CHAPTER 1

SHIFT SPACES

Shift spaces are to symbolic dynamics what shapes like polygons and curves are to geometry. We begin by introducing these spaces, and describing a variety of examples to guide the reader's intuition. Later chapters will concentrate on special classes of shift spaces, much as geometry concentrates on triangles and circles. As the name might suggest, on each shift space there is a shift map from the space to itself. Together these form a "shift dynamical system." Our main focus will be on such dynamical systems, their interactions, and their applications.

In addition to discussing shift spaces, this chapter also connects them with formal languages, gives several methods to construct new shift spaces from old, and introduces a type of mapping from one shift space to another called a sliding block code. In the last section, we introduce a special class of shift spaces and sliding block codes which are of interest in coding theory.

§1.1. Full Shifts

Information is often represented as a sequence of discrete symbols drawn from a fixed finite set. This book, for example, is really a very long sequence of letters, punctuation, and other symbols from the typographer's usual stock. A real number is described by the infinite sequence of symbols in its decimal expansion. Computers store data as sequences of 0's and 1's. Compact audio disks use blocks of 0's and 1's, representing signal samples, to digitally record Beethoven symphonies.

In each of these examples, there is a finite set $\mathcal{A}$ of *symbols* which we will call the *alphabet*. Elements of $\mathcal{A}$ are also called *letters*, and they will typically be denoted by a, b, c, ..., or sometimes by digits like 0, 1, 2, ..., when this is more meaningful. Decimal expansions, for example, use the alphabet $\mathcal{A} = \{0, 1, \dots, 9\}$.

Although in real life sequences of symbols are finite, it is often extremely useful to treat long sequences as infinite in both directions (or *bi-infinite*).

This is analogous to using real numbers, continuity, and other ideas from analysis to describe physical quantities which, in reality, can be measured only with finite accuracy.

Our principal objects of study will therefore be collections of bi-infinite sequences of symbols from a finite alphabet $\mathcal{A}$. Such a sequence is denoted by $x = (x_i)_{i\in\mathbb{Z}}$, or by

$$x = \dots x_{-2}x_{-1}x_0x_1x_2\dots,$$

where each $x_i \in \mathcal{A}$. The symbol x_i is the ith *coordinate* of x, and x can be thought of as being given by its coordinates, or as a sort of infinite "vector." When writing a specific sequence, you need to specify which is the 0th coordinate. This is conveniently done with a "decimal point" to separate the x_i with $i \geqslant 0$ from those with $i < 0$. For example,

$$x = \dots 010.1101\dots$$

means that $x_{-3} = 0$, $x_{-2} = 1$, $x_{-1} = 0$, $x_0 = 1$, $x_1 = 1$, $x_2 = 0$, $x_3 = 1$, and so on.

Definition 1.1.1. If $\mathcal{A}$ is a finite alphabet, then the *full $\mathcal{A}$-shift* is the collection of all bi-infinite sequences of symbols from $\mathcal{A}$. The *full r-shift* (or simply *r-shift*) is the full shift over the alphabet $\{0, 1, \dots, r-1\}$.

The full $\mathcal{A}$-shift is denoted by

$$\mathcal{A}^{\mathbb{Z}} = \{x = (x_i)_{i\in\mathbb{Z}} : x_i \in \mathcal{A} \text{ for all } i \in \mathbb{Z}\}.$$

Here $\mathcal{A}^{\mathbb{Z}}$ is the standard mathematical notation for the set of all functions from $\mathbb{Z}$ to $\mathcal{A}$, and such functions are just the bi-infinite sequences of elements from $\mathcal{A}$. Each sequence $x \in \mathcal{A}^{\mathbb{Z}}$ is called a *point* of the full shift. Points from the full 2-shift are also called *binary sequences*. If $\mathcal{A}$ has size $|\mathcal{A}| = r$, then there is a natural correspondence between the full $\mathcal{A}$-shift and the full r-shift, and sometimes the distinction between them is blurred. For example, it can be convenient to refer to the full shift on $\{+1, -1\}$ as the full 2-shift.

Blocks of consecutive symbols will play a central role. A *block* (or *word*) over $\mathcal{A}$ is a finite sequence of symbols from $\mathcal{A}$. We will write blocks without separating their symbols by commas or other punctuation, so that a typical block over $\mathcal{A} = \{a, b\}$ looks like $aababbabbb$. It is convenient to include the sequence of *no* symbols, called the *empty block* (or *empty word*) and denoted by ε. The *length* of a block u is the number of symbols it contains, and is denoted by $|u|$. Thus if $u = a_1a_2\dots a_k$ is a nonempty block, then $|u| = k$, while $|\varepsilon| = 0$. A *k-block* is simply a block of length k. The set of all k-blocks over $\mathcal{A}$ is denoted $\mathcal{A}^k$. A *subblock* or *subword* of $u = a_1a_2\dots a_k$ is a block of the form $a_i a_{i+1} \dots a_j$, where $1 \leqslant i \leqslant j \leqslant k$. By convention, the empty block ε is a subblock of every block.

If x is a point in $\mathcal{A}^{\mathbb{Z}}$ and $i \leqslant j$, then we will denote the block of coordinates in x from position i to position j by

$$x_{[i,j]} = x_i x_{i+1} \dots x_j \, .$$

If $i > j$, define $x_{[i,j]}$ to be ε. It is also convenient to define

$$x_{[i,j)} = x_i x_{i+1} \dots x_{j-1} \, .$$

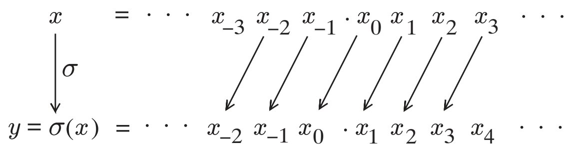
By extension, we will use the notation $x_{[i,\infty)}$ for the *right-infinite sequence* $x_i x_{i+1} x_{i+2} \dots$, although this is not really a block since it has infinite length. Similarly, $x_{(-\infty,i]} = \dots x_{i-2} x_{i-1} x_i$. The *central* $(2k+1)$*-block of* x is $x_{[-k,k]} = x_{-k} x_{-k+1} \dots x_k$. We sometimes will write $x_{[i]}$ for x_i, especially when we want to emphasize the index i.

Two blocks u and v can be put together, or *concatenated*, by writing u first and then v, forming a new block uv having length $|uv| = |u| + |v|$. Note that uv is in general not the same as vu, although they have the same length. By convention, $\varepsilon u = u \varepsilon = u$ for all blocks u. If $n \geqslant 1$, then u^n denotes the concatenation of n copies of u, and we put $u^0 = \varepsilon$. The law of exponents $u^m u^n = u^{m+n}$ then holds for all integers $m, n \geqslant 0$. The point $\dots uuu.uuu \dots$ is denoted by u^∞.

The index i in a point $x = (x_i)_{i \in \mathbb{Z}}$ can be thought of as indicating time, so that, for example, the time-0 coordinate of x is x_0. The passage of time corresponds to shifting the sequence one place to the left, and this gives a map or transformation from a full shift to itself.

Definition 1.1.2. The *shift map* σ on the full shift $\mathcal{A}^{\mathbb{Z}}$ maps a point x to the point $y = \sigma(x)$ whose ith coordinate is $y_i = x_{i+1}$.

The operation σ, pictured below, maps the full shift $\mathcal{A}^{\mathbb{Z}}$ onto itself. There



is also the inverse operation σ^{-1} of shifting one place to the right, so that σ is both one-to-one and onto. The composition of σ with itself $k > 0$ times $\sigma^k = \sigma \circ \dots \circ \sigma$ shifts sequences k places to the left, while $\sigma^{-k} = (\sigma^{-1})^k$ shifts the same amount to the right. This shifting operation is the reason $\mathcal{A}^{\mathbb{Z}}$ is called a full shift ("full" since all sequences of symbols are allowed).

The shift map is useful for expressing many of the concepts in symbolic dynamics. For example, one basic idea is that of codes, or rules, which transform one sequence into another. For us, the most important codes are those that do not change with time. Consider the map $\phi\colon \{0,1\}^{\mathbb{Z}} \to \{0,1\}^{\mathbb{Z}}$ defined by the rule $\phi(x) = y$, where $y_i = x_i + x_{i+1} \pmod 2$. Then ϕ is a coding rule that replaces the symbol at index i with the sum modulo 2 of itself and its right neighbor. The coding operation ϕ acts the same at each coordinate, or is *stationary*, i.e., independent of time.

Another way to say this is that applying the rule ϕ and then shifting gives exactly the same result as shifting and then applying ϕ. Going through the following diagram to the right and then down gives the same result as going down and then to the right.

$$\begin{array}{ccc} x & \xrightarrow{\ \sigma\ } & \sigma(x) \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle\phi} \\ \phi(x) & \xrightarrow{\ \sigma\ } & \sigma(\phi(x)) = \phi(\sigma(x)) \end{array}$$

We can express this as $\sigma \circ \phi = \phi \circ \sigma$, or in terms of the coordinates by $\sigma(\phi(x))_{[i]} = \phi(\sigma(x))_{[i]}$, since both equal $x_{i+1} + x_{i+2} \pmod 2$. Recall that when two mappings f and g satisfy $f \circ g = g \circ f$, they are said to *commute*. Not all pairs of mappings commute (try: f = "put on socks" and g = "put on shoes"). Using this terminology, a code ϕ on the full 2-shift is stationary if it commutes with the shift map σ, which we can also express by saying that the following diagram commutes.

$$\begin{array}{ccc} \{0,1\}^{\mathbb{Z}} & \xrightarrow{\ \sigma\ } & \{0,1\}^{\mathbb{Z}} \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle\phi} \\ \{0,1\}^{\mathbb{Z}} & \xrightarrow{\ \sigma\ } & \{0,1\}^{\mathbb{Z}} \end{array}$$

We will discuss codes in more detail in §1.5.

Points in a full shift which return to themselves after a finite number of shifts are particularly simple to describe.

Definition 1.1.3. A point x is *periodic for* σ if $\sigma^n(x) = x$ for some $n \geqslant 1$, and we say that x has *period* n under σ. If x is periodic, the smallest positive integer n for which $\sigma^n(x) = x$ is the *least period* of x. If $\sigma(x) = x$, then x is called a *fixed point* for σ.

If x has least period k, then it has period $2k, 3k, \dots$, and every period of x is a multiple of k (see Exercise 1.1.5). A fixed point for σ must have the form a^∞ for some symbol a, and a point of period n has the form u^∞ for some n-block u.

Iteration of the shift map provides the "dynamics" in symbolic dynamics (see Chapter 6). Naturally, the "symbolic" part refers to the symbols used to form sequences in the spaces we will study.

EXERCISES

1.1.1. How many points $x \in \mathcal{A}^{\mathbb{Z}}$ are fixed points? How many have period n? How many have least period 12?

1.1.2. For the full $\{+1,-1\}$-shift and $k \geqslant 1$, determine the number of k-blocks having the property that the sum of the symbols is 0.

1.1.3. Let ϕ be the coding rule from this section.

(a) Prove that ϕ maps the full 2-shift *onto* itself, i.e., that given a point y in the 2-shift, there is an x with $\phi(x) = y$.

(b) Find the number of points x in the full 2-shift with $\phi^n(x) = 0^\infty$ for $n = 1, 2$, or 3. Can you find this number for every n?

*(c) Find the number of points x with $\phi^n(x) = x$ for $n = 1, 2$, or 3. Can you find this number for every n?

1.1.4. For each k with $1 \leqslant k \leqslant 6$ find the number of k-blocks over $\mathcal{A} = \{0, 1\}$ having no two consecutive 1's appearing. Based on your result, can you guess, and then prove, what this number is for every k?

1.1.5. Determine the least period of u^∞ in terms of properties of the block u. Use your solution to show that if x has period n, then the least period of x divides n.

1.1.6. (a) Describe those pairs of blocks u and v over an alphabet $\mathcal{A}$ such that $uv = vu$.

*(b) Describe those sequences $u_1, u_2, \dots, u_n$ of n blocks for which all n concatenations $u_1u_2 \dots u_n$, $u_2 \dots u_n u_1$, $\dots$, $u_n u_1 u_2 \dots u_{n-1}$ of the cyclic permutations are equal.

§1.2. Shift Spaces

The symbol sequences we will be studying are often subject to constraints. For example, Morse code uses the symbols "dot," "dash," and "pause." The ordinary alphabet is transmitted using blocks of dots and dashes with length at most six separated by a pause, so that any block of length at least seven which contains no pause is forbidden to occur (the only exception is the SOS signal). In the programming language Pascal, a program line such as `sin(x)***2 := y` is not allowed, nor are lines with unbalanced parentheses, since they violate Pascal's syntax rules. The remarkable error correction in compact audio disks results from the use of special kinds of binary sequences specified by a finite number of conditions. In this section we introduce the fundamental notion of shift space, which will be the subset of points in a full shift satisfying a fixed set of constraints.

If $x \in \mathcal{A}^{\mathbb{Z}}$ and w is a block over $\mathcal{A}$, we will say that w *occurs in* x if there are indices i and j so that $w = x_{[i,j]}$. Note that the empty block ε occurs in every x, since $\varepsilon = x_{[1,0]}$. Let $\mathcal{F}$ be a collection of blocks over $\mathcal{A}$, which we will think of as being the *forbidden blocks*. For any such $\mathcal{F}$, define $X_{\mathcal{F}}$ to be the subset of sequences in $\mathcal{A}^{\mathbb{Z}}$ which do *not* contain any block in $\mathcal{F}$.

Definition 1.2.1. A *shift space* (or simply *shift*) is a subset X of a full shift $\mathcal{A}^{\mathbb{Z}}$ such that $X = X_{\mathcal{F}}$ for some collection $\mathcal{F}$ of forbidden blocks over $\mathcal{A}$.

The collection $\mathcal{F}$ may be finite or infinite. In any case it is at most countable since its elements can be arranged in a list (just write down its blocks of length 1 first, then those of length 2, and so on). For a given shift space there may be many collections $\mathcal{F}$ describing it (see Exercise 1.2.4). Note that the empty set $\varnothing$ is a shift space, since putting $\mathcal{F} = \mathcal{A}$ rules out every point. When a shift space X is contained in a shift space Y, we say that X is a *subshift* of Y.

In the equation $X = \mathsf{X}_{\mathcal{F}}$, the notation X refers to the operation of forming a shift space, while X denotes the resulting set. We will sometimes use similar typographical distinctions between an operation and its result, for example in §2.2 when forming an adjacency matrix from a graph. By use of such distinctions, we hope to avoid the type of nonsensical equations such as "$y = y(x)$" you may have seen in calculus classes.

Example 1.2.2. X is $\mathcal{A}^{\mathbb{Z}}$, where we can take $\mathcal{F} = \varnothing$, reflecting the fact that there are no constraints. □

Example 1.2.3. X is the set of all binary sequences with no two 1's next to each other. Here $X = \mathsf{X}_{\mathcal{F}}$, where $\mathcal{F} = \{11\}$. This shift is called the *golden mean shift* for reasons which will surface in Chapter 4. □

Example 1.2.4. X is the set of all binary sequences so that between any two 1's there are an even number of 0's. We can take for $\mathcal{F}$ the collection

$$\{10^{2n+1}1 : n \geqslant 0\}.$$

This example is naturally called the *even shift*. □

In the following examples, the reader will find it instructive to list an appropriate collection $\mathcal{F}$ of forbidden blocks for which $X = \mathsf{X}_{\mathcal{F}}$.

Example 1.2.5. X is the set of all binary sequences for which 1's occur infinitely often in each direction, and such that the number of 0's between successive occurrences of a 1 is either 1, 2, or 3. This shift is used in a common data storage method for hard disk drives (see §2.5). For each pair (d, k) of nonnegative integers with $d \leqslant k$, there is an analogous (d, k) *run-length limited shift*, denoted by $\mathsf{X}(d, k)$, and defined by the constraints that 1's occur infinitely often in each direction, and there are at least d 0's, but no more than k 0's, between successive 1's. Using this notation, our example is $\mathsf{X}(1, 3)$. □

Example 1.2.6. To generalize the previous examples, fix a nonempty subset S of $\{0, 1, 2, \dots\}$. If S is finite, define $X = \mathsf{X}(S)$ to be the set of all binary sequences for which 1's occur infinitely often in each direction, and such that the number of 0's between successive occurrences of a 1 is an integer in S. Thus a typical point in $\mathsf{X}(S)$ has the form

$$x = \dots 1\, 0^{n_{-1}}\, 1\, 0^{n_0}\, 1\, 0^{n_1}\, 1 \dots,$$

where each $n_j \in S$. For example, the (d, k) run-length limited shift corresponds to $S = \{d, d+1, \dots, k\}$.

When S is infinite, it turns out that to obtain a shift space we need to allow points that begin or end with an infinite string of 0's (see Exercise 1.2.8). In this case, we define $X(S)$ the same way as when S is finite, except that we do *not* require that 1's occur infinitely often in each direction. In either case, we refer to $X(S)$ as the *S-gap shift*.

Observe that the full 2-shift is the S-gap shift with $S = \{0, 1, 2, \dots\}$, the golden mean shift corresponds to $S = \{1, 2, 3, \dots\}$, and the even shift to $S = \{0, 2, 4, \dots\}$. As another example, for $S = \{2, 3, 5, 7, 11, \dots\}$ the set of primes, we call $X(S)$ the *prime gap shift*. □

Example 1.2.7. For each positive integer c, the *charge constrained shift*, is defined as the set of all points in $\{+1, -1\}^{\mathbb{Z}}$ so that for every block occurring in the point, the algebraic sum s of the $+1$'s and -1's satisfies $-c \leqslant s \leqslant c$. These shifts arise in engineering applications and often go by the name "DC-free sequences." See Immink [Imm2, Chapter 6]. □

Example 1.2.8. Let $\mathcal{A} = \{e, f, g\}$, and X be the set of points in the full $\mathcal{A}$-shift for which e can be followed only by e or f, f can be followed only by g, and g can be followed only by e or f. A point in this space is then just a bi-infinite path on the graph shown in Figure 1.2.1 This is an example of a *shift of finite type*. These shifts are the focus of the next chapter. □

Example 1.2.9. X is the set of points in the full shift $\{a, b, c\}^{\mathbb{Z}}$ so that a block of the form $ab^m c^k a$ may occur in the point only if $m = k$. We will refer to this example as the *context-free shift*. □

You can make up infinitely many shift spaces by using different forbidden collections $\mathcal{F}$. Indeed, there are uncountably many shift spaces possible (see Exercise 1.2.12). As subsets of full shifts, these spaces share a common feature called *shift invariance*. This amounts to the observation that the constraints on points are given in terms of forbidden blocks alone, and do not involve the coordinate at which a block might be forbidden. It follows that if x is in $X_{\mathcal{F}}$, then so are its shifts $\sigma(x)$ and $\sigma^{-1}(x)$. This can be neatly expressed as $\sigma(X_{\mathcal{F}}) = X_{\mathcal{F}}$. The *shift map* σ_X on X is the restriction to X of the shift map σ on the full shift.

This shift invariance property allows us to find subsets of a full shift that are not shift spaces. One simple example is the subset X of $\{0, 1\}^{\mathbb{Z}}$

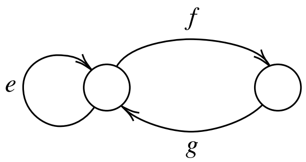


FIGURE 1.2.1. A graph defining a shift space.

consisting of the single point

$$x = \dots 0101.0101 \dots = (01)^\infty.$$

Since $\sigma(x) = (10)^\infty \notin X$, we see that X is not shift invariant, so it is not a shift space.

However, shift invariance alone is not enough to have a shift space. What is missing is a sort of "closure" (see Corollary 1.3.5 and Theorem 6.1.21). This is illustrated by the following example.

Example 1.2.10. Let $X \subseteq \{0,1\}^{\mathbb{Z}}$ be the set of points each of which contains exactly one symbol 1 and the rest 0's. Clearly X is shift invariant. If X were a shift space, then no block of 0's could be forbidden. But then the point $0^\infty = \dots 000.000 \dots$ would necessarily belong to X, whereas it does not. The set X lacks the "closure" necessary for a shift space. □

Since a shift space X is contained in a full shift, Definition 1.1.3 serves to define what it means for $x \in X$ to be fixed or periodic under σ_X. However, unlike full shifts and many of the examples we have introduced, there are shift spaces that contain no periodic points at all (Exercise 1.2.13).

EXERCISES

1.2.1. Find a collection $\mathcal{F}$ of blocks over $\{0,1\}$ so that $\mathsf{X}_{\mathcal{F}} = \varnothing$.

1.2.2. For Examples 1.2.5 through 1.2.9 find a set of forbidden blocks describing the shift space.

1.2.3. Let X be the subset of $\{0,1\}^{\mathbb{Z}}$ described in Example 1.2.10. Show that $X \cup \{0^\infty\}$ is a shift space.

1.2.4. Find two collections $\mathcal{F}_1$ and $\mathcal{F}_2$ over $\mathcal{A} = \{0,1\}$ with $\mathsf{X}_{\mathcal{F}_1} = \mathsf{X}_{\mathcal{F}_2} \neq \varnothing$, where $\mathcal{F}_1$ is finite and $\mathcal{F}_2$ is infinite.

1.2.5. Show that $\mathsf{X}_{\mathcal{F}_1} \cap \mathsf{X}_{\mathcal{F}_2} = \mathsf{X}_{\mathcal{F}_1 \cup \mathcal{F}_2}$. Use this to prove that the intersection of two shift spaces over the same alphabet is also a shift space. Extend this to arbitrary intersections.

1.2.6. Show that if $\mathcal{F}_1 \subseteq \mathcal{F}_2$, then $\mathsf{X}_{\mathcal{F}_1} \supseteq \mathsf{X}_{\mathcal{F}_2}$. What is the relationship between $\mathsf{X}_{\mathcal{F}_1} \cup \mathsf{X}_{\mathcal{F}_2}$ and $\mathsf{X}_{\mathcal{F}_1 \cap \mathcal{F}_2}$?

1.2.7. Let X be the full $\mathcal{A}$-shift.
   (a) Show that if X_1 and X_2 are shift spaces such that $X_1 \cup X_2 = X$, then $X_1 = X$ or $X_2 = X$ (or both).
   (b) Extend your argument to show that if X is the union of any collection $\{X_\alpha\}$ of shift spaces, then there is an α such that $X = X_\alpha$.
   (c) Explain why these statements no longer hold if we merely assume that X is a shift space.

1.2.8. If S is an infinite subset of $\{0,1,2,\dots\}$, show that the collection of all binary sequences of the form

$$x = \dots 1\,0^{n_{-1}}\,1\,0^{n_0}\,1\,0^{n_1}\,1 \dots,$$

where each $n_j \in S$, is not a shift space.

1.2.9. Let X_i be a shift over $\mathcal{A}_i$ for $i = 1, 2$. The *product shift* $X = X_1 \times X_2$ consists of all pairs $(x^{(1)}, x^{(2)})$ with $x^{(i)} \in X_i$. If we identify a pair (x, y) of sequences with the sequence $(\dots(x_{-1}, y_{-1}), (x_0, y_0), (x_1, y_1), \dots)$ of pairs, we can regard $X_1 \times X_2$ as a subset of $(\mathcal{A}_1 \times \mathcal{A}_2)^{\mathbb{Z}}$. With this convention, show that $X_1 \times X_2$ is a shift space over the alphabet $\mathcal{A}_1 \times \mathcal{A}_2$.

1.2.10. Let X be a shift space, and $N \geqslant 1$. Show that there is a collection $\mathcal{F}$ of blocks, all of which have length at least N, so that $X = \mathsf{X}_{\mathcal{F}}$.

1.2.11. For which sets S does the S-gap shift have infinitely many periodic points?

1.2.12. Show there are uncountably many shift spaces contained in the full 2-shift. [*Hint*: Consider S-gap shifts.]

***1.2.13.** Construct a nonempty shift space that does not contain any periodic points.

***1.2.14.** For a given alphabet $\mathcal{A}$, let

$$X = \{x \in \mathcal{A}^{\mathbb{Z}} : x_{i+n^2} \neq x_i \text{ for all } i \in \mathbb{Z} \text{ and } n \geqslant 1\}.$$

(a) If $|\mathcal{A}| = 2$, prove that $X = \varnothing$.
(b) If $|\mathcal{A}| = 3$, show that $X = \varnothing$. [*Hint*: $3^2 + 4^2 = 5^2$.]

§1.3. Languages

It is sometimes easier to describe a shift space by specifying which blocks are allowed, rather than which are forbidden. This leads naturally to the notion of the language of a shift.

Definition 1.3.1. Let X be a subset of a full shift, and let $\mathcal{B}_n(X)$ denote the set of all n-blocks that occur in points in X. The *language of* X is the collection

$$\mathcal{B}(X) = \bigcup_{n=0}^{\infty} \mathcal{B}_n(X).$$

Example 1.3.2. The full 2-shift has language

$$\{\varepsilon, 0, 1, 00, 01, 10, 11, 000, 001, 010, 011, 100, \dots\}. \qquad \square$$

Example 1.3.3. The golden mean shift (Example 1.2.3) has language

$$\{\varepsilon, 0, 1, 00, 01, 10, 000, 001, 010, 100, 101, 0000, \dots\}. \qquad \square$$

The term "language" comes from the theory of automata and formal languages. See [HopU] for a lucid introduction to these topics. Think of the language $\mathcal{B}(X)$ as the collection of "allowed" blocks in X. For a block $u \in \mathcal{B}(X)$, we sometimes use alternative terminology such as saying that u *occurs in* X or *appears in* X or *is in* X or *is allowed in* X.

Not every collection of blocks is the language of a shift space. The following proposition characterizes those which are, and shows that they provide an alternative description of a shift space. In what follows we will denote the complement of a collection $\mathcal{C}$ of blocks over $\mathcal{A}$ relative to the collection of all blocks over $\mathcal{A}$ by $\mathcal{C}^c$.

Proposition 1.3.4.

(1) *Let X be a shift space, and $\mathcal{L} = \mathcal{B}(X)$ be its language. If $w \in \mathcal{L}$, then*
   (a) *every subblock of w belongs to $\mathcal{L}$, and*
   (b) *there are nonempty blocks u and v in $\mathcal{L}$ so that $uwv \in \mathcal{L}$.*

(2) *The languages of shift spaces are characterized by (1). That is, if $\mathcal{L}$ is a collection of blocks over $\mathcal{A}$, then $\mathcal{L} = \mathcal{B}(X)$ for some shift space X if and only if $\mathcal{L}$ satisfies condition (1).*

(3) *The language of a shift space determines the shift space. In fact, for any shift space, $X = \mathsf{X}_{\mathcal{B}(X)^c}$. Thus two shift spaces are equal if and only if they have the same language.*

PROOF: (1) If $w \in \mathcal{L} = \mathcal{B}(X)$, then w occurs in some point x in X. But then every subblock of w also occurs in x, so is in $\mathcal{L}$. Furthermore, clearly there are nonempty blocks u and v such that uwv occurs in x, so that $u, v \in \mathcal{L}$ and $uwv \in \mathcal{L}$.

(2) Let $\mathcal{L}$ be a collection of blocks satisfying (1), and X denote the shift space $\mathsf{X}_{\mathcal{L}^c}$. We will show that $\mathcal{L} = \mathcal{B}(X)$. For if $w \in \mathcal{B}(X)$, then w occurs in some point of $\mathsf{X}_{\mathcal{L}^c}$, so that $w \notin \mathcal{L}^c$, or $w \in \mathcal{L}$. Thus $\mathcal{B}(X) \subseteq \mathcal{L}$. Conversely, suppose that $w = x_0x_1 \dots x_m \in \mathcal{L}$. Then by repeatedly applying (1b), we can find symbols x_j with $j > m$ and x_i with $i < 0$ so that by (1a) every subblock of $x = (x_i)_{i \in \mathbb{Z}}$ lies in $\mathcal{L}$. This means that $x \in \mathsf{X}_{\mathcal{L}^c}$. Since w occurs in x, we have that $w \in \mathcal{B}(\mathsf{X}_{\mathcal{L}^c}) = \mathcal{B}(X)$, proving that $\mathcal{L} \subseteq \mathcal{B}(X)$.

(3) If $x \in X$, no block occurring in x is in $\mathcal{B}(X)^c$ since $\mathcal{B}(X)$ contains all blocks occurring in all points of X. Hence $x \in \mathsf{X}_{\mathcal{B}(X)^c}$, showing that $X \subseteq \mathsf{X}_{\mathcal{B}(X)^c}$. Conversely, since X is a shift there is a collection $\mathcal{F}$ for which $X = \mathsf{X}_{\mathcal{F}}$. If $x \in \mathsf{X}_{\mathcal{B}(X)^c}$, then every block in x must be in $\mathcal{B}(X) = \mathcal{B}(\mathsf{X}_{\mathcal{F}})$, and so cannot be in $\mathcal{F}$. Hence $x \in \mathsf{X}_{\mathcal{F}}$, proving that $X = \mathsf{X}_{\mathcal{F}} \supseteq \mathsf{X}_{\mathcal{B}(X)^c}$. □

This result shows that although a shift X can be described by different collections of forbidden blocks, there is a largest collection $\mathcal{B}(X)^c$, the complement of the language of X. This is the largest possible forbidden collection that describes X. For a minimal forbidden collection, see Exercise 1.3.8. The proposition also gives a one-to-one correspondence between shifts X and languages $\mathcal{L}$ that satisfy (1). This correspondence can be summarized by the equations

$$\mathcal{L} = \mathcal{B}(\mathsf{X}_{\mathcal{L}^c}), \qquad X = \mathsf{X}_{\mathcal{B}(X)^c}. \tag{1–3–1}$$

A useful consequence of part (3) above is that to verify that a point x is in a given shift space X, you only need to show that each subblock $x_{[i,j]}$ is in $\mathcal{B}(X)$. In fact, this gives a characterization of shift spaces in terms of "allowed" blocks.

Corollary 1.3.5. *Let X be a subset of the full $\mathcal{A}$-shift. Then X is a shift space if and only if whenever $x \in \mathcal{A}^{\mathbb{Z}}$ and each $x_{[i,j]} \in \mathcal{B}(X)$ then $x \in X$.*