{q_n} = \frac{p_{n+1}}{q_{n+1}}$ for all $n \geqslant n_0$, and therefore

$$\left| x - \frac{p_{n_0}}{q_{n_0}} \right| \ll M_n$$

for all $n \geqslant n_0$. Letting $n \to \infty$ we obtain

$$x = \frac{p_{n_0}}{q_{n_0}} \in \mathbb{Q} \cap [0,1).$$

To summarize, we have shown for almost every x that $h_{\mu_x^{\mathscr{E}}}(T,\xi) < \delta$ implies that $x \in \mathbb{Q} \cap [0,1)$. Since

$$\mu(\mathbb{Q} \cap [0,1)) = 0$$

we deduce that almost every ergodic component of μ has $h_{\mu_x^{\mathscr{E}}}(T,\xi) \geqslant \delta > 0$ as required. Therefore, Theorem 9.17 implies Theorem 9.19. $\square$

Exercises for Section 9.3

Exercise 9.3.1. After the definition of the σ-algebra $\mathscr{A}$ in the invertible extension X for S_2, S_3 and having obtained (9.15) it is tempting to finish the proof by a much faster argument as follows:
(a) Show that $I_\mu(\xi_X \big| \mathscr{A})(T_3 x) = I_\mu(\xi_X \big| \mathscr{A})(x)$.
(b) Use ergodicity to conclude that $I_\mu(\xi_X \big| \mathscr{A})$ is constant.
(c) Show that $I_\mu(\xi_X \big| \mathscr{A}) = \log 2$ everywhere and conclude that the original measure on $\mathbb{T}$ must have been the Lebesgue measure.
Which part of this outline is not correct?

Exercise 9.3.2. [45] Let μ be an S_3-invariant and ergodic probability measure on $\mathbb{T}$ with positive entropy.
(a) Prove that any weak* limit of $\frac{1}{n}\sum_{j=0}^{n-1}(S_2)_*^j\mu$ is of the form $cm_{\mathbb{T}}+(1-c)\nu$ for some $c > 0$ and $\nu \in \mathscr{M}(\mathbb{T})$.
(b) Prove that μ-almost every point $x \in \mathbb{T}$ has a dense orbit under S_2.

Exercise 9.3.3. Show how to choose, for any $n \geqslant 1$ with $\gcd(n,6) = 1$, a subset S_n of $\{0, \frac{1}{n}, \dots, \frac{n-1}{n}\}$ that is invariant under both $x \mapsto 2x \pmod 1$ and $x \mapsto 3x \pmod 1$ with $|S_n| \geqslant n^\alpha$ for some fixed $\alpha > 0$. Show that for any $\varepsilon > 0$ there exists some $N(\varepsilon)$ such that for $n \geqslant N(\varepsilon)$ the set S_n is ε-dense in $[0,1)$.

Exercise 9.3.4. Show Theorem 9.19 directly for the case $S = \mathbb{N}$ (or for any positive-density semigroup $S \subseteq \mathbb{N}$).

Notes to Chapter 9

(39) (Page 237) The treatment here follows Boshernitzan [20] and Furstenberg [64, Sect. IV] closely.
(40) (Page 240) The argument in Section 9.2 comes from a paper of Einsiedler and Ward [50] and uses half-space entropy ideas from Kitchens and Schmidt [104] and Einsiedler [44]; similar ideas are applied to show isomorphism rigidity for certain $\mathbb{Z}^d$-actions by toral automorphisms in [51], later generalized by Bhattacharya [13]. Related work for $\mathbb{Z}^d$-actions by toral automorphisms has been done by Kalinin and Katok [96], where more refined information is found about joinings and the consequences of the existence of non-trivial joinings.
(41) (Page 243) Determining which subgroups generate in this sense is closely related to the expansive subdynamical structure of the system, which is introduced by Boyle and

Lind [27] and completely determined for commuting automorphisms of compact abelian groups by Einsiedler, Lind, Miles and Ward [46].

(42)(Page 248)In fact Ledrappier's "three dot" example has many invariant measures: Einsiedler [43] constructs uncountably many closed invariant sets, and uncountably many different invariant measures giving positive entropy, to any single map in systems of this sort.

(43)(Page 248) This is a special case of a more general result due to Schmidt [183], where it is shown that Haar measure is the unique "most mixing" invariant probability measure for a large class of systems of this sort.

(44)(Page 249) Other approaches to this result have been found by Feldman [58], Host [85], Lindenstrauss [126], Parry [162] and others. More recently, Bourgain, Lindenstrauss, Michel and Venkatesh [23] have found effective versions of both Rudolph's and Johnson's theorems, relating the size of the assumed positive entropy to effective properties of the invariant measure.

(45)(Page 262) In fact a much stronger statement holds: Johnson and Rudolph [92] showed that the weak* limit is equal to $m_{\mathbb{T}}$.