A Survey of Shift Spaces
Fall 2016

Doğan Çömez

Contents:

1. The n-fold Map T_n.

A topological dynamical system is a pair (X, T), where X is a topological space and $T : X \to X$ is a map compatible with the topology. Typically X is a metric space and T is a continuous map, or a map with finitely many discontinuities. Among such dynamical systems the n-fold (or n-ary) system has some special properties. In this section we will define and provide some basic features of the n-fold system that will pave way to an important class of symbolic dynamical systems called shift spaces.

Let $X = [0, 1)$ with the usual Euclidean metric and, for any $n \geq 2$, let $T_n : [0, 1) \to [0, 1)$ be defined by $T_n(x) = nx(\text{mod} n),\ x \in [0, 1)$. Hence,

$$T_n(x) = \begin{cases} nx \text{ if } 0 \leq x < \dfrac{1}{n} \\ nx - 1 \text{ if } \dfrac{1}{n} \leq x < \dfrac{2}{n} \\ \cdots \\ nx - (n-1) \text{ if } \dfrac{(n-1)}{n} \leq x < 1. \end{cases}$$

The system $([0, 1), T_n)$ is called the *n-fold (or n-ary) system.* When $n = 2$ it is called the *doubling system (map).* Some simple properties of the n-fold map are as follows:

1. T_n has n branches, i-th branch is maps the interval $[\frac{i-1}{n}, \frac{i}{n})$ to $[0, 1),\ 1 \leq i \leq n$.
2. T_n has periodic points of all orders. (A point $x \in X$ is called a *periodic point of period* n of a system (X, T) if $T^n x = x$ and n is the smallest such positive integer.) Periodic points of period one are also called *fixed points.* For the n-fold map, 0 is the only fixed point; and, among others, 1/3 has period 2, 4/7 has period 3, 1/5 has period 4, . . . etc. Indeed, every rational point is a periodic point for the n-fold map. The converse is also valid.

3. T_n is *transitive*, that is, there exists $x \in [0,1)$ such that the set $\{T^k x\}_{k=0}^{\infty}$ is dense in $[0,1)$. In a dynamical system (X,T) the set $O_T(x) = \{T^k x\}_{k=0}^{\infty}$ is called as the *orbit* of the point $x \in X$. Since rational points are those with periodic orbits by (3) above, irrational points are not periodic; hence, the transitive points of the n-fold map are among irrationals. We will not exhibit an example of a transitive point of this system here, it'll be left to the next section when we investigate the shift spaces. If $\overline{O_T(x)} = X$ for all $x \in X$ in a dynamical system (X,T), it is called a *minimal* system. Clearly, $([0,1),T_n)$ is not minimal.

For a dynamical system (X,T), if X is endowed with a measure μ (with the associated sigma algebra of measurable subsets of X), the triple (X,μ,T) is also called a *measurable* dynamical system. If $\mu(E) = \mu(T^{-1}E)$ for all measurable E, then we say that T *preserves* μ (or T is *measure preserving*) and μ is T*-invariant.* If $[0,1)$ is endowed with the Lebesgue measure m, then T_n preserves m. For, it is enough to show the $m(T_n(I)) = m(I)$ for any interval $I \subset [0,1)$. If $I = [a,b)$, then $T^{-1}I = \cup_{i=0}^{n-1}[\frac{a}{n}+\frac{i}{n},\frac{b}{n}+\frac{i}{n})$. Hence,

$$m(T_n^{-1}[a,b)) = \sum_{i=0}^{n-1} m([\frac{a}{n}+\frac{i}{n},\frac{b}{n}+\frac{i}{n})) = b-a = m([a,b)).$$

The Lebesgue measure m is not the only T_n-invariant measure on $[0,1)$, indeed there are uncountably many measures μ on $[0,1)$ (equivalent to m) that are T_n-invariant. Again, we will exhibit such measures in the next section. If μ is the only T-invariant measure of a dynamical system (X,μ,T), then the system is called *uniquely ergodic.* Thus, $([0,1),m,T_n)$ is not uniquely ergodic.

A map T of a dynamical system (X,μ,T) is called *ergodic* if any measurable set E with $E = T^{-1}E$ (called T*-invariant set*) has either measure 0 or full measure. It turns out that T_n is ergodic. In order to prove the ergodicity of T_n, we need a few technicalities.

A class $\mathcal{C}$ of subintervals of $[0,1)$ is called a *covering class* for $[0,1)$ if every subinterval of $[0,1)$ is a countable disjoint union of elements from $\mathcal{C}$.

Lemma. (Knopp's Lemma) Let $B \subset [0,1)$ be a measurable set and $\mathcal{C}$ be a covering class for $[0,1)$. If

$$(*) \qquad \forall A \in \mathcal{C},\ m(A \cap B) \geq \gamma m(A),$$

where $\gamma > 0$ is independent of A, then $m(B) = 1$.

Proof. We will prove the statement by contradiction. Assume that $m(B) < 1$, i.e., $m(B^c) > 0$. Since B is a measurable set, $B = C \cup D$, where C is a Borel set with $m(B) = m(C)$ and $m(D) = 0$. Thus

$m(C^c) > 0$ as well. Now, given $\epsilon > 0$, there a set E_ϵ which is a finite disjoint union of open intervals in $[0, 1)$ such that $m(C^c \triangle E_\epsilon) < \epsilon$. Hence, E_ϵ is a countable disjoint union of elements from $\mathcal{C}$. Therefore, by (*), it follows that $m(C \cap E_\epsilon) \geq \gamma m(E_\epsilon)$. Then,

$$m(C^c \triangle E_\epsilon) \geq m(C \cap E_\epsilon) \geq \gamma m(E_\epsilon) \geq \gamma m(C^c \cap E_\epsilon) > \gamma(m(E_\epsilon) - \epsilon).$$

This implies that $\gamma(m(E_\epsilon) - \epsilon) < m(C^c \triangle E_\epsilon) < \epsilon$. Therefore, we must have $\gamma m(C^c) < \epsilon + \gamma\epsilon$. Since ϵ is arbitrary, this implies that $m(C^c) = 0$, contradiction. ∎

Fact.1 T_n is ergodic.

Proof. We will prove the that T_2 is ergodic for simplicity, ergodicity of T_n follows the same lines. Let B be a T_2-invariant measurable subset of $[0, 1)$. Consider dyadic intervals $D_{n,k} = [\frac{k}{2^n}, \frac{k+1}{2^n})$, where n is a positive integer and $k = 0, 1, 2, \ldots, 2^n - 1$. Then the collection $\mathcal{C}$ of all dyadic intervals is a covering class for $[0, 1)$. Also, $m(D_{n,k}) = 2^{-n}$ and $T_2^n(D_{n,k}) = [0, 1)$ for each $k = 0, 1, 2, \ldots, 2^n - 1$. Also, it follows by induction that, for any measurable set A,

$$m(T_2^{-n} A \cap D_{n,k}) = 2^{-n} m(A) = m(A)m(D_{n,k}), \; k = 0, 1, 2, \ldots, 2^n - 1.$$

Since B is T_2-invariant, this implies that

$$m(B \cap D_{n,k}) = m(B)m(D_{n,k}), \text{ for all } n > 0, \; k = 0, 1, 2, \ldots, 2^n - 1.$$

Taking $\gamma = m(B) > 0$, then it follows that $m(B \cap C) \geq \gamma m(C)$ for any $C \in \mathcal{C}$. Hence, by Knopp's Lemma, $m(B) = 1$, i.e., T_2 is ergodic. ∎

Remark. A map T of a dynamical system (X, μ, T) is called *totally ergodic* if T^m is ergodic for all $m \geq 1$. Since T_n is ergodic and since $T_n^m = T_{nm}$, it follows that T_n is totally ergodic.

A map T of a dynamical system (X, μ, T) is called *mixing* if for all measurable sets A and B, $\lim_{n\to\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B)$. Observe that in the proof of ergodicity of T_2 we have shown that, for any measurable set A,

$$m(T_2^{-n} A \cap D_{n,k}) = m(A)m(D_{n,k}).$$

Since dyadic intervals area generating σ-algebra for measurable sets, it follows that

$$\lim_n m(T_2^{-n} A \cap B) = m(A)m(B),$$

for all measurable sets A and B. Thus T_2 is mixing. By the same proof (adapted to n-adic intervals), it follows that T_n is mixing.

There is an intermediate property between ergodicity and mixing. A map T of a dynamical system (X, μ, T) is called *weakly mixing* if for all measurable sets A and B,

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

By the Fact .2 below we have that

$$\text{mixing} \Rightarrow \text{weak mixing} \Rightarrow \text{ergodicity},$$

hence, T_n is weakly mixing.

Fact.2 For a dynamical system, mixing $\Rightarrow$ weak mixing $\Rightarrow$ ergodicity.
Proof. By the Ergodic Theorem (see [5]) T is ergodic if and only if for all measurable A, B,

$$(1) \qquad \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}A \cap B) = \mu(A)\mu(B).$$

Also, it is known that, for any sequence (a_n) of real numbers,

$$(2) \qquad \lim a_n = 0 \;\Rightarrow\; \lim_n \frac{1}{n} \sum_{k=1}^{n} |a_k| = 0 \;\Rightarrow\; \lim_n \frac{1}{n} \sum_{k=1}^{n} a_k = 0.$$

From (1) and (2) the assertion follows. ∎

Two (topological) dynamical systems (X, T) and (Y, S) are called *conjugate*, and denoted by $X \equiv Y$, if there exists a homeomorphism $\phi : X \to Y$ such that $\phi \circ T = S \circ \phi$. Let $\Sigma_2 = \{0, 1\}^{\mathbb{N}}$, the space of all sequences of 0's and 1's. Define a metric $d : \Sigma_2 \times \Sigma_2 \to \mathbb{R}^+$ by

$$d((a_k), (b_k)) = \sum_{k=1}^{\infty} \frac{|a_k - b_k|}{2^k}, \quad (a_k), (b_k) \in \Sigma_2.$$

Then (Σ_2, d) is a metric space. Since $\{0, 1\}$ is a compact metric space, by Tychonoff's Theorem, so is (Σ_2, d). Now, define a map, called as *the (left) shift map*, $\sigma : \Sigma_2 \to \Sigma_2$ by $\sigma((a_k)) = (a_{k+1})$. The the dynamical system (Σ_2, σ) is called the *2-shift space* (or *shift space*).

Remarks. 1. The metric space (Σ_2, d) has topological dimension zero.
2. Σ_2 is homeomorphic to Cantor set.
3. The shift map $\sigma : \Sigma_2 \to \Sigma_2$ is two-to-one map (in the case $\sigma : \Sigma_n \to \Sigma_n$, it is an n-to-one map).
4. If $\Sigma_n^* = \{0, 1, \ldots, n-1\}^{\mathbb{Z}}$, then the map $\sigma : \Sigma_n^* \to \Sigma_n^*$ is a one-to-one, onto and homeomorphism.

It is well-known that any $x \in [0, 1)$ can be represented in *binary expansion*

$$x = \sum_{k=1}^{\infty} \frac{a_k}{2^k} = 0.a_1a_2a_3 \ldots a_k \ldots,$$

where $a_k \in 0, 1$. This representation is unique except for countably many (rational) numbers, i.e., for those with binary expansion of the form $*1000\ldots0\ldots$, which also has binary expansion of the form $*0111\ldots1\ldots$.

One can remedy this non-uniqueness by accepting the second expansion only. Observe that, for any $x \in [0,1)$ with binary expansion $(a_k)_k$,

$$T_2(x) = 2(\sum_{k=1}^{\infty} \frac{a_k}{2^k})(\text{mod}2) = \sum_{k=2}^{\infty} \frac{a_k}{2^{k-1}}$$
$$= \sum_{k=1}^{\infty} \frac{a_{k+1}}{2^k} = 0.a_2a_3a_4 \ldots a_{k+1} \ldots .$$

Thus, $T_2(0.a_1a_2a_3 \ldots a_k \ldots) = 0.a_2a_3a_4 \ldots a_{k+1} \ldots$. Now, if we define the map $\phi : [0,1) \to \Sigma_2$ by $\phi(x) = (a_k)_{k=1}^{\infty}$, where $x = 0.a_1a_2a_3 \ldots a_k \ldots$, then ϕ is a bijection and a continuos map with continuous inverse. Furthermore,

$$\begin{aligned}\phi(T_2x) &= \phi(0.a_2a_3a_4 \ldots a_{k+1} \ldots) = (a_2a_3a_4 \ldots a_{k+1} \ldots)\\ &= \sigma(a_1a_2a_3 \ldots a_k \ldots) = \sigma(\phi(0.a_1a_2a_3 \ldots a_k \ldots))\\ &= \sigma(\phi(x)).\end{aligned}$$

Hence, $\phi \circ T_2 = \sigma \circ \phi$, i.e., $([0,1), T_2)$ and (Σ_2, σ) are conjugate spaces. Notice that, for any $m \geq 1$, $T_2^m(x) \equiv \sigma^m(a_k)$, where $x = 0.a_1a_2a_3 \ldots a_k \ldots$.

Let $\Sigma_n = \{0,1,2,\ldots,n-1\}^{\mathbb{N}}$, the space of all sequences of letters from the alphabet $\mathcal{A} = \{0,1,2,\ldots,n-1\}$. As in the case of Σ_2, this space is also a compact metric space with the metric

$$d((a_k),(b_k)) = \sum_{k=1}^{\infty} \frac{|a_k - b_k|}{n^k}, \quad (a_k),(b_k) \in \Sigma_n.$$

If we define the shift map as above, the system (Σ_n, σ) is called the *n-shift space* (or *shift space on $\mathcal{A}$*). Similarly, it follows that $([0,1), T_n)$ and (Σ_n, σ) are conjugate spaces with the same conjugacy map ϕ.

2. The Full n-shift Space Σ_n

In this section we will investigate the (symbolic) dynamical system (Σ_n, σ). Since it is conjugate to $([0,1), T_n)$, we will also obtain some (topological) properties of the n-fold map. In what follows, for simplicity, we will only deal with $\Sigma = \Sigma_2$.

As in the case of n-fold map, Σ also has periodic point of all orders. Indeed, it is very easy to exhibit such points: any point $\mathbf{x} \in \Sigma$ (i.e., any periodic sequence) of the form

$$\mathbf{x} = (a_1, a_2, \ldots, a_m, a_1, a_2, \ldots, a_m, a_1, \ldots)$$

is a periodic point of period m. Since $a_i \in \{0,1\}$, there are 2^{m-1} periodic points of period m (we omit the point $\mathbf{1} = (1,1,1,\ldots)$ since it is not in Σ). Consequently, there are only countable many periodic points of Σ and they form a dense set.

Since the n-fold map T_2 is transitive, so is σ. On the other hand it is much easier to exhibit a point with dense orbit for (Σ, σ) than for $([0,1), T_2)$: let

$$\mathbf{x} = (0100011011000001010011100101110111 0000\dots).$$

Observe that $\mathbf{x}$ contains strings of all possible combinations of 0's and 1's in any length.

Claim. $\overline{O_\sigma(\bar{x})} = \Sigma$. For, let $\mathbf{y} = (a_1, a_2, \dots, a_n, \dots) \in \Sigma$ be arbitrary and let $\epsilon > 0$ be given. Pick $n > 1$ such that $\frac{1}{2^n} < \epsilon$. Since $\mathbf{x}$ contains all possible strings of 0's and 1's of length n, applying σ to $\mathbf{x}$ sufficiently many times, say N, we have $\sigma^N \mathbf{x} = (a_1, a_2, \dots, a_n, *****)$. Then $\sigma^N \mathbf{x}$ and $\mathbf{y}$ agree at the first n coordinates; hence,

$$d(\sigma^N \mathbf{x}, \mathbf{y}) \leq \sum_{k=n+1}^{\infty} \frac{1}{2^k} = \frac{1}{2^n} < \epsilon.$$

Thus, the orbit of $\mathbf{x}$ is dense in Σ; hence, (Σ, σ) is transitive. Notice that since it has many periodic points, Σ is not minimal.

Remark. The set of transitive points of Σ form a dense G_δ-set in Σ.

These observations suggest that instead of studying $([0,1), T_n)$, one might study (Σ_n, σ) and transfer the properties that are invariant under conjugacy (i.e., topological invariants) to the former. How about measurable properties? For, one needs to put a (suitable) measurable structure on Σ and show that $([0,1), T_2)$ and (Σ, σ) are measure theoretically conjugate (i.e., *isomorphic*) as well.

Let $[a]_k := \{(x_n) \in \Sigma : x_k = a\}$, where $a \in \{0,1\}$. In general, let

$$[a_0, a_1, \dots, a_m]_k := \{(x_n) \in \Sigma : x_k = a_0, x_{k+1} = a_1, \dots, x_{k+m} = a_m\},$$

where $a_i \in \{0,1\},\ 0 \leq i \leq m$. Notice that

$$[a_0, \dots, a_m]_k = \{0,1\} \times \dots \times \{0,1\} \times \{a_0\} \times \{a_1\} \times \dots \times \{a_m\} \times \{0,1\} \times \dots;$$

hence, such subsets of Σ are called the (basic) *cylinder sets*.

Fact.3 The cylinder sets are both open and closed, and the collection $\mathcal{C}$ of cylinder sets forms a countable base for the topology of Σ.

Proof. Let $C = [a_0, \dots, a_m]_k$. First, we will show that C is open. For, if $\mathbf{x} \in C$, then $\mathbf{x} = (x_1, x_2, \dots, x_{k-1}, a_0, a_1, \dots, a_m, x_{k+m+1}, \dots)$. Let $\epsilon = \frac{1}{2^{k+m+1}}$. Then

$$\begin{aligned} \mathbf{y} \in B(\mathbf{x}, \epsilon) &\iff \sum_{i=1}^{\infty} \frac{|x_i - y_i|}{2^i} < \frac{1}{2^{k+m+1}} \\ &\iff x_i = y_i \text{ for } 1 \leq i \leq n + k + 1, \end{aligned}$$

which implies that $y_i = a_i$ for $k \leq i \leq n+k+1$; equivalently, $\mathbf{y} \in C$. Therefore, $B(\mathbf{x}, \epsilon) \subset C$; i.e., C is open.

Nest, let $(\mathbf{x}^n) \subset C$ be a sequence such that $\mathbf{x}^n \to \mathbf{x}$. So, for all $\epsilon > 0$, $\exists N > 0$ such that $r \geq N$ implies $d(\mathbf{x}^r, \mathbf{x}) < \epsilon$. Thus, for $r \geq N$, we must have $\sum_{i=1}^{\infty} \frac{|x_i - x_i^r|}{2^i} < \epsilon$. Therefore, if ϵ is small enough, this implies that for all except finitely many $r's$ we have $x_i^r = x_i,\ 1 \leq i \leq n+k+1$. Since such $\mathbf{x}^r$'s are in C, for $r \geq N$, we have

$$\mathbf{x} = (x_i) = (*, *, \ldots, *, a_0, a_1, \ldots, a_m, *, * \ldots) \in C.$$

Therefore, C is closed.

Clearly, there are countably many cylinder sets of the form C. Given any $\mathbf{x} \in \Sigma$ and $\epsilon > 0$, there is a cylinder set C such that $\mathbf{x} \in C \subset B(\mathbf{x}, \epsilon)$. Hence the collection $\mathcal{C}$ forms a countable base for the topology of Σ. ∎

Remarks. 1. By the Fact above, every open set is a countable union of cylinder sets. Since the cylinder sets are clopen and generate the topology of Σ, the metric space Σ has topological dimension zero.

2. The collection $\mathcal{C}$ also forms a covering class for Σ; namely, (i) $\emptyset \in \mathcal{C}$, and (ii) $\forall\ A_i \subset \Sigma,\ \exists\ \{C_n\} \in \mathcal{C}$ such that $A \subset \cup_n C_n$.

Since $\mathcal{C}$ is a covering class for Σ, any non-negative $\mathbb{R}^{\#}$-valued set function λ with domain $\mathcal{C}$ satisfying $\lambda(\emptyset) = 0$ gives rise to an outer measure μ^* by

$$\mu * (A) = \inf\{\sum_n \lambda(C_n) : \{C_n\} \subset \mathcal{C},\ A \subset \cup_n C_n\}.$$

This, in turn, defines a measure on the Borel σ-algebra of subsets of Σ. So, let's construct such a measure.

Let $\mathbf{p} = (p_0, p_1)$ such that $0 \leq p_0, p_1 \leq 1$ and $p_0 + p_1 = 1$. Define $\nu(\{0\}) = p_0$ and $\nu(\{1\}) = p_1$. Then $\nu : \{0, 1\} \to \mathbb{R}^+$ is a probability measure (on $\{0, 1\}$). Now, on the collection of cylinder sets $\mathcal{C}$ define a set function $\lambda : \mathcal{C} \to \mathbb{R}^+$ by

$$\lambda([a_0, a_1, \ldots, a_m]_k) = \Pi_{i=0}^m p_{a_i}, \quad a_i \in \{0, 1\}.$$

Let μ be the restriction of the outer measure μ^* generated by λ to the Borel σ-algebra of subsets of Σ. Then, by Kolmogorov Extension Theorem, the measure μ is uniquely defined. Notice that one can construct uncountable many such measures, called *Bernoulli measures*, and denoted by μ_p. Thus, $(\Sigma, \mathcal{B}, \mu_p)$ is a probability measure for each choice of $\mathbf{p}$.

Now, observe that if $C = [a_0, a_1, \ldots, a_m]_k$ is a cylinder set, then

$$\begin{aligned}\sigma^{-1}(C) &= \{(x_n) \in \Sigma :\ \sigma(x_n) \in C\} \\ &= \{(x_n) \in \Sigma :\ x_{k+i+1} = a_i,\ 0 \leq i \leq m\},\end{aligned}$$

which is also a cylinder set. Thus, σ is a measurable function on cylinder sets; hence, it is measurable on $\mathcal{B}$. Furthermore,

$$\mu_p(C) = \Pi_{i=0}^{m} p_{a_i} = \mu_p(\sigma^{-1}(C));$$

hence, σ preserves μ_p. It follows that the system $(\Sigma, \mathcal{B}, \mu_p, \sigma)$ is a measure preserving dynamical system!

3. Subshifts, Subshifts of Finite Type and Sofic Shifts

Let (Σ, σ) be the full shift and $X \subset \Sigma$ be a closed subset. If X is σ-invariant, i.e., $X = \sigma^{-1}(X)$, then the pair (X, σ) is called a *subshift of* (Σ, σ).

Examples. 1. $X = \{\mathbf{0}\}$ and $X = \Sigma$ are the smallest and largest (nonempty) subshifts (of (Σ, σ).

2. The dynamical system (X, σ), where $X = \{\mathbf{x}, \mathbf{y}\} \subset \Sigma$ with $\mathbf{x} = (0, 1, 0, 1, \ldots, 0, 10, 1 \ldots)$ and $\mathbf{y} = (1, 0, 1, 0, \ldots, 1, 0, 1, 0, \ldots)$, is a subshift of (Σ, σ).

3. The dynamical system (X, σ), where $X = \overline{O_\sigma(\mathbf{x})}$, for any $\mathbf{x} \in \Sigma$, is a subshift of (Σ, σ). In the case that $\mathbf{x}$ is a periodic element, the subshift $\overline{O_\sigma(\mathbf{x})} = O_\sigma(\mathbf{x})$ is called a *cyclic shift*.

Often, subshifts are characterized by the strings that are not contained in its elements.

Example.4 Let $X \subset \Sigma$ be defined by

$$X = \{(x_n)_n \in \Sigma : \ (x_n)_n \text{ does not contain the string } 11\}.$$

Consequently, for all $(x_n) \in X$, $\sigma(x_n)$ does not contain 11, either; hence, $\sigma(x_n) \in X$, which implies that X is σ-invariant. Clearly, X is closed. Hence, (X, σ) is a subshift of (Σ, σ), also known as the *golden mean shift (GMS)*.

Example.5 Let $Y \subset \Sigma$ be defined by

$$Y = \{(x_n)_n : \text{any two 1's in } (x_n) \text{ are separated by even number of 0's}\}.$$

Then Y is a closed subset of Σ (Exercise), and the image of nay $(x_n) \in Y$ under σ also satisfies the defining property of Y. Hence, (Y, σ) is a subshift of Σ, also known as the *even shift*.

Example.6 Let $\Pi \subset \Sigma$ be defined by

$$\Pi = \{(x_n)_n : \text{any two 1's in } (x_n) \text{ are separated by prime number of 0's}\}.$$

Then Π is a subshift of Σ known as *prime shift*.

Fact.4 a) The union of finitely many subshifts is also a subshift.
b) The intersection of any collection of subshifts is a subshift.

Proof. Exercise.

The strings that are not allowed to appear in elements of a subshift, as in Example.4, are called *forbidden words*. For the GMS the set of forbidden words is $F = \{11\}$, and for the even shift the set of forbidden words is $F = \{101, 10001, \dots, 10^{2n-1}1, \dots\}$. Subshifts whose forbidden word set is finite are called *subshifts of finite type (SFT)*. Hence, the GMS is a SFT, but the even shift is not.

Definition. Let (X, T) and (Y, S) be topological dynamical systems. A function $\phi : X \to Y$ is called a *factor map*, or *homomorphism*, if it is continuous and $\phi \circ T = S \circ \phi$. In this case, (Y, S) is called a *factor of* (X, T); and (X, T) is called a *extension* of (Y, S). Note that ϕ need not be invertible.

A subshift $X \subset \Sigma$ is called *sofic* if it s a factor of a SFT. The prime example of a sofic shift is the even shift. It is a factor of the GMS; the factor map is given as follows: for any $\mathbf{x} = (x_n)$ in GMS, let $\phi(\mathbf{x}) = \mathbf{y}$ be defined by $y_i = 1 - (x_i + x_{i+1}),\ i \geq 1$.

From the definition, every SFT is sofic, but converse is not valid. For instance, the even shift is sofic but not a SFT. There are also subshifts that are not sofic, the prime shift is such a non-sofic subshift.

Recall that a subshift need not have a periodic point (see Morse system below); however, it is a routine exercise to show that every sofic shift (SFT) has periodic points.

Subshifts inherit their topological and measurable properties from the full shift space (Σ, σ). However, not all properties of the full shift are inherited. For instance, not all subshifts are transitive: let $X = \{\mathbf{0}, \mathbf{a}\}$, where $\mathbf{a} = (01)^\infty$, then (X, σ) is a non transitive shift.

Now we will provide a criterion that ensures transitivity of subshifts.

Fact.5 A subshift $X \subset \Sigma$ is transitive if and only if for any ordered pair (u, v) of allowed words there exists an allowed word w such that the concatenation uwv is an allowed word.

Proof. Let X be transitive and (u, v) be an ordered pair of allowed words in X. Consider the cylinder sets $U = [u]_0$ and $V = [v]_0$. Then, by hypothesis, there exists n, which we can assume WLOG that greater than $l(u)$ =length of u, such that $U \cap \sigma^{-n}(V) \neq \emptyset$. Now, if $\mathbf{z} \in U \cap \sigma^{-n}(V)$, then the word $(z_0, z_1, \dots, z_{n+l(v)-1})$ is an allowed word. Now,

$$(z_0, z_1, \dots, z_{n+l(v)-1}) = u(z_{l(u)}, z_{l(u)+1}, \dots, z_{n-1})v = uwv$$

is allowed, i.e., uwv is allowed, where $w = (z_{l(u)}, z_{l(u)+1}, \dots, z_{n-1})$.

Conversely, assume that for any given ordered pair (u, v) of allowed words there exists an allowed word w such that uwv is an allowed word. We will use the equivalent characterization of transitivity (see [3], [5] or [8] for proof): (X, T) is transitive if and only if for every nonempty open sets U and V, there exists $n \geq 1$ such that $T^{-n}U \cap V \neq \emptyset$. Again, consider $U = [u]_0$ and $V = [v]_0$. By hypothesis, there exists an element $\mathbf{x} \in X$ that contains a word of the form

$$(u_0, u_1, \ldots, u_l, w_1, w_2, \ldots, w_m, v_1, v_2, \ldots, v_n).$$

Hence, $\mathbf{x} \in [u]_0$ and $\mathbf{x} \in \sigma^{-n}([v]_0)$, where $n = l + m$. Thus, $\mathbf{x} \in [u]_0 \cap \sigma^{-n}([v]_0)$, which implies that X is transitive. ∎

Corollary. Every SFT and sofic shift is transitive.

Proof. Let X be SFT with forbidden word set F. Hence, F is finite and no word contains a string including words from F. Let (u, v) be an ordered pair of allowed words; hence neither u nor v contain any element of F as a string. Now, X must contain an element $\mathbf{x}$ that includes the word

$$(u_0, u_1, \ldots, u_{l(u)-1}, w_0, w_1, \ldots, w_k, v_0, v_1, \ldots, v_{l(v)-1}).$$

Since $\mathbf{x}$ does not contain any word from F, it follows that $(w_0, w_1, \ldots, w_k) \notin F$. Thus X is transitive. If X is sofic, from the fact that it is a factor of SFT, it follows that it is also transitive. ∎

Recall that (Σ, σ) is not minimal, so are GMS and the even shift (they have periodic points). On the other hand, one can find minimal subshifts. Clearly, if $\mathbf{x}$ is any periodic point (with finite period, of course), then the orbit of is a minimal (cyclic) subshift. These cyclic subshifts are, in some sense, trivial. Naturally, one asks if there are no-cyclic minimal subshifts. Indeed, there are, but they require some elaborate constructions.

Example: Morse Shift. First, we will construct the PTM (Prouhet-Thue-Morse) sequence inductively. Agree on the substitution rule that: $0 \to 01$, $1 \to 10$. Begin with 0, and applying the substitution rule successively, construct a (one-sided) sequence of 0's and 1's:

$$\begin{array}{l} 0 \\ 01 \\ 0110 \\ 01101001 \\ 0110100110010110 \\ \quad\cdots \end{array}$$

Let $\omega = (\omega_1, \omega_2, \omega_3, \ldots)$ be the resulting one-sided sequence, called the PTM sequence. This sequence ω is not periodic (for a combinatorial proof see[3]). ω does not include any block of 000 or 111. Hence, any two 1's or 0's have at most a gap of two, i.e., it has bounded gaps, since ω consists of strings of 01's and 10's. Now, let $X = \overline{O_\sigma(\omega)}$, the system (X, σ) is called as the *Morse* system. Thus, Morse system is generated by a non-cyclic sequence with bounded gaps. By the Theorem of Birkhoff-Gottschalk [5] (which states that a system $(\overline{O_T(x)}, T)$

is minimal if and only if x has bounded gaps), the Morse system must be a minimal, non-cyclic subshift!

Remark. Morse shift is not sofic, since any sofic shift must have a periodic point, which is not the case for the Morse shift.

We know that the full shift (Σ, σ) is not uniquely ergodic since it has (uncountably) many invariant probability measures (Bernoulli measures). Clearly, cyclic systems are uniquely ergodic. However, there are uniquely ergodic non-minimal subshifts.

Example. Let $\mathbf{x} = (\ldots, 0, \ldots, \underline{0}, 1, 0, \ldots, 0, \ldots)$ be a two-sided sequence (in $\{0,1\}^{\mathbb{Z}}$). Then $X = \overline{O_\sigma(\mathbf{x})} = O_\sigma(\mathbf{x}) \cup \{\mathbf{0}\}$. Therefore, if $\mathbf{y} = \mathbf{0}$, then $\mathbf{y}$ fails to have a dense orbit in (X, σ); and hence, (X, σ) is transitive but not minimal. Now, let $\mu = \delta_{\mathbf{y}}$, where

$$\delta_{\mathbf{y}}([a_1, a_2, \ldots, a_n]_0) = \begin{cases} 1 \text{ if } a_i = 0, 1 \le i \le n \\ 0 \text{ if } \ a_i = 1 \text{ for one } 1 \le i \le n. \end{cases}$$

Then μ is σ-invariant and is the only such measure, since the only other candidates are the usual Bernoulli measures which are identically zero on any measurable set $A \subset X$ (such sets are countable unions of cylinder sets at which Bernoulli measures are zero). Thus the system (X, μ, σ) is uniquely ergodic.

Exercises. 1. Prove that every subshift supports at least one σ-invariant measure (not necessarily unique).
2. Show that if $X = \overline{O_T(x)}$, where x is periodic, then (X, σ) is uniquely ergodic.
3. The Morse system is uniquely ergodic. Note: use the fact that ω is of the form $(a_r\ b_r\ b_r\ a_r\ b_r\ a_r\ a_r\ b_r\ \ldots)$, where a_r and b_r are some words in ω.

If a topological dynamical system has more than one (distinct) points with (distinct) finite orbits, then it cannot be uniquely ergodic. This is due to the simple fact that the orbits of each points supports an invariant measure; hence, the measure on the system cannot be unique. Consequently, SFTs (hence, sofic systems) need not be uniquely ergodic. There are also examples of minimal systems (not necessarily subshifts) that are not uniquely ergodic.

There is also a notion of "mixing" in the topological setting: a (topological) dynamical system (X, T) is called *(topologically) mixing* if $\forall$ open sets $U, V \subset X$, $\exists n_0 \in \mathbb{N}$ such that $T^n U \cap V \neq \emptyset$ whenever $n \ge n_0$. It follows that every topological mixing system is transitive. Of course, not every SFT is topologically mixing; however one can characterize mixing SFT's (see Fact.6 below). For instance GMS is topologically mixing; however the SFT with forbidden word set $F = \{01\}$ is not mixing.

Ergodicity, weak-mixing and mixing are important attributions associated to (measure preserving) dynamical systems. In the case of shift spaces the definitions of these can be restated as follows. A subshift (X, σ) of (Σ, σ) is *mixing* if for any allowed words $u,\ v,\ \exists\ n \geq 1$ such that for any $|k| \geq n,\ \exists\ x^k \in X$ with

$$x^k = \dots u x_i x_{i+1} \dots x_{i+k} v \dots .$$

(That is, given enough time, one can get from u to v within X.) The subshift (X, σ) is *weakly mixing* if for any pair of ordered pairs (u_1, v_1) and (u_2, v_2) of allowed words, there exists an allowed word w such that both $u_1 w v_1$ and $u_2 w v_2$ are allowed words. It follows that if a subshift if mixing, it is also weakly mixing, and every weakly mixing subshift is transitive (exercise).

Examples. 1. The GMS is mixing (exercise); hence, it is also weakly-mixing.

2. Recall that the PTM sequence is non-periodic. Therefore, given any block w there exists no n such that $x = \dots\ w\ x_i x_{i+1} \dots x_{i+k}\ w\ \cdots \in X$ for $|k| \geq n$, where X is the Morse dynamical system. Thus, X is not mixing. Next, let $U = \{x\}$ and $V = \{\sigma x\}$ for some $x \in X$. Then, there exists no $n \geq 1$ such that $\sigma^n V \cap U \neq \emptyset$; consequently, there exists no allowed word w such that both $u_1 w v_1$ and $u_2 w v_2$ are allowed words, where $u_1 w v_1$ and $u_2 w v_2$ are allowed words in x. Therefore, the Morse system is not weakly mixing either.

In the rest of this section we will provide some basic properties of SFTs and sofic shifts.

Let X_F be a SFT with the set of forbidden words F. Let N be the length of the longest word in F. If F' is the set of all words of length N containing some word of F, then the SFT $X_{F'}$ with the set of forbidden words F' is the same as X_F (Exercise). That is why, we can assume that all forbidden word of a SFT has the same length. A subshift is called *of order N* (or *an N-step SFT*) if all words in its set F of forbidden words consists of words of length $N+1$. It is easy to see that a SFT of order N is a SFT of order K for all $K \geq N$.

An important feature of SFTs is that any SFT X of order N can be described by an $(N+1) \times (N+1)$ matrix with entries in $\{1, 2, \dots, N\}$. When $n = 1$, the set of all allowed words of length 2 can be defined by a function $A : \{0,1\}^2 \to \{0,1\}$ with

$$A(a,b) = \begin{cases} 1 \text{ if } (a,b) \text{ is an allowed pair} \\ 0 \text{ otherwise.} \end{cases}$$

Hence, A is essentially a 2×2 matrix $\left[a_{ij}\right]$ having $\{0,1\}$ as a set of indices for its entries and with entries in $\{0,1\}$. So,

$$X := X^A = \{x \in \Sigma_2 :\ \forall\ i \in \mathbb{N}\ A(x_i, x_{i+1}) = 1\}.$$

The matrix A is also called as the *transition matrix for X*.

Example. Let X be the GMS. So, $X = X_F$ with $F = \{11\}$. Hence, the set of allowed pairs (or words of length 2) is $W = \{(0,0),(0,1),(1,0)\}$. Consequently,

$$A(x_i, x_{i+1}) = \begin{cases} 1 \text{ if } x_i = 0 = x_{i+1} \text{ or } x_i = 0, x_{i+1} = 1 \text{ or } x_i = 1, x_{i+1} = 0 \\ 0 \text{ otherwise.} \end{cases}$$

Thus, $a_{00} = a_{01} = a_{10} = 1$ and $a_{11} = 0$, implying that the transition matrix of GMS is

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

The converse of the process above can be used to construct SFTs; namely, an $N \times N$ matrix with entries in $\{0,1\}$ can be viewed as the transition matrix for a SFT of order $N-1$. This view is particularly useful in describing SFTs of any order.

Examples. 1. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Then $W = \{(0,0)\}$; hence, $\forall\ x \in X,\ x = \mathbf{0}$, i.e., $X = X^A = \{\mathbf{0}\}$. (Also, $X = X_F$ with $F = \{01, 10, 11\}$.)

2. If $A = \begin{bmatrix}\mathbf{0}\end{bmatrix}_{2\times 2}$, then $X^A = \emptyset$. Similarly, if

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{or} \quad A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

then $X^A = \emptyset$.

3. If $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, then $W = \{(0,0),(1,1)\}$; hence $X^A = \{\mathbf{0}, \mathbf{1}\}$, where $\mathbf{1} = 1^\infty$. By the same token,if $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, then $X^A = \{\mathbf{x}, \mathbf{y}\}$, where $\mathbf{x} = (0,1,0,1,0,1,0,\dots)$ and $\mathbf{y} = \sigma\mathbf{x} = (1,0,1,0,1,0,\dots)$.

Exercise. Describe X^A if A is

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Transition matrix of a SFT is a very convenient tool to study its properties. One prime example of this fact is that the number of periodic points of a SFT is determined by the characteristic polynomial of its transition matrix (See pp: 23-25, [8]).

Typically, on a subshift X (hence on SFTs) the natural metric is the metric on Σ (restricted to X). In regards to measurable structure, SFTs are endowed with Markov measures by a matrix the same size as the matrix defining the shift. Say X_A is a SFT, where A is an $n \times n$ matrix. Let P be an $n \times n$ stochastic matrix, i.e., $P = \begin{bmatrix} p_{ij} \end{bmatrix}$, with

(i) $p_{ij} \geq 0$,
(ii) $\sum_j p_{ij} = 1, \ \forall \ i = 1, 2, \ldots, n$, (all row sums are 1),
(iii) $p_{ij} = 0$ whenever $a_{ij} = 0$.

Notice that (ii) implies that $(1, 1, \ldots, 1)$ is a right eigenvector for P. Let $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ be a left eigenvector with the eigenvalue 1 (i.e., $\mathbf{p}P = \mathbf{p}$). Observing that the cylinder sets in a SFT form a countable basis for the topology, now, define ν on cylinder sets by

$$\nu([a_1, \ldots, a_k]_{n_1}) = p_{a_1} p_{a_1 a_2} p_{a_2 a_3} \ldots p_{a_{k-1} a_k}.$$

Then ν is a finitely additive set function on cylinder sets of X_A; hence, by Kolmogorov extension Theorem, it extends to a measure on the sigma algebra of subsets of X_A. Note that $\mathbf{p}$ is needed for consistency, i.e., that

$$\begin{aligned}\nu([a_2, \ldots, a_k]_{n_2}) &= p_{a_2} p_{a_2 a_3} \ldots p_{a_{k-1} a_k} \\ &= \Big(\sum_{1 \leq a_1 \leq k, A_{a_1 a_2} = 1} \frac{p_{a_1}}{p_{a_2}} p_{a_1 a_2} \Big) p_{a_1} p_{a_1 a_2} p_{a_2 a_3} \ldots p_{a_{k-1} a_k} \\ &= \sum_{1 \leq a_1 \leq k, \ A_{a_1 a_2} = 1} \nu([a_1, \ldots, a_k]_{n_1}).\end{aligned}$$

Remarks. 1. By construction and the consistency (i.e., by $\mathbf{p}P = \mathbf{p}$ or, equivalently, $\sum_{i=1}^{k} p_i p_{ij} = p_j$), the measure ν is shift invariant.

2. As in the case of full shift, (nontrivial) SFTs also have uncountably many σ-invariant measures.

The transition matrix and the associated set W of allowed words of length N for a SFT can also be viewed an an oriented graph describing the allowable (forbidden) words; hence, all possible strings appearing in elements of the SFT. Consequently, any SFT of order $N-1$ can be represented by (doubly) infinite walks on this oriented graph. The SFTs defined by such graphs G are also called *vertex shifts* and denoted by X^G, the graph G is also called as the *transition graph* for the SFT.

Examples. 1. $X_A = \Sigma$ is given by $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ with the associated graph G as

$$\circlearrowright \mathbf{0} \longleftrightarrow \mathbf{1} \circlearrowleft .$$

2. Let X be the GMS. Then

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and the associated graph } G \text{ is:} \quad \circlearrowright \mathbf{0} \longleftrightarrow \mathbf{1}.$$

3. If X_A is given by $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, then the associated graph G is

$$\mathbf{0} \longleftrightarrow \mathbf{1} \circlearrowleft .$$

4. If $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, then the associated graph G is

$$\circlearrowright \mathbf{0} \longleftarrow \mathbf{1} \circlearrowleft,$$

and if $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, then the associated graph G is

$$\circlearrowright \mathbf{0} \longrightarrow \mathbf{1} \circlearrowleft .$$

5. If $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$, then the associated graph G is

$$\begin{array}{ccc} \mathbf{2} & \longleftrightarrow & \mathbf{0} \circlearrowleft \\ & \searrow & \\ & & \mathbf{1} \circlearrowleft . \end{array}$$

Remarks. 1. It can be observed from the graphs that if X_A is given by $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, then $X_A = X^G \approx$ GMS.

2. There are also "edge shifts" induced by a graph.

Another benefit of the transition matrix representation is that it provides another criteria for the corresponding dynamical system. For, recall that a matrix $A = \left[a_{ij}\right]$ of nonnegative integer entries is called *irreducible* if $\forall\ i, j,\ \exists\ n \geq 0$ such that $(A^n)_{ij} > 0$; and it is called *aperiodic* (or *primitive*) if $\exists\ n \geq 0$ such that $(A^n)_{ij} > 0\ \forall\ i, j$. Notice that if A is aperiodic, then it is irreducible.

Fact.6 a) A is irreducible if and only if (X_A, σ) is transitive.

b) A is aperiodic if and only if (X_A, σ) is mixing.

Proof. See [3].

Recall that every SFT is sofic; however, it is not easy to show that a subshift is sofic, for one needs some "sophisticated" tools of characterization. One such characterization is via "follower sets."

Definition. Let X be a subshift and u be an allowable word in X. The *follower set* $W_X(u)$ of u in X is the set of all (allowable) words v that can follow u in X; i.e., $W_X(u) = \{v : uv \text{ is allowed in } X\}$. The collection of all follower sets in X is denoted by

$$\mathcal{C}_X = \{W_X(u) : u \text{ is in } X\}.$$

Theorem.1 A shift $X \subset \Sigma_A$ is sofic if and only if $|\mathcal{C}_X| < \infty$.

Proof. See [3].

Example. The even shift is sofic. For, let

(i) u_1 be a word that contains no 1's,
(ii) u_2 be a word that ends in 10^{2k}, for some $k \geq 0$,
(iii) u_3 be a word that ends in 10^{2k+1}, for some $k \geq 0$.

Then

(i) $W_X(u_1) = \{0, 1, 00, 01, 10, 11, 000, 001, 010, 100, \dots\}$,
(ii) $W_X(u_2) = \{0, 1, 00, 10, 11, 000, 001, 100, \dots\}$, and
(iii) $W_X(u_3) = \{0, 00, 01, 000, 010, 011, \dots\}$.

These are the only follower sets; hence, $|\mathcal{C}_X| = 3$. Thus, the even shift is sofic.

The even shift is not a SFT; a sofic shift which is not a SFT is called *strictly sofic*. Hence, the even shift is strictly sofic.

Exercises. 1. A factor of a sofic shift is sofic.
2. A shift conjugate to a sofic shift is sofic.
3. If X is sofic, then it has periodic points.

Remark. It follows from Exercise.3 above that the Morse system is not sofic.

Let Y be a set of words on an alphabet $\mathcal{A}$, such a set is also called a *language* on $\mathcal{A}$. Define

$$Y^\infty = \{(a_n) \in \mathcal{A}^{\mathbb{N}} : (a_n) = w_1 w_2 \dots w_k \dots \ , w_i \in Y\}.$$

A subshift X for which there exists a (not necessarily unique) language Y such that $X = \overline{Y^\infty}$ is called a *coded system.* It follows from Fact.4 we obtain that

Corollary. All coded systems are transitive.

Furthermore, we have the following feature of the sofic systems.

Fact.7 All transitive sofic shifts are coded systems.

Proof. Let X be a sofic system X with language $Y(X)$. Assume that the words of X are obtained from the finite oriented graph G. All allowable words of X are those that can be spelled by paths in G; i.e., the language $Y(X)$ is recognizable in this fashion. Since X is transitive, all elements $\mathbf{a}$ of X are obtained as $\mathbf{a} = w_1 w_2 \dots w_k \dots \ , w_i \in Y(X)\}$. ∎

The converse of Fact.7 is not valid, however. For instance, let $X = \overline{Y}$, where $Y = \{0^n 1^n : n \in \mathbb{N}\}$. By definition, X is coded. On the other hand, the follower set of all words in X are infinite; hence, X is not sofic.

4. β-transformations T_β and β-shifts D_β.

a) β-transformations.

Let $\beta \in \mathbb{R}^+,\ \beta > 1$, be non-integer and $X = [0,1]$. Let $T_\beta : X \to X$ be a map defined by $T_\beta x = \beta x (\mod 1)$. Then, the map T_β on X is called the *β-transformation.* the β-transformation T_β has $[\beta] + 1$ branches; $[\beta]$ full branches and one "non-full" branch. (In the case that $\beta = n \in \mathbb{Z}^+,\ T_n$ has n full branches only.) Due to this nature, β-transformations have significantly distinct features in comparison to n-fold maps. To begin with, let $(a,b) \subset [0,1]$ be such that $\beta - [\beta] < a < b < 1$, then

$$m(T_\beta^{-1}(a,b) = [\beta](\frac{b-a}{\beta}) < b - a = m((a,b))!$$

Hence, T_β does not preserve the Lebesgue measure. On the other hand, since, by Krylov-Bogulibov Theorem, any compact dynamical system carries at least one invariant Borel probability measure, (X, T_β) has an invariant Borel probability measure. Hence, the question is the existence of T_β-invariant Borel probability measure absolutely continuous w.r.t. (or equivalent to) Lebesgue measure. It turns out that the answer to this question is affirmative: the existence of such a measure ν_β for each $\beta > 1$ was shown by A. Renyi in 1957, and the explicit form of such measures was obtained by W. Parry in 1960.

Example. Let $\beta = \frac{\sqrt{5}+1}{2}$, the golden mean ($\approx 1.618\ldots$). Recall that β is the (larger) root of $x^2 - x - 1 = 0$. So, $\frac{1}{\beta} = \beta - 1 = \frac{\sqrt{5}-1}{2}$ ($\approx 0.61\ldots$). In this case, if

$$f(x) = \begin{cases} \dfrac{5+3\sqrt{5}}{10} & \text{if } 0 < x < \dfrac{1}{\beta} \\ \dfrac{5+\sqrt{5}}{10} & \text{if } \dfrac{1}{\beta} < x < 1, \end{cases}$$

then the measure ν_β given by $\nu_\beta(E) = \int_E f(x)dm$ for Borel subsets E of $[0,1)$, is well-defined. Furthermore, it is straightforward to show that ν_β is also a T_β-invariant Borel probability measure which is equivalent to the Lebesgue measure m.

It also turns out that β-transformations are ergodic.

Fact.8 For any $\beta > 1,\ T_\beta$ is ergodic (w.r.t. Lebesgue measure).

Proof. We will call any interval $I \subset [0,1]$ a *full interval of rank n* if $m(I) = \frac{1}{\beta^n}$. (For example, $T_\beta^{-n}[0,1]$ consists of 2^n distinct subintervals, each of which is a full interval of rank n.) Let $B \in \mathcal{F}$ such that $T_\beta^{-1}B \subset B$ and $m(B) > 0$. We will show that $m(B) = 1$. Let E be a full interval of rank n. Then, $\forall\ C \in \mathcal{F},\ m(T_\beta^{-n}C \cap E) = \frac{m(C)}{\beta^n}$. Hence,

$$\frac{m(B \cap E)}{m(E)} = \frac{m(T_\beta^{-n}B \cap E)}{m(E)} = \frac{m(B)}{\beta^n m(E)} = m(B).$$

So, $M(B \cap E) = m(B)m(E)$. Letting $\gamma = m(B)$, we have $m(B) = 1$ by Knopp's Lemma. ∎

As in the case of the n-fold map and the n-shift, which are conjugates of each other, the conjugate of the β-transformation is called as the *β-shift*. The β-shift can be defined in two ways. For, first we need to introduce a partial order, called as the *lexicographic order,* on $\mathcal{A}^{\mathbb{N}}$, where $\mathcal{A} = \{0, 1, \dots, [\beta]\}$. Given $\mathbf{a} = (a_1, a_2, \dots, a_k, \dots)$, $\mathbf{b} = (b_1, b_2, \dots, b_k, \dots) \in \mathcal{A}^{\mathbb{N}}$, we define

$$\mathbf{a} <_L \mathbf{b} \text{ if and only if } a_n < b_n,\ n = \min\{k : a_k \neq b_k\}.$$

Now, for a positive real number $\beta > 1$, let's define the β-expansion of $x \in [0, 1]$ by

$$x = \sum_{k=1}^{\infty} \frac{a_k}{\beta^k}, \quad \text{where, } a_k \in \mathcal{A}.$$

This expansion is unique if $\beta \in \mathbb{Z}$; however, if β is non-integer x may have more than one (indeed infinitely many) β-expansions. For example, let $\beta = \frac{1+\sqrt{5}}{2}$ (the golden mean), then

$$\begin{aligned} 1 &= \frac{1}{\beta} + \frac{1}{\beta^2} \quad (\text{since } \beta^2 - \beta - 1 = 0 \ \Rightarrow\ 1 - \frac{1}{\beta} - \frac{1}{\beta^2} = 0), \text{ or} \\ 1 &= \frac{1}{\beta} + \frac{1}{\beta^3} + \frac{1}{\beta^4}, \text{ or} \\ 1 &= \frac{1}{\beta} + \frac{1}{\beta^3} + \frac{1}{\beta^5} + \frac{1}{\beta^6}, \text{ or} \\ &\dots \\ 1 &= \frac{1}{\beta} + \frac{1}{\beta^3} + \frac{1}{\beta^5} + \dots = \sum_{k=0}^{\infty} \frac{1}{\beta^{2k+1}}. \end{aligned}$$

Hence, as element of $\mathcal{A}^{\mathbb{N}}$, 1 has β-expansions

$$\begin{gathered} 110^{\infty}, \text{ or} \\ 10110^{\infty}, \text{ or} \\ 1010110^{\infty}, \text{ or} \\ \dots \\ (10)^{\infty}. \end{gathered}$$

Which one will we accept? Given all possible β-expansions of a real number $x \in [0, 1]$, we will call the maximal one (w.r.t. lexicographic order) as the *greedy β-expansion* (or simply *greedy expansion*) and the minimal one as the *lazy β-expansion* (or simply *lazy expansion*) of x. We will choose the greedy expansion of x as its β-expansion.

In the lexicographical order,

$$(10)^\infty <_L \cdots <_L 1010110^\infty <_L 10110^\infty <_L 110^\infty.$$

Hence, $(10)^\infty$ is the lazy expansion and 110^∞ is the greedy expansion of 1; hence we will accept 110^∞ as the β-expansion of 1 for $\beta = \frac{1+\sqrt{5}}{2}$. In the same way, for a real number $x \in [0,1]$, one can obtain various expansions w.r.t. β, from which the greedy one is defined as the β-expansion of x.

Example. Let $\beta = \frac{3}{2}$. Then all β-expansions of $x = \frac{1}{3}$ are the same: $(000100010001\ldots)$. However, for $x = \frac{1}{\sqrt{2}}$ the greedy β-expansion is $(1010\ldots)$, whereas lazy β-expansion is $(1000000100100\ldots)$.

In particular, the algorithm for the greedy expansion is as follows: given $x \in [0,1]$, let

$$\begin{aligned}
&a_1 = [\beta x],\ r_1 = \beta x - a_1,\\
&a_2 = [\beta r_1],\ r_2 = \beta r_1 - a_2,\\
&a_3 = [\beta r_2],\ r_3 = \beta r_2 - a_3,\\
&\quad\ldots\\
&a_k = [\beta r_{k-1}],\ r_k = \beta r_{k-1} - a_k, \quad k \geq 2.
\end{aligned}$$

Then $x = \sum_{k=1}^\infty \frac{a_k}{\beta^k}$ is the greedy expansion of x. For example, for $\beta = \frac{3}{2}$, the greedy expansion of $x = \frac{1}{2}$ is $(01000001001\ldots)$.

The other way of defining β-expansion is via β-transformation, which is also known as the *Renyi β-expansion* (or simply *Renyi expansion*). For $x \in [0,1]$ and $\beta > 1$ a real number, let

$$\begin{aligned}
&a_1 = [\beta x],\\
&a_2 = [\beta T_\beta x],\\
&a_3 = [\beta T_\beta^2 x],\\
&\quad\ldots\\
&a_k = [\beta T_\beta^{k-1} x] \quad k \geq 2,
\end{aligned}$$

where T_β is the β-transformation. Then $x = \sum_{k=1}^\infty \frac{a_k}{\beta^k}$ is the Renyi expansion of x. Clearly, for any $\beta > 1$, the greedy expansion and the Renyi expansion of a real number $x \in [0,1]$ are the same.

A. Renyi proved in 1957 that for any $0 \leq x < 1$ the greedy and Renyi expansions coincide; however, if $x = 1$, then they differ. Namely, while the greedy expansion of 1 is 10^∞, the Renyi expansion of 1 is a sequence $(a_1 a_2 \ldots a_k \ldots)$ where $a_k = [\beta T_\beta^{k-1} x]$ $k \geq 1$. For example, if $\beta = \sqrt{2}$, then 1 has the greedy expansion 10^∞ and the Renyi expansion $(10010\ldots)$, which is a non repeating sequence. The Renyi β-expansion of 1is denoted by 1_β. It is interesting to note that in 1994 Erdös, Joo and Komanek proved that $\forall\ N \geq 1,\ \exists$ uncountable number of $\beta \in [1,2]$ such that 1 has exactly N different β-expansions.

Fact.9 Let $x, y \in [0,1]$ and let $\hat{x},\ \hat{y}$ be their Renyi β-expansions (in $\mathcal{A}^{\mathbb{N}}$), respectively. Then

$$x < y \iff \hat{x} <_L \hat{y}.$$

Proof. Exercise.

It follows that Renyi β-extension of 1 determines a type of subshifts of $\{0, 1, \ldots, [\beta]\}^{\mathbb{N}}$ in a unique manner. Now, we turn to study some basic features of such subshifts.

b) Basic Properties of β-shifts.

Observe that when $n \in \mathbb{N}$, the shift space $\Sigma_n = \{0, 1, \ldots, n-1\}^{\mathbb{N}}$ does not contain $\mathbf{1} = (1, 1, \ldots 1 \ldots) = 1^\infty$. On the other hand, for any $\mathbf{a} = (a_1 a_2 \ldots a_k \ldots) \in \Sigma_n$, we have $\mathbf{a} <_L \mathbf{1}$. Indeed, we can re-characterize the shift space Σ_n as

$$\Sigma_n = \{(a_k) \in \mathcal{A}^{\mathbb{N}} : (a_k) <_L \mathbf{1}\},$$

where $\mathcal{A} = \{0, 1, \ldots, [\beta]\}$, which is known to be a compact metric space closed under the shift map.

From the discussion above, we can deduce that not every sequence of the form $(a_1 a_2 \ldots a_k \ldots) \in \mathcal{A}^{\mathbb{N}}$ is the β-expansion of a real number $x \in [0,1]$. Given $\beta > 1$ non integer, let D_β be the collection of all sequences of the form $(a_1 a_2 \ldots a_k \ldots)$ such that $\sum_{k=1}^{\infty} \frac{a_k}{\beta^k}$ is the β expansion of a real number $x \in [0,1]$. Clearly, sequences $(a_k)_k$ obtained via Renyi expansion algorithm belong D_β. It turns out that D_β is a closed and shift invariant subspace of $\mathcal{A}^{\mathbb{N}}$. Furthermore,

$$D_\beta = \{\mathbf{a} = (a_1 a_2 \ldots a_k \ldots) \in \mathcal{A}^{\mathbb{N}} :\ \mathbf{a} <_L 1_\beta\}.$$

The pair (D_β, σ) is called the *β-shift space*, or simply the *β-shift.* W. Parry proved in 1960 that the sequence 1_β uniquely determines the shift space D_β, and it is associated to theβ-transformation T_β on $[0,1]$ via the association $x \leftrightarrow (a_k)_{\text{Renyi}}$.

Examples. 1. If $\beta = \frac{1+\sqrt{5}}{2}$, then $1_\beta = 110^\infty$, a repeating sequence; so

$$D_\beta = \{(a_k) : (a_k) <_L 110^\infty\}.$$

2. For $\beta = \frac{1}{2},\ 1_\beta = (101110000110\ldots)$, a non repeating sequence.
3. For $\beta = \frac{5}{2},\ 1_\beta = (2101\ldots)$, a non repeating sequence.
4. For $\beta = \frac{3}{2},\ 1_\beta = (1010000010\ldots)$, a non repeating sequence.

5. For β largest root of $x^3 - x^2 - 2x + 1 = 0$ in $[1, 2]$, then

$$1_\beta = 1(10)^\infty, \text{ a repeating sequence.}$$

As seen in the examples above, the sequence $1_\beta \in \mathcal{A}^{\mathbb{N}}$ may be periodic, eventually periodic, or non periodic. This feature of 1_β characterizes the nature of space D_β. Call a sequence (a_k) finite if $a_n = 0\, \forall n \geq n_0$ for some $n_0 \geq 1$.

Theorem.2 (W. Parry, 1960) Let $\beta > 1$ be a non-integer.

a) If 1_β is not finite, then

$$(a_k) \in D_\beta \iff \sigma^n(a_k) <_L 1_\beta,\ \forall n \geq 0.$$

b) If 1_β is finite, say $1_\beta = b_1 \dots b_k 0^\infty$, then

$$(a_k) \in D_\beta \iff \sigma^n(a_k) <_L b_1 \dots b_{k-1}(b_k - 1)^\infty,\ \forall n \geq 0.$$

Proof. See [4].

Thus, when β is the golden mean,

$$D_\beta = \{(a_k) : \sigma^n(a_k) <_L (10)^\infty\},\ n \geq 0\}.$$

It follows that no sequence containing the string $\dots 11 \dots$ is in D_β. Hence, D_β is the GMS, which is a SFT! Notice also that in each of the examples above $D_\beta \neq \mathcal{A}^{\mathbb{N}}$. Hence, (D_β, σ) is a proper subshift in each of these cases; indeed, this is the case for every non-integer $\beta > 1$.

Observe that, by Theorem.2, D_β is the set of all one-sided sequences that are lexicographically less that 1_β; hence, it is the set of all sequences that are concatenations of words of the form $b_1 \dots b_{k-1} d$, where $d \in \mathcal{A}$ is such that $d < b_k,\ k \in \mathbb{N}$. Let Y_β be the collection of all such words; hence, $D_\beta = \overline{Y_\beta^\infty}$. This proves that

Fact.10 All β-shifts are coded systems.

Since T_β is ergodic for each $\beta > 1$, it follows that β-shifts are transitive (Exercise: Prove this fact). On the other hand, they need not be uniquely ergodic. For example, let β ($\approx 2.65897\dots$) be the largest root of $x^3 - 2x^2 - x - 2 = 0$. Then $1_\beta = \overline{211}$; hence, $x = (02)^\infty <_L 1_\beta$. Since x is periodic, it supports an invariant measure, 1_β also supports an invariant measure; hence, D_β cannot be uniquely ergodic.

Let $Per(\beta)$ be the set of all $x \in [0, 1]$ such that $O_{T^\beta}(x)$ is finite (i.e., the set of periodic points). (These are also called β *numbers*.) Let $\mathbb{Q}(\beta)$ be the smallest field containing $\mathbb{Q}$ and β. (When $\beta \in \mathbb{N},\ \mathbb{Q}(\beta) = \mathbb{Q}$.)

Fact. 11 For $\beta > 1,\ Per(\beta) \subset \mathbb{Q}(\beta) \cap [0, 1)$.

Proof. When $\beta \in \mathbb{N}$, $Per(\beta) = \mathbb{Q}$, proving the case for β is integer. So, assume that β is non-integer and $x \in Per(\beta)$. Then

$$x = a_1 a_2 \dots a_{m-1} \overline{a_m a_{m+1} \dots a_{m+l-1}}.$$

Letting $y = \overline{a_m a_{m+1} \dots a_{m+l-1}}$, it follows that

$$y = \frac{a_m}{\beta} + \frac{a_{m+1}}{\beta^2} + \dots + \frac{a_{m+l-1}}{\beta^l} + \frac{y}{\beta^l}.$$

Hence,

$$y = \frac{\beta^{l-1}}{\beta^l - 1} a_m + \frac{\beta^{l-2}}{\beta^l - 1} a_{m+1} + \dots + \frac{1}{\beta^l - 1} a_{m+l-1} \in \mathbb{Q}(\beta) \cap [0, 1).$$

This implies the assertion. ∎

The converse of this statement is highly nontrivial. It is proved by A. Bertrand (1977) and K. Schmidt (1980) that if $\beta > 1$ is a Pisot number (see definition below), then $Per(\beta) = \mathbb{Q}(\beta) \cap [0, 1)$. Hence, if $\beta > 1$ is Pisot, then every number of $\mathbb{Q}(\beta)$ must have an eventually periodic β-expansion.

The golden mean example above suggest that there is a close connection between the number theoretical properties of β and the nature of the β-shift D_β. This is best seen by the interplay between the properties of 1_β (i.e., the orbit of 1_β under T_β) and the ergodic properties of D_β. First, we note the following

Fact.12 If (a_n) is an expansion of 1 for some β such that $\sigma^k(a_n) <_L (a_n)$, $k > 0$, then β is unique.

Proof. By hypothesis and Theorem.2, β must be the unique solution of the equation $1 = \sum_{k=0}^{\infty} a_k \beta^{-k}$. It is easy to check that β-expansion of 1 is equal to (a_n). ∎

For technical purposes (needed for the following theorem and after), we extend the order $<_L$ to finite blocks:

$$a_1 \dots a_l <_L b_1 \dots b_m \iff a_1 \dots a_l 0^\infty <_L b_1 \dots b_m [\beta][\beta] \dots.$$

Theorem.3 (Parry, 1960) 1_β is finite if and only if (D_β, σ) is a SFT.

Proof. 1_β is finite, say $1_\beta = b_1 b_2 \dots b_n 0^\infty$, is equivalent to saying that, by Theorem.1, $(a_n) \in D_\beta$ iff $(a_n) <_L \overline{b_1 b_2 \dots b_{n-1}(b_n - 1)}$. Hence, the set of forbidden words of D_β are those $d_1 \dots d_j$, $1 \leq j \leq n$, with $b_1 b_2 \dots b_j <_L d_1 \dots d_j$, which is a finite set of words. ∎

Theorem.4 (Bertrand-Mathis, 1988) 1_β is eventually periodic if and only if (D_β, σ) is a sofic.

Proof. (Sketch) Let $1_\beta = a_1 a_2 \dots a_l \overline{b_{l+1} b_{l+2} \dots b_{l+n}}$. The associated shift D_β is a factor of the shift generated by $1_\gamma = b_{l+1} b_{l+2} \dots b_{l+n-1} d 0^\infty$, where $d = (b_{l+n} + 1)(\mod [\beta])$. ∎

In the case that 1_β not eventually periodic, the situation is more complicated. First, consider the case that 1_β contains bounded strings of 0s but not eventually periodic.

Definition. A subshift $X \subset \{0, 1, \ldots, [\beta]\}^{\mathbb{N}}$ is called *specified* if $\exists\ k$ such that $\forall\ u, v$ allowed words, $\exists$ an allowed word w with length k such that uwv is allowed.

Definition. A subshift $X \subset \{0, 1, \ldots, [\beta]\}^{\mathbb{N}}$ is called *synchronizing* if $\exists$ a word $u \in Y(X)$ such that if $vu \in Y(X)$, then $vwu \in Y(X) \iff uw \in Y(X)$.

Theorem.5 (Bertrand-Mathis, 1988) a) D_β is specified if and only if $\exists n \in \mathbb{N}$ such that all strings of 0s in 1_β have length less than n (equivalently, **0** is not an accumulation point of $O_{T_\beta}(1)$.

b) D_β is synchronizing if and only if 1_β does not contain some allowed word (i.e., $O_{T_\beta}(1)$ is not dense in [0,1]).

Proof. See [1].

As it can be surmised by the properties seen above, β-shifts have some interesting number theoretical properties. Below, we will see some of the interesting interplay between types of numbers $\beta > 1$ and associated subshifts. First, we will give some definitions. Recall that any number α which is the root of a polynomial with integer coefficients is called an algebraic number. (This polynomial is called as the *characteristic polynomial of* α.)

Definition. An algebraic number whose characteristic polynomial has a term of highest degree with coefficient 1 is called an *algebraic integer.*

Examples. $\sqrt{2}$ is an algebraic integer (root of $x^2 - 2 = 0$); however any rational number $\frac{p}{q}$, $q \neq 1$, is not an algebraic integer (since it is the root of $qx - p = 0$).

Definition. An algebraic integer $\alpha > 1$ all conjugates of which have absolute value less than α is called a *Perron number.*

Examples. $\sqrt{2}$ is not a Perron number (its conjugate is $-\sqrt{2}$). $\alpha = \frac{1}{2}(5 + \sqrt{5})$ is a Perron number (since its characteristic polynomial is $x^2 - 5x + 5$ with the conjugate root $\frac{1}{2}(5 - \sqrt{5})$.

Definition. A Perron number α all conjugates of which have absolute value less than 1 is called a *Pisot number.*

Examples. All integers are Pisot numbers. $\frac{1}{2}(1 + \sqrt{5})$ is a Pisot number (since its characteristic polynomial is $x^2 - x - 1$ with conjugate $|\frac{1}{2}(1 - \sqrt{5})| < 1$). On the other hand, $\frac{1}{2}(5 + \sqrt{5})$ is not a Pisot number.

Definition. A Perron number α whose all conjugates have absolute value less than or equal to 1 with at least one conjugate $|b| = 1$ is called a *Salem number.*

Examples. Salem numbers are not easy to find. Smallest known Salem number is the largest root of (Lehmer's) polynomial

$$x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1,$$

which is approximately 1.17628... The number $\frac{1}{2}(5 + \sqrt{5})$ is not a Salem number.

The following statements exhibit two of the interesting interplay between type of numbers and the type of subshifts, we will state them without proof. (For proofs we refer the reader to [1], [2] or [7].)

Proposition.1 (W. Parry, 1960) If β is a Pisot number, then D_β is sofic.

The following partial converse statement is known:

Proposition.2 (Lind, 1984) (Denker, Grillenberger and Sigmund, 1976) If D_β is sofic, then β is a Perron number.

Note: If $\beta = \frac{p}{q}$ (non-integer), then β is not an algebraic integer; hence, D_β is not sofic!

The following problems are still open.

Problem.1 Characterize the set of all real numbers β for which D_β is sofic.

Problem.2 Does the set of all β such that D_β is sofic contain all Salem numbers?

K. Schmidt proved in 1980 that if $\beta > 1$ is such that $\mathbb{Q} \cap [0, 1) \subset Per(\beta)$, then β is either Pisot number or a Salem number.

Conjecture. (K. Schmidt) Let $\beta > 1$ be a Salem number. Then $Per(\beta) = \mathbb{Q}(\beta) \cap [0, 1)$.

This conjecture has been shown to be true for all Salem numbers of degree 4 by D. Boyd. The general case is still open.

Since we deal with $\beta \in (1, \infty)$, naturally one inquires existence of interplay between the measure/geometric properties of sets of β and classes of subshifts. For, let's define the following subsets of $(1, \infty)$:

$$\begin{aligned}
C_1 &= \{\beta \in (1, \infty) : \ D_\beta \text{ is SFT}\} \\
C_2 &= \{\beta \in (1, \infty) : \ D_\beta \text{ is sofic}\} \\
C_3 &= \{\beta \in (1, \infty) : \ D_\beta \text{ is specified}\} \\
C_4 &= \{\beta \in (1, \infty) : \ D_\beta \text{ is synchronizing}\} \\
C_5 &= \{\beta \in (1, \infty) : \ D_\beta \text{ is none of the above}\}.
\end{aligned}$$

Obviously,

$$C_1 \subset C_2 \subset C_3 \subset C_4 \subset (0, \infty) \ \text{ and } C_5 = (0, \infty) \setminus C_4.$$

From the definition and the properties of algebraic numbers, and using Propositions 1 and 2, we immediately have

Proposition.3 C_2 is at most countable.

The following is a list of some facts on the size of these sets; we will state them without proof. (For proofs we refer the reader to [2], [4] and [7].)

Proposition.4 (W. Parry, 1960) C_1 is dense in $(0, \infty)$.

Proposition.5 (J. Schmeling, 1997) C_3 has Hausdorff dimension 1.

Definition. A subset $A \subset (0, \infty)$ is called *residual* iff it contains a countable intersections of open and dense sets. A is called *meager* if A^c is residual.

Proposition. 6 (J. Schmeling, 1997) The class C_5 is residual in $(0, \infty)$ and has full Lebesgue measure.

Furthermore, J. Schmeling (1997) provided a complete list for the sizes of these sets as:

a) C_1 and C_2 have Hausdorff dimension 0, whereas C_3, C_4 and C_5 have Hausdorff dimension 1.
b) C_1, C_2, C_3 and C_4 have Lebesgue measure 0, and C_5 has full Lebesgue measure.
c) C_1 and C_2 are countable and dense in $(0, \infty)$; C_3 and C_4 are meager and dense in $(0, \infty)$; and C_5 is residual in $(0, \infty)$.

The following is a list of some of the open questions.

Q.1 Characterize numbers belonging to classes C_3, C_4 and C_5.
Q.2 Are there rational numbers in C_3, C_4 and C_5, or only in some of them?
Q.3 Which classes do π, e, $\frac{3}{2}$ belong?
Q.4 Are there transcendental numbers in C_4 and C_5?

It should be remarked here that there are transcendental numbers β for which D_β is specified (J.-P. Allouche).

References

[1] A. Bertrand-Mathis, *Développements en base θ et répartition modulo 1 de la suite $(x\theta^n)$,* Bull. Soc. Math. Fr., **114** (1986), 271-324.
[2] F. Blanchard, *β-expansions and symbolic dynamics,* Theor. Comp. Sci., **65** (1989), 131-141.
[3] B. Marcus and D.A. Lind, **An Introduction to Symbolic Dynamics and Coding**, Cambridge Univ. Press, 1995.
[4] W. Parry, *On the β-expansion of real numbers,* Acta Math. Acad. Sci. Hung., **11** (1960), 401-416.

[5] K. Petersen, **Ergodic Theory**, Cambridge Univ. Press, 1983.

[6] A. Rényi, *Representations for real numbers,* Acta Math. Acad. Sci. Hung., **8** (1957), 477-493.

[7] J. Schmeling, *Symbolic dynamics for β-shifts and self-normal humbers*, Erg. Th. & Dynam. Syst., **17** (1997), 675-694.

[8] B.P. Kitchens, **Symbolic Dynamics**, Springer, 1998.

Notes. [5] is one of the best sources for general information on ergodic theory and [3] is the best and most comprehensive source on symbolic dynamics.

[7] is the article that contains proofs of some major results on β-shifts while providing many relatively recent results on number theoretical, measure theoretical and topological properties of β-shifts.