Chapter 1
General Framework

Valérie Berthé and Michel Rigo

Abstract This introductory chapter briefly presents some of the main notions that appear in the subsequent chapters of this book. We recap a few definitions and results from combinatorics on groups and words, formal language theory, morphic words, k-automatic and k-regular sequences, and dynamical systems. Our aim is not to be exhaustive. The reader can consult this chapter when studying other parts of this book.

1.1 Conventions

The set of nonnegative integers (respectively integers, rational numbers, real numbers, and complex numbers) is written $\mathbb{N}$ (respectively, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$). In particular, the set $\mathbb{N}$ is $\{0, 1, 2, \ldots\}$. We use the notation $[\![i, j]\!]$ for the set of integers $\{i, i+1, \ldots, j\}$. The *floor* of a real number x is $\lfloor x \rfloor = \sup\{z \in \mathbb{Z} \mid z \leq x\}$, whereas $\{x\} = x - \lfloor x \rfloor$ stands for the *fractional part* of x. Recall that $\lceil \cdot \rceil$ denotes the *ceiling function*, i.e., $\lceil x \rceil = \inf\{z \in \mathbb{Z} \mid z \geq x\}$. The characteristic sequence χ_X of a set $X \subset \mathbb{N}^d$ takes its values in $\{0, 1\}$ and satisfies $\chi_X(n) = 1$ if and only if $n \in X$.

Let us recall the notation about asymptotics. Let $f, g : \mathbb{R} \to \mathbb{R}$ be two functions. The definitions given below can also be applied to functions defined on another domain like $\mathbb{R}_{>a}$, $\mathbb{N}$ or $\mathbb{Z}$. We assume implicitly that the following notions are defined for $x \to +\infty$. We write $f \in \mathcal{O}(g)$, if there exist two constants x_0 and $C > 0$ such that, for all $x \geq x_0$, $|f(x)| \leq C|g(x)|$. We also write $f \ll g$ or $g \gg f$, or else $g \in \Omega(f)$. Note that we can write either $f \in \mathcal{O}(g)$ or $f = \mathcal{O}(g)$. Be aware that in the literature, authors sometimes give different meanings to the notation $\Omega(f)$. Here we consider a bound, for all large enough x, but there exist

V. Berthé (✉)
IRIF, UMR 8243, CNRS & Université Paris Diderot, Case 7014, F-75205 Paris Cedex 13, France
e-mail: berthe@irif.fr

M. Rigo
Department of Mathematics, University of Liège, Allée de la découverte 12 (B37),
B-4000 Liège, Belgium
e-mail: M.Rigo@ulg.ac.be

V. Berthé, M. Rigo (eds.), *Sequences, Groups, and Number Theory*,
Trends in Mathematics, https://doi.org/10.1007/978-3-319-69152-7_1

variants where the bound holds only for an increasing sequence $(x_n)_{n\geq 0}$ of reals, i.e., $\limsup_{x\to+\infty} |g(x)|/|f(x)| > 0$.

If g belongs to $\mathcal{O}(f) \cap \Omega(f)$, i.e., there exist constants x_0, C_1, C_2 with $C_1, C_2 > 0$ such that, for all $x \geq x_0$, $C_1|f(x)| \leq |g(x)| \leq C_2|f(x)|$, then we write $g \in \Theta(f)$. As an example, the function $x^2 + \sin 6x$ is in $\Theta(x^2)$ and $x^2|\sin(4x)|$ is in $\mathcal{O}(x^2)$ but not in $\Theta(x^2)$.

1.2 Algebraic Structures

We briefly recall the basic definitions of monoid, (semi)group, (semi)ring, field, ideal, vector space, and module.

Definition 1.2.1. Let $\mathbb{S}$ be a set equipped with a single binary operation

$$\star : \mathbb{S} \times \mathbb{S} \to \mathbb{S}.$$

It is convenient to call this operation a *multiplication* over $\mathbb{S}$, and the product of $x, y \in \mathbb{S}$ is usually denoted by xy.

If this multiplication is *associative*, i.e., for all $x, y, z \in \mathbb{S}$, $(xy)z = x(yz)$, then the algebraic structure given by the pair $(\mathbb{S}, \star)$ is a *semigroup*.

If, moreover, multiplication has an identity element, i.e., there exists some element $1 \in \mathbb{S}$ such that, for all $x \in \mathbb{S}$, $x1 = x = 1x$, then $(\mathbb{S}, \star)$ is a *monoid*.

In addition if every element $x \in \mathbb{S}$ has an *inverse*, i.e., there exists $y \in \mathbb{S}$ such that $xy = 1 = yx$, then $(\mathbb{S}, \star)$ is a *group*.

Definition 1.2.2. A *semiring* is a set R equipped with two binary operations $+$ and $\cdot$ such that

1. $(R, +)$ is a commutative monoid with identity element 0.
2. $(R, \cdot)$ is a monoid with identity element 1.
3. The product is distributive with respect to the sum.
4. For all $r \in R$, $0 \cdot r = 0 = r \cdot 0$.

If, moreover, $\cdot$ is commutative, then the semiring is said to be *commutative*. A *ring* is a semiring where $(R, +)$ is a commutative group. A *field* is a commutative ring where $(R, \cdot)$ is a group.

Definition 1.2.3. A (two-sided) *ideal* of a ring $(R, +, \cdot)$ is a nonempty subset I of R, such that $(I, +)$ is a subgroup of $(R, +)$ and for all $i \in I$ and all $r \in R$, $i \cdot r$ and $r \cdot i$ belong to I.

Definition 1.2.4. Let K be a field with identity element 1 for its multiplication. A *vector space* over K is a set V equipped with a binary operation $+ : V \times V \to V$ such that $(V, +)$ is a commutative group and a binary operation $\cdot : K \times V \to V$ such that, for all $k, \ell \in K$ and all $x, y \in V$,

1. $k \cdot (\ell \cdot x) = (k\ell) \cdot x$
2. $1 \cdot x = x$
3. $(k + \ell) \cdot x = k \cdot x + \ell \cdot x$
4. $k \cdot (x + y) = k \cdot x + k \cdot y$

A *K-module* is similarly defined but it is built over a ring K instead of a field.

We now consider natural notions specific to group and semigroup theory (see also Section 9.3.1 for further basic definitions on group theory and Chapter 11).

For a given property $\mathscr{P}$ of groups (abelian, free, nilpotent, soluble, ...), group G is called *virtually* $\mathscr{P}$ if G contains a finite-index subgroup satisfying property $\mathscr{P}$. See also Definition 9.3.36 and Section 9.3.4.1 for properties of virtually free groups such as the decidability for the word problem (Theorem 9.3.37).

Schreier graphs generalize Cayley graphs. Let G be a group generated by S and acting on a set X, the vertices of its *Schreier graph* (depending on S) are the elements of X, and there is an edge from x to y if y is the image of x under the action of some element of S. By considering the action of the group on itself by right multiplication, this graph coincides with its *Cayley graph*. See also Definition 10.3.1 and Section 11.3.

Let G be a finitely generated group with a generator system given by $S = \{g_1, \dots, g_m\}$. The *length* of $g \in G$ (with respect to S) is the smallest integer ℓ such that g can be represented by a product of the form

$$g = g_{i_1}^{\pm 1} \cdots g_{i_\ell}^{\pm 1},$$

i.e., the length of the shortest decomposition of g. The *growth* of the group G (with respect to S) is the map

$$\gamma_S : \mathbb{N} \to \mathbb{N},\ n \mapsto \mathrm{Card}\{g \in G \mid d_S(g) \leq n\},$$

where $d_S(g)$ is the length of g with respect to S. This definition can be made independent of S by noticing that the growths corresponding to two generating sets are equivalent [409]. Note that a finite group has a bounded growth, an infinite abelian group has a polynomial growth, and a non-abelian free group has an exponential growth. The growth of a finitely generated group can also be seen as the growth of its Cayley graph: we count the vertices which are within distance n of the identity element. This notion is considered in Sections 10.3.4.1, 11.3.1, and 11.4.

1.3 Words

This section is intended to give basic definitions about words either finite or infinite. Words are ubiquitous when encoding a piece of information. As an example, a finite word over the alphabet of digits $\{0, \dots, k-1\}$ can be seen as the k-ary expansion of an integer. On the other hand, an infinite word over $\{0, 1\}$ could be used as the characteristic sequence for a subset of $\mathbb{N}$. For material not covered here, see the classical Lothaire's textbooks on finite or infinite words and their properties are [385–387]. Also see Allouche and Shallit's book [14] about automatic sequences or, Queffélec's book [488] for a dynamical point of view. For a quick overview, the reader can have a look at the chapter [150] or the tutorial [75]. The book [504] is also intended to serve as introductory lecture notes on the subject.

1.3.1 Finite Words

An *alphabet* is a finite nonempty set. Its elements are called *symbols* or *letters*.

Definition 1.3.1. A (finite) *word* over Σ is a finite sequence of letters from Σ. The empty sequence is called the *empty word* and it is denoted by ε. The sets of all finite words (respectively, finite nonempty words) over Σ are denoted by Σ^* (respectively, Σ^+). A word $w = w_0 w_2 \cdots w_{n-1}$ where $w_i \in \Sigma$, $0 \leq i < n$, can be seen as a function $w : \{0, 1, \ldots, n-1\} \to \Sigma$ in which $w(i) = w_i$ for all i. The empty word is the word whose domain is the empty set.

Let $u = u_0 \cdots u_{m-1}$ and $v = v_0 \cdots v_{n-1}$ be two words over Σ. The *concatenation* of u and v is the word $w = w_0 \cdots w_{m+n-1}$ defined by $w_i = u_i$ if $0 \leq i < m$, and $w_i = v_{i-m}$ otherwise. We write $u \cdot v$ or simply uv to express the concatenation of u and v. The concatenation (or catenation) of words is an associative operation, i.e., given three words u, v and w, $(uv)w = u(vw)$. Hence, parenthesis can be omitted. In particular, the set Σ^* (respectively, Σ^+) equipped with the concatenation product is a monoid (respectively, a semigroup).

The *length* of a word w, denoted by $|w|$, is the number of occurrences of the letters in w. In other words, if $w = w_0 w_2 \cdots w_{n-1}$ with $w_i \in \Sigma$, $0 \leq i < n$, then $|w| = n$. In particular, the length of the empty word is zero. The set of words of length k (respectively, at most k) over Σ is denoted by Σ^k (respectively, $\Sigma^{\leq k}$). For $a \in \Sigma$ and $w \in \Sigma^*$, we write $|w|_a$ for the number of occurrences of a in w. Therefore, we have

$$|w| = \sum_{a \in \Sigma} |w|_a \, .$$

If u and v are two words over Σ such that $|u|_a = |v|_a$ for all $a \in \Sigma$, then u is obtained by permuting the letters of v: u and v are said to be *abelian equivalent*. These are anagrams.

A word u is a *factor* of a word v (respectively, a *prefix* or a *suffix*), if there exist words x and y such that $v = xuy$ (respectively, $v = uy$, or $v = xu$). A factor (respectively, a prefix or a suffix) u of a word v is called *proper* if $u \neq v$ and $u \neq \varepsilon$. Prefixes and suffixes are sometimes called initial and terminal factors. Thus, for example, if $w =$ `concatenation`, then `con` is a prefix, `ate` is a factor, and `nation` is a suffix of w. If $w = w_0 \cdots w_n$ and u is a factor of w such that $u = w_i \cdots w_{i+|u|-1}$, we say that u *occurs* in w at position i. For instance, in abbabaabbaab, the factor ab occurs at positions 0, 3, 6, 10. The set of factors of u (respectively, of prefixes of u) is denoted by $\mathrm{Fac}(u)$ (respectively, $\mathrm{Pref}(u)$).

The *mirror* (sometimes called *reversal*) of a word $u = u_0 \cdots u_{m-1}$ is the word $\tilde{u} = u_{m-1} \cdots u_0$. It can be defined inductively on the length of the word by $\tilde{\varepsilon} = \varepsilon$ and $\widetilde{au} = \tilde{u}a$ for $a \in \Sigma$ and $u \in \Sigma^*$. Notice that for $u, v \in \Sigma^*$, $\widetilde{uv} = \tilde{v}\tilde{u}$. A *palindrome* is a word u such that $\tilde{u} = u$. For instance, the palindromes of length at most 3 in $\{0, 1\}^*$ are $\epsilon, 0, 1, 00, 11, 000, 010, 101, 111$.

1.3.2 Infinite Words

Instead of considering finite sequences of elements belonging to an alphabet Σ, considering infinite sequences of elements in Σ is also relevant.

Definition 1.3.2. An (one-sided right) *infinite word* is a map from $\mathbb{N}$ to Σ. If $\mathbf{w}$ is an infinite word, we often write

$$\mathbf{w} = a_0 a_1 a_2 \cdots ,$$

where each $a_i \in \Sigma$. The set of all infinite words of Σ is denoted Σ^ω (one can also find the notation $\Sigma^{\mathbb{N}}$).

Example 1.3.3. Consider the infinite word $\mathbf{x} = x_0 x_1 x_2 \cdots$ where the letters $x_i \in \{0, \dots, 9\}$ are given by the digits appearing in the usual decimal expansion of $\pi - 3$,

$$\pi - 3 = \sum_{i=0}^{+\infty} x_i \, 10^{-i-1},$$

i.e., $\mathbf{x} = 14159265358979323846264338327950288419 \cdots$ is an infinite word.

The notions of *factor*, *prefix*, or *suffix* introduced for finite words can be extended to infinite words. Factors and prefixes are finite words, but a suffix of an infinite word is also infinite. We still make use of the notation $\mathrm{Fac}(\mathbf{w})$ and $\mathrm{Pref}(\mathbf{w})$.

Definition 1.3.4. The *language* of the infinite word $\mathbf{x}$ is the set of all its factors. It is denoted by $\mathrm{Fac}(\mathbf{x})$. The set of factors of length n occurring in $\mathbf{x}$ is denoted by $\mathrm{Fac}_n(\mathbf{x})$.

Definition 1.3.5. The *complexity function*, or *factor complexity*, of an infinite word $\mathbf{x}$ maps $n \in \mathbb{N}$ onto the number $p_{\mathbf{x}}(n) = \mathrm{Card}(\mathrm{Fac}_n(\mathbf{x}))$ of distinct factors of length n occurring in $\mathbf{x}$.

Example 1.3.6. The Thue–Morse word $\mathbf{t} = t_0 t_1 t_2 \cdots$ (ubiquitous word encountered in combinatorics on words [18]) can be defined over $\{\mathtt{a}, \mathtt{b}\}$ by $t_n = \mathtt{a}$ if and only if there is an even number of ones in the base-2 expansion of $n \geq 0$. Otherwise stated, if the sum of base-2 digits of n is even. Thus a prefix of $\mathbf{t}$ is given

$$\mathtt{abbabaabbaababbabaababbaabbabaab} \cdots .$$

If we replace a with 1 and b with 0, then we get the characteristic sequence χ_E of the set of integers whose sum of base-2 digits is even. The factor complexity of the Thue–Morse word **t** is well known [107, 391]. See also [78, p. 225] where a chapter is devoted to the factor complexity of morphic words. We have

$$p_{\mathbf{t}}(n) = \begin{cases} 4n - 2 \cdot 2^m - 4, & \text{if } 2 \cdot 2^m < n \leq 3 \cdot 2^m; \\ 2n + 4 \cdot 2^m - 2, & \text{if } 3 \cdot 2^m < n \leq 4 \cdot 2^m. \end{cases}$$

Definition 1.3.7. A two-sided or *bi-infinite word* is a map from $\mathbb{Z}$ to Σ. The set of all bi-infinite words is denoted ${}^{\omega}\Sigma^{\omega}$ (one can also find the notation $\Sigma^{\mathbb{Z}}$).

Definition 1.3.8. An infinite word $\mathbf{x} = x_0 x_1 \cdots$ is *(purely) periodic* if there exists a finite word $u = u_0 \cdots u_{k-1} \neq \epsilon$ such that $x = u^{\omega}$, i.e., for all $n \geq 0$, we have $x_n = u_r$ where $n = dk + r$ with $r \in \{0, \ldots, k-1\}$. An infinite word x is *eventually periodic* (or *ultimately periodic*) if there exist two finite words $u, v \in \Sigma^*$, with $v \neq \epsilon$ such that $x = uvvv \cdots = uv^{\omega}$. Notice that purely periodic words are special cases of eventually periodic words. For any eventually periodic word x, there exist words u, v of shortest length such that $x = uv^{\omega}$, then the integer $|u|$ (respectively $|v|$) is referred to as the *preperiod* (respectively *period*) of x. An infinite word is said to be *nonperiodic* if it is not eventually periodic.

Let us mention the next result called Morse–Hedlund theorem.

Theorem 1.3.9. *Let* **w** *be an infinite word over a finite alphabet. The word* **w** *is eventually periodic if and only if there exists some integer N such that $p_{\mathbf{w}}(N) \leq N$.*

Among the nonperiodic words of low factor complexity, Sturmian words play a special role and have been extensively studied. An infinite word **x** is *Sturmian* if $p_{\mathbf{x}}(n) = n + 1$ for all $n \geq 0$. Note that Sturmian words are over a 2-letter alphabet. For general references, see [386, Chapter 2] or [487, Chapter 6]. They will be considered in Chapter 6.

Definition 1.3.10. An infinite word **x** is *recurrent* if all its factors occur infinitely often in **x**. It is *uniformly recurrent* if it is recurrent and for every factor u of **x**, for the infinite set

$$\left\{ i_1^{(u)} < i_2^{(u)} < i_3^{(u)} < \cdots \right\}$$

of positions where u occurs in **x**, there exists a constant C_u such that, for all $j \geq 1$,

$$i_{j+1}^{(u)} - i_j^{(u)} \leq C_u \, .$$

Note that, by Furstenberg's theorem, for any infinite word **w**, there is a uniformly recurrent word **r** over the same alphabet such that every finite factor of **r** is a factor of **w**, i.e., $\mathrm{Fac}(\mathbf{r}) \subseteq \mathrm{Fac}(\mathbf{w})$ (see Theorem 4.4.9).

Let **x** be an infinite word, the function $R_{\mathbf{x}} : \mathrm{Fac}(\mathbf{x}) \to \mathbb{N} \cup \{\infty\}$ maps a factor u of **x** to the smallest k such that every factor of **x** of length k contains u, or ∞ if no such k exists. Otherwise stated, an infinite word $\mathbf{x}$ is uniformly recurrent, if for every factor u of $\mathbf{x}$, $R_{\mathbf{x}}$ is finite. The *recurrence function* maps $n \in \mathbb{N}$ to

$$R_{\mathbf{x}}(n) = \max_{u \in L_n(\mathbf{x})} R_{\mathbf{x}}(u)\,.$$

Otherwise stated, if $\mathbf{x}$ is uniformly recurrent, then for every factor of length n of $\mathbf{x}$, $R_{\mathbf{x}}(n)$ is finite and u occurs in all factors of length $R_{\mathbf{x}}(n)$ of $\mathbf{x}$.

Assume that Σ is totally ordered: $(\Sigma, <)$. Let $\mathbf{x}, \mathbf{y}$ be two infinite words over Σ. We say that $\mathbf{x}$ is *lexicographically less* than $\mathbf{y}$ if there exists N such that $x_i = y_i$ for all $i < N$ and $x_N < y_N$.

Definition 1.3.11. One can endow Σ^ω with a *distance* d defined as follows. Let $\mathbf{x}, \mathbf{y}$ be two infinite words over Σ. Let $\mathbf{x} \wedge \mathbf{y}$ denote the longest common prefix of $\mathbf{x}$ and $\mathbf{y}$. Then the distance d is given by

$$d(\mathbf{x}, \mathbf{y}) := \begin{cases} 0, & \text{if } \mathbf{x} = \mathbf{y}, \\ 2^{-|\mathbf{x} \wedge \mathbf{y}|}, & \text{otherwise.} \end{cases}$$

This notion of distance extends to $\Sigma^{\mathbb{Z}}$. Notice that the topology on Σ^ω is the product topology (of the discrete topology on Σ). The space Σ^ω is a compact *Cantor set*, that is, a totally disconnected compact space without isolated points. Since Σ^ω is a (complete) metric space, it is therefore relevant to speak of convergent sequences of infinite words. The sequence $(\mathbf{z}_n)_{n \geq 0}$ of infinite words over Σ *converges* to $\mathbf{x} \in \Sigma^\omega$, if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that, for all $n \geq N$, $d(\mathbf{z}_n, \mathbf{x}) < \epsilon$. To express the fact that a sequence of finite words $(w_n)_{n \geq 0}$ over Σ converges to an infinite word $\mathbf{y}$, it is assumed that Σ is extended with an extra letter $c \notin \Sigma$. Any finite word w_n is replaced with the infinite word $w_n ccc \cdots$, and if the sequence of infinite words $(w_n ccc \cdots)_{n \geq 0}$ converges to $\mathbf{y}$, then the sequence $(w_n)_{n \geq 0}$ is said to converge to $\mathbf{y}$.

Let $(u_n)_{n \geq 0}$ be a sequence of nonempty finite words. If we define, for all $\ell \geq 0$, the finite word v_ℓ as the concatenation $u_0 u_1 \cdots u_\ell$, then the sequence $(v_\ell)_{\ell \geq 0}$ of finite words converges to an infinite word. This latter word is said to be the concatenation of the elements in the infinite sequence of finite words $(u_n)_{n \geq 0}$. In particular, for a constant sequence $u_n = u$ for all $n \geq 0$, $v_\ell = u^{\ell+1}$ and the concatenation of an infinite number of copies of the finite word u is denoted by u^ω.

We have discussed the fact that a (finite) word u may appear as a factor of an infinite word $\mathbf{x}$. It may occur a finite number of times, infinitely often, or even in such a way that $R_{\mathbf{x}}(u)$ is finite. But we could also introduce the *frequency* of a factor u occurring in $\mathbf{x}$ as the following limit, if it exists,

$$\lim_{n \to +\infty} \frac{\mathrm{Card}\left(\{i \leq n - |u| \mid x_i \cdots x_{i+|u|-1} = u\}\right)}{n}\,.$$

For instance, for the infinite word $\mathbf{w} = 01\,0011\,0^4 1^4\,0^8\,1^8\,0^{16} 1^{16} \cdots$ where we have longer and longer blocks of consecutive zeroes followed by longer and longer blocks of ones. The frequencies of 0 and 1 do not exist. Frequency appears naturally in the definition of normal numbers given below. See also Theorem 1.6.10 about the frequency of symbols in automatic sequences and morphic words. Frequencies are also considered in Chapter 5 in the framework of repetitions, and in Chapter 7 and 8 in the framework of normality.

1.3.3 Number Representations

We refer the reader to Frougny's chapter [386] or to [227] for a general presentation of numeration systems. The book [503] can also serve as an introduction to the subject. We also mention the survey [36]. More details are also discussed in Section 3.2 of this book.

Let $k \geq 2$ be an integer. Let us recall how base-k expansion of integers may be computed. For any positive integer n, there exist $\ell \geq 0$ such that $k^\ell \leq n < k^{\ell+1}$ and unique coefficients $c_0, \dots, c_\ell \in \{0, \dots, k-1\}$ such that

$$n = \sum_{i=0}^{\ell} c_i\, k^i \text{ and } c_\ell \neq 0\,.$$

The coefficients $c_\ell, \dots, c_0$ can be computed by successive Euclidean divisions. Set $n_0 := n$. We have $n_0 = c_\ell\, k^\ell + n_1$ with $n_1 < k^\ell$ and for $i = 1, \dots, \ell, n_i = c_{\ell-i}\, k^{\ell-i} + n_{i+1}$ with $n_{i+1} < k^{\ell-i}$. The word $c_\ell \cdots c_0$ is said to be the *k-ary representation* or *k-ary expansion* of n (sometimes called *greedy* representation) and denoted by $\mathrm{rep}_p(n)$. If $d_\ell \cdots d_0$ is a word over an alphabet of digits included in $\mathbb{Z}$, we define

$$\mathrm{val}_k(d_\ell \cdots d_0) = \sum_{i=0}^{\ell} d_i\, k^i\,.$$

If one replaces the sequence $(k^n)_{n\geq 0}$ with an increasing sequence $(U_n)_{n\geq 0}$ of integer such that $U_0 = 1$, then a similar algorithm may be applied. The corresponding U-expansions are over the alphabet $\{0, \dots, \sup\lceil \frac{U_{n+1}}{U_n} \rceil - 1\}$. One finds the general terminology *positional numeration system*. It is also possible to extend the procedure to represent real numbers. Let $x \in (0, 1)$. There exists a decomposition of the form

$$x = \sum_{i=1}^{+\infty} c_i\, k^{-i}$$

where $c_i \in \{0, \dots, k-1\}$ for all $i \geq 1$. If we forbid sequences where $c_i = k - 1$ for all large enough i, then the sequence $(c_i)_{i\geq 1}$ is unique. Given $x \in [0, 1)$, the algorithm in Table 1.1 provides the corresponding sequence $(c_i)_{i\geq 0}$ of digits.

Table 1.1 An algorithm for computing the base-k expansion of $x \in [0,1)$.

```
$i \leftarrow 0$
$y \leftarrow x$
REPEAT FOREVER
    $c_i \leftarrow \lfloor ky \rfloor$
    $y \leftarrow \{ky\}$
    INCREMENT $i$
END-REPEAT.
```

In this algorithm, we iterate a map from the interval $[0, 1)$ onto itself, i.e.,

$$T_k : [0,1) \to [0,1), y \mapsto \{ky\} \tag{1.1}$$

and the value taken by the image determines the next digit in the expansion. This yields a dynamical system such as discussed in Section 1.7. The interval $[0, 1)$ is thus split into k subintervals $[j/k, (j+1)/k)$, for $j = 0, \ldots, k-1$. For all $i \geq 0$, if $T_k^i(x)$ belongs to the subinterval $[j/k, (j+1)/k)$, then the digit c_i occurring in $\text{rep}_k(x)$ is equal to j. It is indeed natural to consider such subintervals. If y belongs to $[j/k, (j+1)/k)$, then ky has an integer part equal to j and the map T_k is continuous and increasing on every subinterval $[j/k, (j+1)/k)$. Note also that the range of T_k on any of these subintervals is $[0, 1)$. So applying T_k to a point in one of these subintervals can lead to a point belonging to any of these subintervals (later on, we shall introduce some other transformation, e.g., β-transformations, where a restriction appears on the intervals that can be reached). So to speak, the base-k expansion of x can be derived from the trajectory of x under T_k, i.e., from the sequence $(T_k^n(x))_{n\geq 0}$.

As an example, consider the base $k = 3$ and the expansion of $x = 3/10$. The point lies in the interval $[0, 1/3)$; thus the first digit of the expansion is 0. Then $T_3(3/10) = 9/10$ lies in the interval $[2/3, 1)$; thus the second digit is 2. If we apply again T_3, we get $T_3^2(3/10) = \{27/10\} = 7/10$, which belongs again to $[2/3, 1)$ giving the digit 2. Then $T_3^3(3/10) = 1/10$ giving the digit 0 and finally $T_3^4(3/10) = 3/10$. So $\text{rep}_3(3/10) = (0220)^\omega$.

A natural generalization of base-k expansion (discussed in Section 3.6 and in Example 8.1.2) is to replace the base k with a real number $\beta > 1$. In particular, the transformation T_k will be replaced by the so-called β-transformation. Note that we shall be concerned with expansions of numbers in $[0, 1)$. If $x \geq 1$, then there exists a smallest d such that x/β^d belongs to $[0, 1)$. It is therefore enough[1] to concentrate on $[0, 1)$.

Definition 1.3.12 (β-Expansions). We will only represent real numbers in the interval $[0, 1)$. Let $\beta > 1$ be a real number. The representations discussed here

[1] If the β-expansion of x/β^d is $d_0 d_1 \cdots$, then using an extra decimal point, the expansion of x is conveniently written $d_0 \cdots d_{\ell-1} \bullet d_\ell d_{\ell+1} \cdots$. Note that the presentation in Chapter 1 is not entirely consistent with our present treatment if x belongs to $[0, 1/(\beta-1)] \setminus [0, 1)$.

are a direct generalization of the base-k expansions. Every real number $x \in [0, 1)$ can be written as a series

$$x = \sum_{i=0}^{+\infty} c_i\, \beta^{-i-1} \tag{1.2}$$

where c_i belong to $\{0, \lceil\beta\rceil - 1\}$. Note that if β is an integer, then $\lceil\beta\rceil - 1 = \beta - 1$. For integer base-$b$ expansions, a number may have more than one representation, namely, those ending with 0^ω or $(b-1)^\omega$. For a real base β, we obtain many more representations. Consider the Golden mean ϕ, which satisfies $\phi^2 - \phi - 1 = 0$, and thus

$$\frac{1}{\phi^n} = \frac{1}{\phi^{n+1}} + \frac{1}{\phi^{n+2}}, \quad \forall n \geq 0\,.$$

As an example, the number $1/\phi$ has thus infinitely many representations as a power series with negative powers of ϕ and coefficients 0 and 1:

$$\frac{1}{\phi} = \frac{1}{\phi^2} + \frac{1}{\phi^3} = \frac{1}{\phi^2} + \frac{1}{\phi^4} + \frac{1}{\phi^5} = \frac{1}{\phi^2} + \frac{1}{\phi^4} + \frac{1}{\phi^6} + \frac{1}{\phi^7} = \cdots\,.$$

To get a canonical expansion for a real $x \in [0, 1)$, we just have to replace the integer base b with β and consider the so-called *β-transformation*

$$T_\beta : [0, 1) \to [0, 1),\ x \mapsto \{\beta x\}$$

in the algorithm from Table 1.1. For $i = 0, 1, \ldots$, the idea is to remove the largest integer multiple c_i of β^{-i-1} and then repeat the process with the remainder and the next negative power of β to get (1.2). Note that c_i is less than $\lceil\beta\rceil$ because of the greediness of the process. Otherwise, one could have removed a larger multiple of the power of β at a previous step. The corresponding infinite word $c_0c_1\cdots$ is called the *β-expansion* of x and is usually denoted by $\mathsf{d}_\beta(x)$. Any word $d_0d_1\cdots$ over a finite alphabet of nonnegative integers satisfying

$$x = \sum_{i=0}^{+\infty} d_i\, \beta^{-i-1}$$

is said to be a *β-representation* of x. Thus, the β-expansion of x is the lexicographically maximal word among the β-representations of x.

The greediness of the algorithm can be reformulated as follows.

Lemma 1.3.13. *A word $d_0d_1\cdots$ over $\{0, \ldots, \lceil\beta\rceil - 1\}$ is the β-expansion of a real number $x \in [0, 1)$ if and only if, for all $j \geq 0$,*

$$\sum_{i=j}^{+\infty} d_i\, \beta^{-i-1} < \beta^{-j}\,.$$

Proposition 1.3.14. *Let x, y be real numbers in $[0,1)$. We have $x < y$ if and only if $\mathsf{d}_\beta(x)$ is lexicographically less than $\mathsf{d}_\beta(y)$.*

1.3.4 Normality

Now that number representations and the frequency of a factor have been introduced, we can define normal numbers.

A real number x is *simply normal* with respect to base $b \geq 2$ if in the base-b expansion of x (which is an infinite word over $\{0,\dots,b-1\}$), the frequency of every digit $d \in \{0,1,\dots,b-1\}$ exists and is equal to $1/b$. Furthermore x is *normal* in base b if it is simply normal with respect to the bases b, b^2, b^3,.... An equivalent definition is to say that for all $k \geq 1$ and every word $u = u_1 \dots u_k \in \{0,1,\dots,b-1\}^k$, the frequency of u in the base-b expansion of x exists and is equal to $1/b^k$. A real number x is *absolutely normal* if x is normal to every integer base greater than or equal to 2.

Normality can also be expressed in terms of uniform distribution modulo 1 [578] (see Section 7.6 for corresponding definitions). Indeed, a real number x is normal to base b if and only if the sequence $(b^j x)_{j\geq 0}$ is uniformly distributed modulo 1.

These notions were introduced by Borel [99] and are discussed in Chapters 2, 7, and 8. In particular, constructions of normal numbers are provided in Sections 7.7 and 7.8. See also Theorem 7.4.1 (the so-called Hot Spot Lemma according to [101]) for a further convenient characterization of normality in terms of limsups instead of limits. For a dynamical viewpoint, see Section 8.2, where the definition of a normal number is transferred to symbolic dynamical systems, and constructions with concatenation of words for languages with specification are provided.

1.3.5 Repetitions in Words

In combinatorics on words, a question that naturally arises is to study the repetitions that should occur or may be avoided in words. See in particular Chapter 5 and Chapters 4 and 5 in [79].

Concatenating a word w with itself k times is abbreviated by w^k. In particular, $w^0 = \varepsilon$. Furthermore, for an integer m and a word $w = w_1 w_2 \cdots w_n$, where $w_i \in \Sigma$ for $1 \leq i \leq n$ (here it is convenient to start indexing with 1), the *rational power*

$$w^{m/n}$$

is $w^q w_1 w_2 \cdots w_r$, where $m = qn + r$ for $0 \leq r < n$. For instance, we have

$$(\texttt{abbab})^{9/5} = \texttt{abbababba}\,.$$

Consider definitions that have to do with repetitions in words. A *square* is a nonempty word of the form xx, where $x \in \Sigma^*$. An example of a square in English is the word murmur with x equal to mur. An *overlap* is a word of the form $axaxa$, where $a \in \Sigma$ and $x \in \Sigma^*$. The word alfalfa is an example of an overlap in English with x equal to lf. It is obvious that every overlap has a square as prefix. For any positive integer $k \geq 2$, a *k-power* is a nonempty word of the form x^k. Thus a 2-power is a square, and a 3-power is a *cube*. A nonempty word that is not a k-power for any $k \geq 2$ is *primitive*.

Let us say a few words about avoidance (which is the topic of Chapter 5). It is an easy exercise to show that over a 2-letter alphabet, every word of a length of at least 4 contains a square. This raises several questions. Over a 3-letter alphabet, can we build longer words with no square as a factor? In particular, does there exist an infinite word with no square in it? Also over a 2-letter alphabet, if squares cannot be avoided, could we avoid cubes or even overlaps?

We say that a word w (finite or infinite) is *square-free* (or avoids squares) if no factor of w is a square. A finite or infinite word is *overlap-free* if it contains no factor that is an overlap. Thue [563] was the first to show the existence of an infinite overlap-free binary word. The Thue–Morse word (see Example 1.3.6) is overlap-free. See [79, Chapter 4] for more on avoidable repetitions and regularities in words. More generally, a (finite or infinite) word is *k-power-free* (or *avoids k-powers*) if none of its factors is a k-power. For instance, one can check that abbabaabbaab is overlap-free. (It is indeed a prefix of the Thue–Morse word). The goal of Chapter 5 is to present general techniques to prove positive or negative results about the appearance of a repetition pattern. The general question is to know whether an infinite word without a given pattern exists over an alphabet of a given size. Another question is to consider the growth function (in the sense of Definition 1.5.7) of the language of finite words avoiding a particular pattern.

Many variations on these topics exist. For instance, an abelian square is a word of the form uv where u and v are abelian equivalent. One can check that over a 3-letter alphabet, every long enough finite word contains an abelian square.

In Chapter 6, the addressed question is this: given a nonperiodic word $\mathbf{x} \in \Sigma^\omega$, does there exist a finite nonempty set C and a mapping $\varphi : \Sigma^+ \to C$ such that for each factorization $\mathbf{x} = u_1 u_2 u_3 \cdots$ there exist $i, j \geq 1$ such that $\varphi(u_i) \neq \varphi(u_j)$?

1.4 Morphisms

Infinite words of particular interest can be obtained by iterating morphisms of free monoids. They have many interesting combinatorial properties and can be generated by a simple mean.

Definition 1.4.1. A map $h : \Sigma^* \to \Delta^*$, where Σ and Δ are alphabets, is called a *morphism* if h satisfies $h(xy) = h(x)h(y)$ for all $x, y \in \Sigma^*$. In particular, we have $h(\varepsilon) = \varepsilon$. When $\Sigma = \Delta$, morphisms are also called *substitutions*.

A morphism may be specified by providing the values $h(a)$ for all $a \in \Sigma$. For example, we may define the morphism $t : \{0, 1\}^* \to \{0, 1\}^*$ by

$$\begin{aligned} 0 &\mapsto 01 \\ 1 &\mapsto 10. \end{aligned} \tag{1.3}$$

This morphism is often referred to as the *Thue–Morse morphism*. The domain Σ^* of a morphism h is easily extended to the set Σ^ω of (one-sided) infinite words. Let $h : \Sigma^* \to \Delta^*$ be a morphism and $\mathbf{x} = x_0x_1x_2 \cdots$ be an infinite word over Σ. Simply consider the sequence of finite words $(h(x_0 \cdots x_n))_{n\geq 0}$ of images of the prefixes of $\mathbf{x}$. The limit of this sequence is $h(\mathbf{x})$. In particular, if $h : \Sigma^* \to \Sigma^*$ and $\mathbf{x}$ is an infinite word such that $h(\mathbf{x}) = \mathbf{x}$, then $\mathbf{x}$ is said to be a *fixed point* of h.

A morphism $h : \Sigma^* \to \Sigma^*$ such that $h(a) = ax$ for some $a \in \Sigma$ and $x \in \Sigma^*$ with $h^i(x) \neq \epsilon$ for all i is said to be *prolongable on a*. The Thue–Morse morphism t given by (1.3) is prolongable on 0 (and also on 1). The first few iterations of t are

$$\begin{aligned} t(0) &= 01 \\ t^2(0) &= 0110 \\ t^3(0) &= 01101001 \\ t^4(0) &= 0110100110010110 \\ &\vdots \end{aligned}$$

Since $|t(0)| = |t(1)| = 2$, we have $|t^n(0)| = 2^n$ for all $n \geq 0$. It is easy to prove that $t^n(0)$ is a proper prefix of $t^{n+1}(0)$, and thus the sequence $(t^n(0))_{n\geq 0}$ converges to an infinite word. So we get the fixed point of t

$$t^\omega(0) = 0110100110010110 \cdots .$$

One can prove that the fixed point $t^\omega(0)$ is the *Thue–Morse word* introduced in Example 1.3.6.

More generally, if $h : \Sigma^* \to \Sigma^*$ is a morphism prolongable on a, we may then repeatedly iterate h to obtain the infinite *fixed point*

$$h^\omega(a) = a\,x\,h(x)\,h^2(x)\,h^3(x) \cdots .$$

This infinite word is said to be *purely morphic*.

The factor complexity of purely morphic word is well known. The next result was stated by Pansiot in [467] and then generalized in [468]. For a comprehensive presentation, see [78, Section 4.7]. Recall that the case of eventually periodic words is settled by Morse–Hedlund theorem.

Theorem 1.4.2. *Let* $\mathbf{w}$ *be a pure morphic word. If* $\mathbf{w}$ *is not eventually periodic, then its factor complexity* $p_{\mathbf{w}}$ *belongs to* $\Theta(n)$*,* $\Theta(n \log \log n)$*,* $\Theta(n \log n)$*, or* $\Theta(n^2)$.

Definition 1.4.3. A morphism h is *non-erasing* if $h(a) \neq \epsilon$ for all $a \in \Sigma$. Otherwise it is *erasing*. A morphism is k*-uniform* if $|h(a)| = k$ for all $a \in \Sigma$; it is *uniform* if it is k-uniform for some k. A 1-uniform morphism is often said to be a *letter-to-letter morphism* or a *coding*.

The Thue–Morse morphism t given in (1.3) is 2-uniform.

Example 1.4.4 (Fibonacci Word). Another significant example of a purely morphic word is the *Fibonacci word*. It is obtained from the non-uniform morphism defined over the alphabet $\{0, 1\}$ by $\sigma : 0 \mapsto 01, 1 \mapsto 0$,

$$\sigma^{\omega}(0) = (x_n)_{n \geq 0} = 0100101001001010010100100101001001010010100 \cdots .$$

It is a Sturmian word and can be obtained as follows. Let $\phi = (1 + \sqrt{5})/2$ be the Golden mean. For all $n \geq 1$, if $\lfloor (n + 1)\phi \rfloor - \lfloor n\phi \rfloor = 2$, then $x_{n-1} = 0$; otherwise $x_{n-1} = 1$.

An infinite word $\mathbf{x}$ over Δ is *morphic* if there exists a purely morphic word $\mathbf{y}$ over Σ and a morphism $g : \Sigma^* \to \Delta^*$ such that $\mathbf{x} = g(\mathbf{y})$.

We can always restrict ourselves to non-erasing prolongable morphisms and codings. This result was already stated in [154]. J.-J. Pansiot also considered this result in [466]. For a proof, see [14]. An alternative short proof is given in [298]. This result is also discussed in detail in [134] and [146].

Theorem 1.4.5. *Let* $f : \Sigma^* \to \Sigma^*$ *be a (possibly erasing) morphism that is prolongable on a letter* $a \in \Sigma$*. Let* $g : \Sigma^* \to \Gamma^*$ *be a (possibly erasing) morphism. If the word* $g(f^{\omega}(a))$ *is infinite, there exists a non-erasing morphism* $h : \Delta^* \to \Delta^*$ *prolongable on a letter* $c \in \Delta$ *and a coding* $j : \Delta^* \to \Gamma^*$ *such that* $g(f^{\omega}(a)) = j(h^{\omega}(c))$.

1.5 Languages and Machines

Formal languages theory is mostly concerned with the study of the mathematical properties of sets of words. For a comprehensive exposition on regular (or rational) languages and automata theory, see, for instance, Sakarovitch's book [518]. For the connections with infinite words, see [476]. For an overview see the chapter [590]. Finally see [555], Hopcroft and Ullman's classic book [301], or its updated version [300] for general books on formal languages theory.

1.5.1 Languages of Finite Words

Let Σ be an alphabet. A subset L of Σ^* is said to be a *language*. Since a language is a *set* of words, we can apply all the usual set operations like union, intersection, or set difference: $\cup$, $\cap$, or $\setminus$. The concatenation of words can be extended to define an operation on languages. If L, M are languages, LM is the language of the words obtained by concatenation of a word in L and a word in M, i.e.,

$$LM = \{uv \mid u \in L, v \in M\}\,.$$

We can of course define the concatenation of a language with itself, so it permits us to introduce the power of a language. Let $n \in \mathbb{N}$, Σ be an alphabet, and $L \subseteq \Sigma^*$ be a language. The language L^n is the set of words obtained by concatenating n words in L. We set $L^0 := \{\epsilon\}$. In particular, we recall that Σ^n denotes the set of words of length n over Σ, i.e., concatenations of n letters in Σ. The *(Kleene) star* of the language L is defined as

$$L^* = \bigcup_{i \geq 0} L^i\,.$$

Otherwise stated, L^* contains the words that are obtained as the concatenation of an arbitrary number of words in L. Notice that the definition of Kleene star is compatible with the notation Σ^* introduced to denote the set of finite words over Σ. We also write $L^{\leq n}$ as a shorthand for

$$L^{\leq n} = \bigcup_{i=0}^{n} L^i\,.$$

Note that if the empty word belongs to L, then $L^{\leq n} = L^n$. We recall that $\Sigma^{\leq n}$ is the set of words over Σ of length at most n.

Example 1.5.1. Let $L = \{a, ab, aab\}$ and $M = \{a, ab, ba\}$ be two finite languages. We have

$$L^2 = \{aa, aab, aaab, aba, abab, abaab, aaba, aabab, aabaab\}$$

and

$$M^2 = \{aa, aab, aba, abab, abba, baa, baab, baba\}\,.$$

One can notice that $\mathrm{Card}\,(L^2) = (\mathrm{Card}\,L)^2$ but $\mathrm{Card}\,(M^2) < (\mathrm{Card}\,M)^2$. This is due to the fact that all words in L^2 have a unique factorization as concatenation of two elements in L, but this is not the case for M, where $(ab)a = a(ba)$. We can notice that

$$L^* = \{a\}^* \cup \{a^{i_1}ba^{i_2}b\cdots a^{i_n}ba^{i_{n+1}} \mid \forall n \geq 1, i_1, \ldots, i_n \geq 1,\ i_{n+1} \geq 0\}\,.$$

Since languages are sets of (finite) words, a language can be either *finite* or *infinite*. For instance, a language L differs from $\emptyset$ or $\{\epsilon\}$ if and only if the language L^* is infinite. Let L be a language, we set $L^+ = LL^*$. The mirror operation can also be extended from words to languages: $\tilde{L} = \{\tilde{u} \mid u \in L\}$.

Definition 1.5.2. A language is *prefix-closed* (respectively *suffix-closed*) if it contains all prefixes (respectively suffixes) of any of its elements. A language is *factorial* if it contains all factors of any of its elements.

Obviously, any factorial language is prefix-closed and suffix-closed. The converse does not hold. For instance, the language $\{a^nb \mid n > 0\}$ is suffix-closed but not factorial.

Example 1.5.3. Connected with the Thue–Morse word (see Example 1.3.6), the set of words over $\{0, 1\}$ containing an even number of ones is the language

$$\begin{aligned} E &= \{w \in \{0,1\}^* \mid |w|_1 \equiv 0 \pmod 2\} \\ &= \{\epsilon, 0, 00, 11, 000, 011, 101, 110, 0000, 0011, \ldots\}. \end{aligned}$$

This language is closed under mirror, i.e., $\tilde{L} = L$. Notice that the concatenation $E\{1\}E$ is the language of words containing an odd number of ones and $E \cup E\{1\}E = E(\{\epsilon\} \cup \{1\}E) = \{0,1\}^*$. Notice that E is neither prefix-closed, since $1001 \in E$ but $100 \notin E$, nor suffix-closed.

Definition 1.5.4. The set of factors of a language L is denoted as $\mathrm{Fac}(L)$, whereas the set of prefixes of a language L is denoted as $\mathrm{Pref}(L)$. The notation $w^{-1}L$ stands for $w^{-1}L = \{u \mid wu \in L\}$.

If a language L over Σ can be obtained by applying to some finite languages a finite number of operations of union, concatenation, and Kleene star, then this language is said to be a *regular language*. This generation process leads to *regular expressions* which are well-formed expressions used to describe how a regular language is built in terms of these operations.

Note that the *Chomsky–Schützenberger hierarchy* introduced in the theory of formal languages provides a classification depending on the machine needed to recognize an infinite language of finite words. From a computational perspective, the simplest languages are the regular languages. They are accepted (or recognized) by finite automata, and described by regular expressions. One then has context-free languages that are recognized by non-deterministic pushdown automata, context-sensitive languages recognized by linear-bounded non-deterministic Turing machines, and lastly, recursively enumerable languages recognized by Turing machines. See Section 2.1.2 for a similar hierarchy for Mahler functions and regular sequences.

From the definition of a regular language, the following result is immediate.

Theorem 1.5.5. *The class of regular languages over Σ is the smallest subset of 2^{Σ^*} (for inclusion) containing the languages $\emptyset$, $\{a\}$ for all $a \in \Sigma$ and closed under union, concatenation, and Kleene star.*

Example 1.5.6. For instance, the language L over $\{0,1\}$ whose words do not contain the factor 11 is regular. It is called the *Golden mean shift*. This language can be described by the regular expression $L = \{0\}^*\{1\}\{0,01\}^* \cup \{0\}^*$. Otherwise stated, it is generated from the finite languages $\{0\}$, $\{0,01\}$, and $\{1\}$ by applying union, concatenation, and star operations. Its complement in Σ^* is also regular and is described by the regular expression $\Sigma^*\{11\}\Sigma^*$. The language E from Example 1.5.3 is also regular; we have the following regular expression $\{0\}^*(\{1\}\{0\}^*\{1\}\{0\}^*)^*$ describing E.

Definition 1.5.7. Let $L \subseteq \Sigma^*$ be a language over the alphabet Σ. The *growth function* of L is the map

$$g_L : \mathbb{N} \to \mathbb{N},\ n \mapsto \mathrm{Card}(L \cap \Sigma^n)\,.$$

In particular, $g_L(n) \leq (\mathrm{Card}\,\Sigma)^n$ for all $n \geq 0$. Note that the complexity function of an infinite word $\mathbf{x}$ (see Definition 1.3.5) is exactly the growth function of the language $\mathrm{Fac}(\mathbf{x})$ of $\mathbf{x}$.

1.5.2 Formal Series

Let R be a semiring (see Definition 1.2.2). We can consider a map m from Σ^* to R. This map can be represented as a formal series

$$S = \sum_{w \in \Sigma^*} m(w)\, w\,.$$

This means that the coefficient (S, w) of the series S for the word w is given by $m(w)$. The sets of those formal series is denoted by $R\langle\langle \Sigma^* \rangle\rangle$ and has a semiring structure for the two operations defined as follows:

$$(S + T, w) = (S, w) + (T, w)$$

and

$$(ST, w) = \sum_{uv=w} (S, u)(T, v)\,.$$

In particular, a finite word w of length n can be factored in $n+1$ concatenation products. This means that the sum above is finite. When R is limited to the Boolean semiring $\mathbb{B}$, then $\mathbb{B}\langle\langle\Sigma^*\rangle\rangle$ is just the set of languages over Σ. As a prominent example, Mahler functions are studied in details in Chapter 2.

1.5.3 Codes

A subset $X \subset \Sigma^+$ is a *code* if every word in X^* has a unique factorization with factors in X, i.e.,

$$(x_1 \cdots x_m = y_1 \cdots y_n,\ x_1, \ldots, x_m, y_1, \ldots, y_n \in X) \Rightarrow (m = n \text{ and } x_i = y_i\ \forall i)\,.$$

As an example, the set $X = \{a, ab, ba\}$ is not a code because the word aba has two X-factorizations: $a(ba)$ and $(ab)a$. The language $\{a^i b \mid i \geq 0\}$ is clearly a code. For an introduction to codes, see Bruyère's chapter in [386].

Let X be a set of words where no word in X is a proper prefix of another word in X. Then X is said to be a *prefix code*. The terminology of code comes from the following proposition.

Proposition 1.5.8. *A subset $X \subset \Sigma^+$ is a code if and only if any morphism $f : \Gamma^* \to \Sigma^*$ induced by a one-to-one correspondence (i.e., bijection) from Γ to X is one to one (injective).*

The notion can be extended to deal with infinite words. A subset $X \subset \Sigma^+$ is an *ω-code* if every word in Σ^ω has at most one factorization with words in X. As an example, $X = \{a, ab, bb\}$ is a code but it is not an ω-code. The infinite word $abbb\cdots$ has two X-factorizations $(a, bb, bb, \ldots)$ and $(ab, bb, bb, \ldots)$.

1.5.4 Automata

As we shall briefly explain in this section, the regular languages are exactly the languages recognized by finite automata. We start with non-deterministic automata in Definition 1.5.9, then we present the deterministic ones in Definition 1.5.13. Finally, we introduce automata with output in Definition 1.5.17. The notions recalled here will be used in particular in Section 7.5 in connection with normality, and in Chapter 10 with the notion of Mealy automaton.

Definition 1.5.9. A *finite automaton* is a labeled graph given by a 5-tuple $\mathscr{A} = (Q, \Sigma, E, I, T)$ where Q is the (finite) *set of states*, $E \subseteq Q \times \Sigma^* \times Q$ is the finite set of *edges* defining the *transition relation*, $I \subseteq Q$ is the set of *initial states*, and T is the *set of terminal (or final) states*. A *path* in the automaton is a sequence

$$(q_0, u_0, q_1, u_1, \ldots, q_{k-1}, u_{k-1}, q_k)$$

such that, for all $i \in \{0, \ldots, k-1\}$, $(q_i, u_i, q_{i+1}) \in E$, $u_0 \cdots u_{k-1}$ is the *label* of the path. Such a path is *successful* if $q_0 \in I$ and $q_k \in T$. The language $L(\mathscr{A})$ *recognized* (or *accepted*) by $\mathscr{A}$ is the set of labels of all successful paths in $\mathscr{A}$.

Any finite automaton $\mathscr{A}$ gives a partition of Σ^* into $L(\mathscr{A})$ and $\Sigma^* \setminus L(\mathscr{A})$. When depicting an automaton, initial states are marked with an incoming arrow and terminal states are marked with an outgoing arrow. A transition like (q, u, r) is represented by a directed edge from q to r with label u, $q \xrightarrow{u} r$.

Example 1.5.10. In Figure 1.1 the automaton has two initial states p and r and three terminal states q, r, and s. For instance, the word ba is recognized by the automaton. There are two successful paths corresponding to the label ba: (p, b, q, a, s) and (p, b, p, a, s). For this latter path, we can write $p \xrightarrow{b} p \xrightarrow{a} s$. On the other hand, the word $baab$ is not recognized by the automaton.

Example 1.5.11. The automaton in Figure 1.2 recognizes exactly the language E of the words having an even number of 1 from Example 1.5.3.

Definition 1.5.12. Let $\mathscr{A} = (Q, \Sigma, E, I, T)$ be a finite automaton. A state $q \in Q$ is *accessible* (respectively *co-accessible*) if there exists a path from an initial state to q (respectively from q to some terminal state). If all states of $\mathscr{A}$ are both accessible and co-accessible, then $\mathscr{A}$ is said to be *trim*.

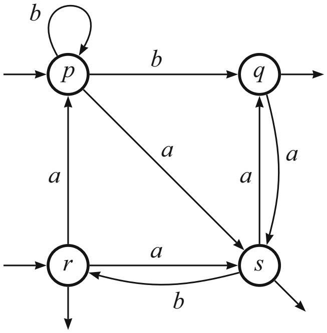

Fig. 1.1 A finite automaton.

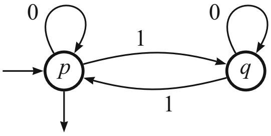

Fig. 1.2 An automaton recognizing words with an even number of 1.

Definition 1.5.13. A finite automaton $\mathscr{A} = (Q, \Sigma, E, I, T)$ is said to be *deterministic* (*DFA*) if it has only one initial state q_0, if E is a subset of $Q \times \Sigma \times Q$ and for each $(q, a) \in Q \times \Sigma$ there is at most one state $r \in Q$ such that $(q, a, r) \in E$. In that case, E defines a partial function $\delta_{\mathscr{A}} : Q \times \Sigma \to Q$ that is called the *transition function* of $\mathscr{A}$. The adjective *partial* means that the domain of $\delta_{\mathscr{A}}$ can be a strict subset of $Q \times \Sigma$. To express that the partial transition function is total, the DFA can be said to be *complete*. To get a total function, one can add to Q a new "sink state" s and, for all $(q, a) \in Q \times \Sigma$ such that $\delta_{\mathscr{A}}$ is not defined, set $\delta_{\mathscr{A}}(q, a) := s$. This operation does not alter the language recognized by $\mathscr{A}$. We can extend $\delta_{\mathscr{A}}$ to be defined on $Q \times \Sigma^*$ by $\delta_{\mathscr{A}}(q, \epsilon) = q$ and, for all $q \in Q$, $a \in \Sigma$, and $u \in \Sigma^*$, $\delta_{\mathscr{A}}(q, au) = \delta_{\mathscr{A}}(\delta_{\mathscr{A}}(q, a), u)$. Otherwise stated, the language recognized by $\mathscr{A}$ is $L(\mathscr{A}) = \{u \in \Sigma^* \mid \delta_{\mathscr{A}}(q_0, u) \in F\}$ where q_0 is the initial state of $\mathscr{A}$. If the automaton is deterministic, it is sometimes convenient to refer to the 5-tuple $\mathscr{A} = (Q, \Sigma, \delta_{\mathscr{A}}, I, T)$.

As explained by the following result, for languages of finite words, finite automata and deterministic finite automata recognize exactly the same languages. The following result is referred to as Rabin–Scott theorem [489].

Theorem 1.5.14. *If L is recognized by a finite automaton $\mathscr{A}$, there exists a DFA which can be effectively computed from $\mathscr{A}$ and recognizing the same language L.*

A proof and more details about classical results in automata theory can be found in textbooks like [300, 518] or [539]. For standard material in automata theory, we shall not refer again to these references below.

One important result is that the set of regular languages coincides with the set of languages recognized by finite automata. The following result is referred to as Kleene's theorem [349].

Theorem 1.5.15. *A language is regular if and only if it is recognized by a (deterministic) finite automaton.*

Observe that if L and M are two regular languages over Σ, then $L \cap M$, $L \cup M$, LM, and $L \setminus M$ are also regular languages. In particular, a language over Σ is regular if and only if its complement in Σ^* is regular.

Example 1.5.16. The regular language $L = \{0\}^*\{1\}\{0, 01\}^* \cup \{0\}^*$ introduced in Example 1.5.6 is recognized by the DFA depicted in Figure 1.3. Notice that the state s is a *sink*: a non-terminal state and all transitions remain in s.

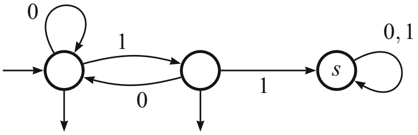

Fig. 1.3 A DFA accepting words without factor 11.

We introduce the notion of automaton with output (see also more generally Definition 7.5.1 for the notion of a transducer). It generalizes the classical DFA: if the output function takes at most two values, then it is a DFA. The extra output function will take care of the extra coding.

Definition 1.5.17. A *deterministic finite automaton with output* or DFAO for short is given by a 5-tuple $\mathscr{A} = (Q, q_0, A, \delta, \mu)$ where Q is a finite set of states, $q_0 \in Q$ is the initial state, $\delta : Q \times A \to Q$ is the transition function, and $\mu : Q \to B$ is the output map (where B is some finite set).

Finite automata accepting languages of infinite words are not presented here. Büchi automata (where an accepting run goes infinitely often through an accepting state) are introduced in Section 3.6.

1.6 Sequences and Machines

1.6.1 Automatic Sequences

We now consider how finite automata can be used to generate sequences with values in a finite alphabet, namely, we present the *automatic sequences*. As we shall soon see, they are particular morphic words and are deeply linked with the integer base-k numeration system. They were introduced by A. Cobham [156] under the name *uniform tag sequences*. Automatic sequences will appear in Chapters 2, 3, and 4. See in particular Section 2.2 for definitions, properties and examples, and connections with Mahler functions. We will recall that automatic sequences may be obtained as the image under a coding of the fixed point of a k-uniform morphism. Equivalently, for all $n \geq 0$, the nth symbol of such a sequence is the output of a deterministic finite automaton with output fed with the k-ary expansion of n.

Definition 1.6.1. Let $k \geq 2$. Consider an infinite word $\mathbf{w} = g(f^{\omega}(a))$ where $f : \Sigma^* \to \Sigma^*$ is a k-uniform morphism prolongable on a and $g : \Sigma^* \to \Gamma^*$ is a coding. We say that $\mathbf{w}$ is *k-automatic*.

Observe that $|f^n(a)| = k^n$ for all $n \geq 0$. We first consider the "internal sequence," i.e., the fixed point $\mathbf{x} = f^{\omega}(a) = x_0x_1x_2\cdots$. Let j such that $k \leq j < k^2$; then $j = kq + r$ with $1 \leq q < k$ and $0 \leq r < k$. The symbol x_j is the $(r+1)$st symbol occurring in $f(x_q)$. As depicted in Figure 1.4, this simply comes from one iteration of the k-uniform morphism.

We obtain the following result by induction on $m \geq 0$. Even though it is not surprising, it has an important consequence about how the word can be obtained.

Lemma 1.6.2. *Let j such that $k^m \leq j < k^{m+1}$, for some $m \geq 0$. Then $j = kq + r$ with $k^{m-1} \leq q < k^m$ and $0 \leq r < k$ and the symbol x_j is the $(r+1)$st symbol occurring in $f(x_q)$.*

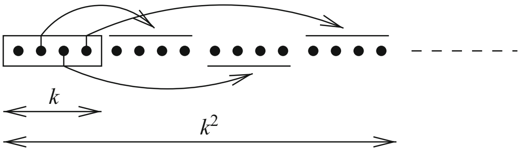

Fig. 1.4 Iterating a k-uniform morphism (with $k = 4$).

$$f : \begin{cases} \mathtt{a} \mapsto \mathtt{abc} \\ \mathtt{b} \mapsto \mathtt{cbc} \\ \mathtt{c} \mapsto \mathtt{bca} \end{cases}$$

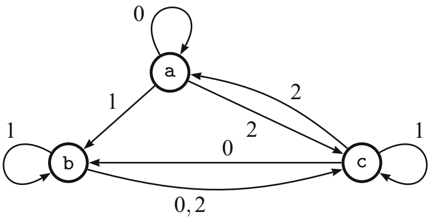

Fig. 1.5 A 3-uniform morphism and the associated automaton $\mathscr{A}_f$.

The quotient $\lfloor j/k \rfloor$ of the Euclidean division of j by k is denoted by j DIV k. So to speak, for any symbol x_j occurring in $\mathbf{x} = f^\omega(a)$, we can track its history: x_j has been produced by f from $x_{j \text{ DIV } k}$. The latter symbol appears itself in the image by f of $x_{(j \text{ DIV } k) \text{ DIV } k}$, and so on and so forth.

Note that if the base-k expansion of j is $\text{rep}_k(j) = c_i \cdots c_1 c_0$, then the base-$k$ expansion of j DIV k is $c_i \cdots c_1$. This simple observation permits one to easily track the past of a given symbol by considering the prefixes of $\text{rep}_k(j)$. Consider, for instance, the symbol $\mathbf{t}_{28}$ occurring in the Thue–Morse word:

$$\mathbf{t} = 0\underline{1}1\overline{0}100\underline{1}100101\overline{1}0100101100110\underline{1}001\cdots.$$

Since $\text{rep}_2(28) = 11100$, this symbol comes from $\mathbf{t}_{14}$ because $\text{rep}_2(14) = 1110$. Then $\mathbf{t}_{14}$ appears in the image of $\mathbf{t}_7$, itself appearing in the image of $\mathbf{t}_3$ and finally in the image of $\mathbf{t}_1$.

But Lemma 1.6.2 provides some extra knowledge. Let j such that $j = kq + r$ with $k^{m-1} \le q < k^m$ and $0 \le r < k$, for some $m \ge 0$. We have just explained how x_j comes from x_q. But the knowledge of x_q and r entirely determines x_j. It is thus time to explain where does the term of automatic sequence come from.

We can associate with a k-uniform morphism $f : \Sigma^* \to \Sigma^*$ and a letter $a \in \Sigma$, a DFA $\mathscr{A}_f = (\Sigma, a, [\![0, k-1]\!], \delta_f, \Sigma)$ where $\delta_f(b, i) = w_{b,i}$ if $f(b) = w_{b,0} \cdots w_{b,k-1}$. Note that the alphabet Σ is the set of states of this automaton.

Example 1.6.3. Consider the morphism f and the associated automaton depicted in Figure 1.5.

The next propositions explain the terminology of automatic sequences.

Proposition 1.6.4. *Let* $\mathbf{x} = f^{\omega}(a) = x_0x_1\cdots$ *with* f *a* k*-uniform morphism. With the above notation, for all* $j \geq 0$,

$$x_j = \delta_f(a, \mathrm{rep}_k(j)) .$$

Proof. This is a direct consequence of Lemma 1.6.2. □

The converse also holds.

Proposition 1.6.5. *Let* $(\Sigma, a, [\![0, k-1]\!], \delta, \Sigma)$ *be a DFA such that* $\delta(a, 0) = a$*. Then the word* $\mathbf{x} = x_0x_1x_2\cdots$ *defined by* $x_j = \delta(a, \mathrm{rep}_k(j))$*, for all* $j \geq 0$*, is the fixed point of a* k*-uniform morphism* f *prolongable on* a *where* $f(b) = \delta(b, 0)\cdots\delta(b, k-1)$ *for all* $b \in \Sigma$.

Proof. This is again a direct consequence of Lemma 1.6.2. □

The reader will object that we have not taken into account that an extra coding can be applied to $\mathbf{x} = f(\mathbf{x})$. This does not require many changes. We simply have to make use of automata with output as stated below in Cobham's theorem on automatic sequences [156].

Theorem 1.6.6. *Let* $\mathbf{w} = w_0w_1w_2\cdots$ *be an infinite word over an alphabet* Γ*. It is of the form* $g(f^{\omega}(a))$ *where* $f : \Sigma^* \to \Sigma^*$ *is a* k*-uniform morphism prolongable on* $a \in \Sigma$ *and* $g : \Sigma^* \to \Gamma^*$ *is a coding if and only if there exists a DFAO*

$$(\Sigma, a, [\![0, k-1]\!], \delta, \mu : \Sigma \to \Gamma)$$

such that $\delta(a, 0) = a$ *and, for all* $j \geq 0$, $w_j = \mu(\delta(a, \mathrm{rep}_k(j)))$.

Proof. Proceed as above and the coding g coincides with the output function μ. □

Example 1.6.7. From the morphism t given in (1.3) generating the Thue–Morse word, we derive the automaton depicted in Figure 1.2. Again considering 28, which is written 11100 in base 2, if we start from the initial state p and we read consecutively the symbols in $\mathrm{rep}_2(28)$ from left to right, then we follow some path in the automaton, and the state q finally reached gives the symbol $\mathbf{t}_{28}$. The output function maps p to 0 and q to 1.

Example 1.6.8. Let us consider a more intricate example where a coding, and thus an output function, is used. The morphism f and the coding g are given in Figure 1.6. The corresponding automaton is represented on the right of the same figure. We have

$$f^{\omega}(\mathtt{a}) = \mathtt{acabaccaacababacacabaccaaccaacab}\cdots$$

and

$$g(f^{\omega}(\mathtt{a})) = 00010000000101000001000000000001\cdots .$$

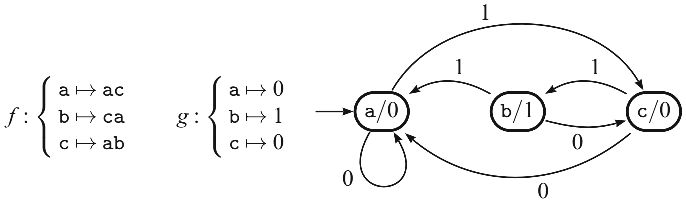

Fig. 1.6 A 2-uniform morphism, a coding and the corresponding DFAO.

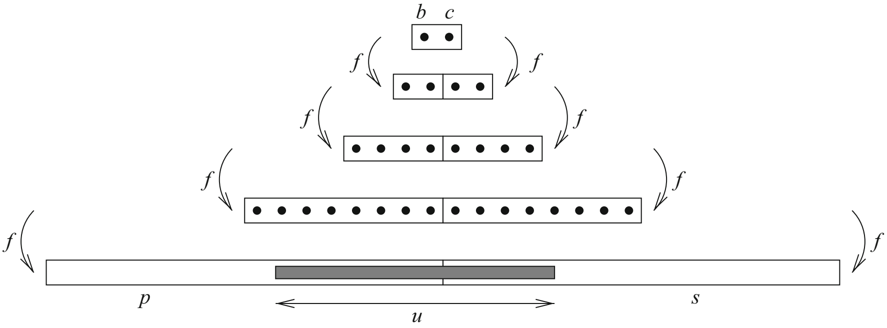

Fig. 1.7 Iterating a 2-uniform morphism.

Again, the jth symbol in $g(f^\omega(\mathtt{a}))$ can be readily obtained from $\mathrm{rep}_2(j)$ fed to the DFAO represented in Figure 1.6 where the states contain the information about the value of the output function.

Now we turn to the factors occurring in an automatic sequence $\mathbf{w} = g(\mathbf{x})$, where $\mathbf{x}$ is a fixed point of the k-uniform morphism $f : \Sigma^* \to \Sigma^*$. Let u be a factor of length n occurring in $\mathbf{x}$. There exists i such that $k^{i-1} \leq n < k^i$. Note that $|f^i(b)| = k^i$ for all $b \in \Sigma$. We consider the factorization of $\mathbf{x}$ into consecutive blocks of length k^i of the form $f^i(b)$. Therefore, the factor u either occurs inside some $f^i(b)$ or it overlaps two images, i.e., u occurs in $f^i(bc)$ for some letters $b, c \in \Sigma$. Actually, there exist two letters b and c such that $f^i(bc) = pus$ with $|p| < k^i$. This last condition tells us that u starts inside $f^i(b)$. Such a simple observation, where we look backwards at the images of the morphism, as suggested by Figure 1.7, is sometimes called a *desubstitution*. It provides us with an upper bound on the number of factors of length n that may occur in $\mathbf{x}$: the number of pairs of letters (b, c) is $(\mathrm{Card}\,\Sigma)^2$ and u should start in one of the k^i symbols of $f^i(b)$. Therefore, the number of factors of length n in $\mathbf{x}$ is at most

$$(\mathrm{Card}\,\Sigma)^2\, k^i \leq (\mathrm{Card}\,\Sigma)^2\, k\, n\,.$$

We can even replace $(\mathrm{Card}\,\Sigma)^2$ with $p_{\mathbf{x}}(2)$ because only the factors bc occurring in $\mathbf{x}$ give factors of the form $f^i(b)f^i(c)$ occurring in $\mathbf{x} = f^i(\mathbf{x})$. Since applying a coding

g cannot increase the number of factors, we get

$$\mathrm{Card}(\mathrm{Fac}(\mathbf{x}) \cap \Sigma^n) \geq \mathrm{Card}\{g(u) \mid u \in \mathrm{Fac}(\mathbf{x}) \cap \Sigma^n\}\,,$$

and so we have obtained the following result.

Theorem 1.6.9. *Let* $\mathbf{w}$ *be a k-automatic sequence. Then* $p_{\mathbf{w}}(n)$ *is in* $\mathcal{O}(n)$*.*

A proof of the following result can be found in [14, Section 8.4].

Theorem 1.6.10. *If the frequency of a letter in a morphic sequence exists, then it is an algebraic number. If the frequency of a letter in an automatic sequence exists, then it is a rational number.*

To conclude this section, we present another characterization of k-automatic sequences. This is not the last one; in Chapter 3, Section 3.3, a logical characterization of k-automatic sequences will be discussed, whereas Chapter 4 will provide an algebraic characterization in terms of polynomial identities (see Corollary 4.5.3).

Definition 1.6.11. Let $k \geq 2$ be an integer. Given a sequence $s = (s(n))_{n\geq 0}$, we define a particular set of subsequences called the *k-kernel* of s

$$\mathrm{Ker}_k(s) := \left\{(s(k^\ell n + r))_{n\geq 0} \mid \ell \geq 0,\ 0 \leq r < k^\ell\right\}\,.$$

An equivalent definition of the k-kernel is to introduce k operators of *k-decimation* acting on the set of sequences and defined, for $r \in \{0, \ldots, k-1\}$, by

$$\rho_{k,r}((s(n))_{n\geq 0}) = (s(kn + r))_{n\geq 0}\,.$$

Thus $\mathrm{Ker}_k(s)$ is the set of sequences of the form

$$\rho_{k,r_1} \circ \cdots \circ \rho_{k,r_m}((s(n))_{n\geq 0}) \tag{1.4}$$

for all $m \geq 0$ and $r_1, \ldots, r_m \in \{0, \ldots, k-1\}$. These decimation operators are close to the Cartier operators discussed in Chapter 2. The following result appeared in Eilenberg's book [211]. Note that if a sequence t belongs to $\mathrm{Ker}_k(s)$, then $\rho_{k,r}(t)$ also belongs to $\mathrm{Ker}_k(s)$.

Theorem 1.6.12. *A sequence is k-automatic if and only if its k-kernel is finite.*

Example 1.6.13. The 2-kernel of the Thue–Morse sequence contains exactly two sequences (the sequence itself and its "complement"). Indeed, let $s_2(n)$ be the sum of digits of the binary expansion of n, we have

$$s_2(2n) = s_2(n), \quad s_2(2n+1) = s_2(n) + 1\,. \tag{1.5}$$

1.6.2 Regular Sequences

We have seen that k-automatic sequences may be defined through the finiteness of their k-kernel (Theorem 1.6.12). This characterization is used to extend the notion to sequences taking infinitely many values. Allouche and Shallit considered sequences taking values in a ring R containing a commutative Noetherian ring R' (i.e., every ideal of R' is finitely generated). Examples of such rings R' are given by all finite rings, all principal ideal domains, and in particular $\mathbb{Z}$, the ring of polynomials with coefficients in a field, or all fields. We may consider linear combinations with coefficients in R' (R'-linear combinations) of sequences in $R^{\mathbb{N}}$. Endowed with pointwise addition and multiplication by an element in R', the set $R^{\mathbb{N}}$ has a R'-module structure: if $r = (r(n))_{n\geq 0}$ and $s = (s(n))_{n\geq 0}$ belong to $R^{\mathbb{N}}$ and α belongs to R', then, for all $n \in \mathbb{N}$,

$$(r + s)(n) = r(n) + s(n)$$

and

$$(\alpha \cdot r)(n) = \alpha \cdot r(n).$$

In this short section, we mainly consider sequences in $\mathbb{Z}^{\mathbb{N}}$, i.e., $R = R' = \mathbb{Z}$. We will encounter regular sequences in Chapters 2, 3, and 4 of this book. To have stand-alone chapters, these notions will also be repeated there. In Chapter 3 (see in particular Section 3.4.1), k-regularity will be extended to sequences taking values in a semiring.

Regular sequences appeared in [16]. Many examples are given in [15]. See also [14, Chapter 16] and the updated version of Berstel and Reutenauer's book [77] where a chapter is devoted to regular sequences and linked with rational series.

Let M be a R-module and a subset $X \subset M$. The *submodule generated* by X is the intersection of all submodules of M containing X. It is denoted by $\langle X\rangle$. It is the set of all finite R-linear combinations of elements in X. A module is *finitely generated* (over R) when it is generated by a finite set (i.e., it is the R-span of a finite set). One also says that the module is of *finite type* or even *finite over* R. Note that the finite set of generators is not necessarily a basis.

Definition 1.6.14. Let $k \geq 2$ be an integer. A sequence $s = (s(n))_{n\geq 0}$ taking integer values is *k-regular* if the $\mathbb{Z}$-module generated by its k-kernel $\langle \mathrm{Ker}_k(s)\rangle$ is finitely generated, i.e., there exists a finite number of sequences in $\mathbb{Z}^{\mathbb{N}}$

$$t_1 = (t_1(n))_{n\geq 0}, \ldots, t_\ell = (t_\ell(n))_{n\geq 0}$$

such that

$$\langle \mathrm{Ker}_k(s)\rangle = \langle t_1, \ldots, t_\ell\rangle.$$

In particular, every sequence in $\mathrm{Ker}_k(s)$ is a $\mathbb{Z}$-linear combination of the t_js. For all $i \geq 0$ and for all $r \in \{0, \ldots, k^i - 1\}$, there exist integers $c_{i,1}, \ldots, c_{i,\ell}$ such that

$$\forall n \geq 0, \quad s(k^i n + r) = \sum_{j=1}^{\ell} c_{i,j}\, t_j(n).$$

One can consider another point of view. A sequence is said to be k-regular if its orbit under the action of the operators of k-decimation remains in a finite dimensional vector space. Indeed, $\mathbb{Z}$ is included in fields such as $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$. Thus the sequences can be seen as elements of $\mathbb{Q}^{\mathbb{N}}$ which is a $\mathbb{Q}$-vector space.

Remark 1.6.15. The original definition in [16] was formulated differently. Let R be a ring containing a commutative Noetherian ring R'. A sequence $s = (s(n))_{n\geq 0}$ in $R^{\mathbb{N}}$ is *(R',k)-regular* if there exists a finite number of sequences in $R^{\mathbb{N}}$

$$t_1 = (t_1(n))_{n\geq 0}, \dots, t_\ell = (t_\ell(n))_{n\geq 0}$$

such that every sequence in $\mathrm{Ker}_k(s)$ is an R'-linear combination of $t_1, \dots, t_\ell$. Thus the definition means that $\langle \mathrm{Ker}_k(s)\rangle \subseteq \langle t_1, \dots, t_\ell\rangle$. Otherwise stated, $\langle \mathrm{Ker}_k(s)\rangle$ is a submodule of a finitely generated R'-module (in general, this does not imply that the submodule itself is finitely generated). Since R' is assumed to be Noetherian, one can show that every submodule of a finitely generated R'-module is finitely generated[2], and thus $\langle \mathrm{Ker}_k(s)\rangle$ is finitely generated. This was the point of view adopted in Definition 1.6.14. In particular, if the setting does not assume that R' is Noetherian (in particular, if R or R' is a semiring), then Definition 1.6.14 would be stronger than simply requiring $\langle \mathrm{Ker}_k(s)\rangle \subseteq \langle t_1, \dots, t_\ell\rangle$.

Example 1.6.16. The base-2 sum-of-digits function s_2 gives the sequence

$$(s_2(n))_{n\geq 0} = 0, 1, 1, 2, 1, 2, 2, 3, 1, 2, 2, 3, 2, 3, 3, 4, 1, 2, 2, 3, 2, 3, 3, 4, 2, 3, 3, 4, \dots .$$

(Notice that we can interchange the words function and sequence and also speak of k-regular functions when defined over $\mathbb{N}$.) Clearly this sequence is unbounded: $s_2(2^n - 1) = n$ for all n. Nevertheless, in view of (1.5), the $\mathbb{Z}$-module generated by its 2-kernel is generated by the sequence $(s_2(n))_{n\geq 0}$ itself and the constant sequence $(1)_{n\geq 0}$.

Obviously, every k-automatic sequence is k-regular.

Proposition 1.6.17. *Let s be a sequence taking finitely many different values, i.e., there exists a finite alphabet Σ such that $s \in \Sigma^\omega$. Let $k \geq 2$. The sequence is k-automatic if and only if it is k-regular.*

There is an intermediate class of sequences between k-automatic and k-regular sequences [130].

[2] An R'-module M is *Noetherian* if every submodule of M is finitely generated. Let R' be a Noetherian ring. An R'-module M is Noetherian if and only if it is finitely generated.

Definition 1.6.18. Let $k \geq 2$ be an integer. The map rep_k is extended to $\mathbb{N} \times \mathbb{N}$ as follows. For all $m, n \in \mathbb{N}$,

$$\text{rep}_k(m,n) = \left(0^{M-|\text{rep}_k(m)|}\text{rep}_k(m), 0^{M-|\text{rep}_k(n)|}\text{rep}_k(n)\right)$$

where $M = \max\{|\text{rep}_k(m)|, |\text{rep}_k(n)|\}$. The idea is that the shortest word is padded with leading zeroes to get two words of the same length.

A sequence $(s(n))_{n\geq 0}$ of integers is said to be *k-synchronized* if the language $\{\text{rep}_k(n, s(n)) \mid n \in \mathbb{N}\}$ is accepted by some finite automaton reading pairs of digits.

As an example, the complexity function $(p_{\mathbf{x}}(n))_{n\geq 0}$ of a k-automatic sequence $\mathbf{x}$ is k-synchronized [522]; we refer to Proposition 3.4.16. More results of this form are provided in Section 3.4. For results on the growth of regular sequences, see Section 2.3.

Proposition 1.6.19. *Let s be a sequence taking finitely many different values, i.e., there exists a finite alphabet Σ such that $s \in \Sigma^\omega$. Let $k \geq 2$. The sequence is k-automatic if and only if it is k-synchronized.*

Similarly to recognizable formal series, with every k-regular sequence $(s(n))_{n\geq 0} \in \mathbb{Z}^{\mathbb{N}}$ is associated *linear representation* (λ, μ, ν). There exist a positive integer r, a row vector $\lambda \in \mathbb{Z}^{1\times r}$ and a column vector $\nu \in \mathbb{Z}^{r\times 1}$, a matrix-valued morphism $\mu : \{0, \ldots, k-1\} \to \mathbb{Z}^{r\times r}$ such that

$$s(n) = \lambda\mu(c_0 \cdots c_\ell)\nu$$

for all $c_\ell, \ldots, c_0 \in \{0, \ldots, k-1\}^*$ such that $\text{val}_k(c_\ell \cdots c_0) = \sum_{i=0}^{\ell} c_i k^i = n$. The converse also holds, if there exists a linear representation associated with the canonical k-ary expansion of integers (one has to take into account the technicality of representations with leading zeros), then the sequence is k-regular. See, for instance, [14, Theorem 16.2.3]. As a corollary, the nth term of a k-regular sequence can be computed with $\lfloor \log_k(n) \rfloor$ matrix multiplications.

Proof. Let $s = (s(n))_{n\geq 0} \in \mathbb{Z}^{\mathbb{N}}$ be a k-regular sequence. By definition, there exists a finite number of sequences $t_1, \ldots, t_\ell$ such that $\langle \text{Ker}_k(s) \rangle = \langle t_1, \ldots, t_\ell \rangle$. In particular, each t_j is a $\mathbb{Z}$-linear combination of elements in the k-kernel of s. We have finitely many t_js, so $t_1, \ldots, t_\ell$ are linear combinations of finitely many elements in $\text{Ker}_k(s)$. Thus we can assume that $\langle \text{Ker}_k(s) \rangle$ is generated by finitely many elements from $\text{Ker}_k(s)$ itself. Without loss of generality, we will now assume that $t_1, \ldots, t_\ell$ belong to $\text{Ker}_k(s)$.

From (1.4), for all $r \in \{0, \ldots, k-1\}$ and all $i \in \{1, \ldots, \ell\}$, $\rho_{k,r}(t_i)$ is a sequence in $\text{Ker}_k(s)$, and thus, there exist coefficients $(A_r)_{1,i}, \ldots, (A_r)_{\ell,i}$ such that

$$\rho_{k,r}(t_i) = \sum_{j=1}^{\ell} (A_r)_{j,i}\, t_j\,.$$

Notice that A_r is an $\ell \times \ell$ matrix. Roughly, if we were in a vector space setting, this means that the matrices A_r represent the linear operators $\rho_{k,r}$ in the basis $t_1, \dots, t_\ell$. Let $p \geq 0$ be an integer. Notice that if $\mathrm{rep}_k(p) = r_m \cdots r_0$, then $s(p)$ is the first term, i.e., corresponding to the index 0, of the sequence

$$(s(b^{m+1}n + p))_{n \geq 0} = \rho_{k,r_0} \circ \cdots \circ \rho_{k,r_m} \left((s(n))_{n \geq 0}\right) .$$

We will use the fact that $\rho_{k,r}$ is linear, i.e., if α, β are coefficients and v, w are two sequences, then $\rho_{k,r}(\alpha v + \beta w) = \alpha \rho_{k,r}(v) + \beta \rho_{k,r}(w)$. It is easy to see that

$$\rho_{k,r_0} \circ \cdots \circ \rho_{k,r_m} (t_i) = \sum_{j=1}^{\ell} (A_{r_0} \cdot \cdots \cdot A_{r_m})_{j,i}\, t_j .$$

If we have the following decomposition of s (in a vector space setting, we would have a unique decomposition of s in the basis $t_1, \dots, t_\ell$)

$$s = \sum_{i=1}^{\ell} \sigma_i\, t_i$$

then, by linearity,

$$(s(b^{m+1}n + p))_{n \geq 0} = \sum_{i=1}^{\ell} \sigma_i \sum_{j=1}^{\ell} (A_{r_0} \cdot \cdots \cdot A_{r_m})_{j,i}\, (t_j(n))_{n \geq 0} = \sum_{j=1}^{\ell} \tau_j\, (t_j(n))_{n \geq 0}$$

where

$$\begin{pmatrix} \tau_1 \\ \vdots \\ \tau_\ell \end{pmatrix} = A_{r_0} \cdot \cdots \cdot A_{r_m} \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_\ell \end{pmatrix} .$$

Consequently, $s(p)$ is obtained as

$$s(p) = \sum_{i=1}^{\ell} \tau_i\, t_i(0) = \big(t_1(0) \cdots t_\ell(0)\big)\, A_{r_0} \cdot \cdots \cdot A_{r_m} \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_\ell \end{pmatrix} .$$

□

For a reader familiar with rational series, the previous result can be reformulated as follows. A sequence $s(n)$ is k-regular if and only if the formal series

$$\sum_{w\in\{0,\dots,k-1\}^*} s(\mathrm{val}_k(w))\, w$$

is recognizable (with the terminology of [77]; see Definition 3.4.1).

Example 1.6.20. For the sum-of-digits function given in Example 1.6.16, the sequence $s_2 = (s_2(n))_{n\geq 0}$ has a (base-2) linear representation given by

$$\lambda = \begin{pmatrix} 0 & 1 \end{pmatrix},\ \mu(i) = \begin{pmatrix} 1 & 0 \\ i & 1 \end{pmatrix},\ \nu = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

We let $\mathbf{1}$ denote the constant sequence. It does not belong to the 2-kernel of s_2, but it belongs to the $\mathbb{Z}$-module generated by it because it is equal to $\rho_{2,1}(s_2) - s_2$. Nevertheless, it is enough to see that $\rho_{2,0}(\mathbf{1}) = \rho_{2,1}(\mathbf{1}) = \mathbf{1}$ and take s_2 and $\mathbf{1}$ as generators to proceed as in the proof above. From the following relations we derive the two columns of matrix $\mu(0)$

$$\rho_{2,0}(s_2) = 1\cdot s_2 + 0\cdot \mathbf{1},\quad \rho_{2,0}(\mathbf{1}) = 0\cdot s_2 + 1\cdot \mathbf{1}$$

and for $\mu(1)$

$$\rho_{2,1}(s_2) = 1\cdot s_2 + 1\cdot \mathbf{1},\quad \rho_{2,1}(\mathbf{1}) = 0\cdot s_2 + 1\cdot \mathbf{1}\,.$$

The vector λ is given by $s_2(0) = 0$ ans $\mathbf{1}(0) = 1$. The vector ν is obtained from $s_2 = 1\cdot s_2 + 0\cdot \mathbf{1}$. To compute $s_2(19)$, observe that $\mathrm{rep}_2(19) = 10011$. Thus we compute

$$\begin{pmatrix} 0 & 1 \end{pmatrix} \mu(1)\mu(1)\mu(0)\mu(0)\mu(1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 3\,.$$

Example 1.6.21. A less trivial example is considered in [201] by counting the number of odd numbers in the first n rows of the Pascal triangle. This sequence has a (base-2) linear representation given by

$$\lambda = \begin{pmatrix} 0 & 1 \end{pmatrix},\ \mu(0) = \begin{pmatrix} 3 & 6 \\ 0 & 1 \end{pmatrix},\ \mu(1) = \begin{pmatrix} 0 & -6 \\ 1 & 5 \end{pmatrix},\ \nu = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Remark 1.6.22. In [15, Section 6], a practical procedure to guess relations a possibly k-regular sequence will satisfy is described. Consider a sequence $(s(n))_{n\geq 0}$. The idea is to construct a matrix in which the rows represent truncated versions of elements of the k-kernel of $(s(n))_{n\geq 0}$, together with row reduction. Start with a matrix having a single row, say, corresponding to the first m elements of the sequence. Then repeatedly add subsequences of the form $(s(k^\ell n+r))_{n\geq 0}$ not linearly dependent of the previous stored sequences. From this, you have candidate relations that remain to be proven.

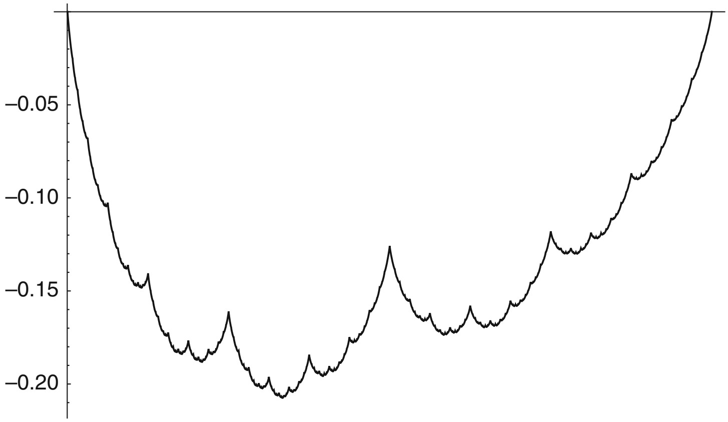

Fig. 1.8 The periodic function $\mathscr{G}$ on [0, 1].

Considering again the sum-of-digit function, Delange [191] showed that the summatory function of s_2 exhibits a particular behavior (also see [14, Thm. 3.5.4]).

$$\frac{1}{N}\sum_{j=0}^{N-1} s_2(j) = \frac{1}{2}\log_2 N + \mathscr{G}(\log_2 N) \tag{1.6}$$

where $\mathscr{G}$ is a continuous nowhere differentiable periodic function of period 1 (Figure 1.8).

General results do exist for summatory function of k-regular sequences. The result below can be found in [14, Thm. 16.4.1].

Theorem 1.6.23. *Let* $\mathsf{a} = (a(n))_{n\geq 0}$ *and* $\mathsf{b} = (b(n))_{n\geq 0}$ *be k-regular sequences. Then* $\mathsf{c} = \mathsf{a} \star \mathsf{b}$, *where, for all* $n \geq 0$, $c(n) = \sum_{i=0}^{n} a_i\, b_{n-i}$, *is k-regular.*

Corollary 1.6.24. *Let* $\mathsf{a} = (a(n))_{n\geq 0}$ *be a k-regular sequence. The sequence of partial sums*

$$\left(\sum_{i=0}^{n} a_i\right)_{n\geq 0}$$

is k-regular.

Proof. One simply takes for b the constant sequence $(1)_{n\geq 0}$ in Theorem 1.6.23. □

A linear representation of the summatory sequence can easily be deduced from the linear representation of the sequence itself, see [201, Lemma 1] or Proposition 2.2.11 in Chapter 2. Let us state the following result obtained by Dumas [201, 202] (see also Theorem 2.3.13).

Theorem 1.6.25. *Let $k \geq 2$ be an integer. The summatory function of a k-regular sequence $(u(n))_{n\geq 0}$ with a linear representation given by the matrices $\Gamma_0, \dots, \Gamma_{k-1}$ admits an asymptotic expansion which is a sum of terms of the form*

$$N^{\log_k \rho} \binom{\log_k N}{m} e^{i\theta \log_k N} \varphi(\log_k N)$$

for the eigenvalues $\rho e^{i\theta}$ of $\Gamma := \Gamma_0 + \cdots + \Gamma_{k-1}$ whose modulus ρ is larger than the joint spectral radius of $\Gamma_0, \dots, \Gamma_{k-1}$ and where m is an integer bounded by the maximal size of a Jordan block associated with $\rho e^{i\theta}$ and φ is a periodic function of period 1. For this asymptotic expansion, there is an error term in $\mathcal{O}(N^{\log_k r})$ for every r larger than the joint spectral radius of the matrices $\Gamma_0, \dots, \Gamma_{k-1}$.

Definition about the joint spectral radius will be given in Chapter 2; see also Chapter 11.8.1. Similar results are also discussed by Drmota and Grabner in [78, Theorem 9.2.15]. Let us also mention another result (see [14, Theorem 3.5.1]) with stronger assumptions but avoiding error terms. In this result, if v belongs to $\mathbb{C}^d$, then the notation $||v||$ stands for the Euclidean norm of v defined by $\left(\sum_{i=1}^d |v_i|^2\right)^{\frac{1}{2}}$. Moreover, if M is a square matrix of dimension d with entries in $\mathbb{C}$, then by $||M||$ we mean the L^2 norm, which is the matrix norm associated with the usual Euclidean norm on $\mathbb{C}^d$ by the formula $||M|| = \sup_{||x||=1} ||Mx||$.

Theorem 1.6.26. *Let $k \geq 2$ be an integer. Suppose there exist an integer $d \geq 1$, a sequence of vectors $(V_n)_{n\geq 0}$, $V_n \in \mathbb{C}^d$, defined by*

$$V_n = \begin{pmatrix} V_n^{(1)} \\ V_n^{(2)} \\ \vdots \\ V_n^{(d)} \end{pmatrix},$$

and k square matrices $\Gamma_0, \Gamma_0, \dots, \Gamma_{k-1}$ of dimension d such that

1. *$V_{kn+r} = \Gamma_r V_n$ for all $n \geq 0$ and all r, $0 \leq r < k$.*
2. *$||V_n|| = O(\log n)$.*
3. *There exist a $d \times d$ matrix Λ and a constant $c > 0$ such that either $||\Lambda|| < c$ or Λ is nilpotent, such that $\Gamma := \Gamma_0 + \Gamma_1 + \cdots + \Gamma_{k-1} = cI + \Lambda$.*

The matrix Γ being clearly invertible, if $||\Gamma^{-1}|| < 1$, then there exists a continuous function $G : \mathbb{R} \to \mathbb{C}^d$ of period 1 such that

$$\sum_{0\leq n<N} V_n = N^{\log_k c}(I + c^{-1}\Lambda)^{\log_k N} G(\log_k N).$$

1.7 Dynamical Systems

There are two main types of dynamical systems, namely, topological ones and measure-theoretic ones. Dynamical systems will be considered in particular in Chapters 8, 9, and 11.

1.7.1 Topological Dynamical Systems

Definition 1.7.1. A *topological dynamical system* (X, T) is defined as a compact metric space X together with a continuous map T defined onto the set X.

We are interested in iterating the map T, and we look at the *orbits* $\mathscr{O}(x)$ of $x \in X$ defined as

$$\mathscr{O}(x) = \{T^n(x) : n \in \mathbb{N}\}.$$

under the action T. The *trajectory* of $x \in X$ is the sequence $(T^n(x))_{n \geq 0}$.

A topological dynamical system (X, T) is *minimal* if, for all $\mathbf{x}$ in X, the orbit of $\mathbf{x}$, i.e., the set $\{T^n\mathbf{x} \mid n \in \mathbb{N}\}$, is dense in X. Let us note that if (X, S) is a subshift, and if X is furthermore assumed to be minimal, then X is periodic if and only if X is finite.

Two dynamical systems (X_1, T_1) and (X_2, T_2) are said to be *topologically conjugate* (or *topologically isomorphic*) if there exists an homeomorphism f from X_1 onto X_2 which conjugates T_1 and T_2, that is:

$$f \circ T_1 = T_2 \circ f.$$

If f is only onto, then (X_1, T_1) is said to factor onto (X_2, T_2), (X_2, T_2) is a factor of (X_1, T_1), and f is called a *factor map*.

1.7.2 Measure-Theoretic Dynamical Systems

We have considered here the notion of dynamical system, that is, a map acting on a given set, in a topological context. This notion can be extended to measurable spaces; we thus get measure-theoretic dynamical systems. For more details, one can refer, for instance, to [579]. See also Section 11.11.3.

Definition 1.7.2. A *measure-theoretic dynamical system* is defined as a system $(X, \mathscr{B}, \mu, T)$, where $\mathscr{B}$ is a σ-algebra, μ a probability measure defined on $\mathscr{B}$, and $T : X \to X$ is a measurable map which preserves the measure μ, i.e., for all $B \in \mathscr{B}$,

$\mu(T^{-1}(B)) = \mu(B)$. Such a measure is said to be *T-invariant* and the map T is said to preserve the measure μ.

The transformation T (or the system $(X, \mathscr{B}, \mu, T)$) is *ergodic* if for every $B \in \mathscr{B}$ such that $T^{-1}(B) = B$, then B has either zero measure or full measure.

Let (X, T) be a topological dynamical system. A topological system (X, T) always has an invariant probability measure. The case where there exists only one T-invariant measure is of particular interest. A topological dynamical system (X, T) is said to be *uniquely ergodic* if there exists one and only one T-invariant Borel probability measure over X. In particular, a uniquely ergodic topological dynamical system yields an ergodic measure-theoretic dynamical system.

A measure-theoretic ergodic dynamical system satisfies the *Birkhoff ergodic theorem*, also called *individual ergodic theorem*. Let us recall that the abbreviation a.e. stands for "almost everywhere": a property holds almost everywhere if the set of elements for which the property does not hold is contained in a set of zero measure.

Theorem 1.7.3. *Let $(X, \mathscr{B}, \mu, T)$ be a measure-theoretic dynamical system. Let $f \in L^1(X, \mathbb{R})$. Then the sequence $(\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k)_{n\geq 0}$ converges a.e. to a function $f^* \in L^1(X, \mathbb{R})$. One has $f^* \circ T = f^*$ a.e. and $\int_X f^*\, d\mu = \int_X f\, d\mu$. Furthermore, if T is ergodic, since f^* is a.e. constant, one has:*

$$\forall f \in L^1(X, \mathbb{R}) \ , \quad \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k \xrightarrow[n\to\infty]{\mu-a.e.} \int_X f\, d\mu \, .$$

Note that the notions of conjugacy and factor between two topological dynamical systems extend in a natural way to the measure-theoretic context.

1.7.3 Symbolic Dynamics

Let us introduce some basic notions in symbolic dynamics. For expository books on the subject, see [167, 348, 381, 475] and [488]. For references on ergodic theory, also see, e.g., [579]. These notions will be central in particular in Chapters 8 and 9.

Let S denote the following map defined on Σ^ω, called the *one-sided shift*:

$$S((x_n)_{n\geq 0}) = (x_{n+1})_{n\geq 0} \, .$$

In particular, if $x = x_0x_1x_2\cdots$ is an infinite word over Σ, then for all $n \geq 0$, its suffix $x_nx_{n+1}\cdots$ is simply $S^n(x)$. The map S is uniformly continuous, onto but not one to one on Σ^ω. This notion extends in a natural way to $\Sigma^{\mathbb{Z}}$. In this latter case, the shift S is one to one. We thus get *symbolic dynamical systems*. Here symbolic refers to the fact that they are defined on words.

The definitions given below correspond to the *one-sided shift*, but they extend in a natural way to the *two-sided shift*.

Definition 1.7.4. Let x be an infinite word over the alphabet Σ. The *symbolic dynamical system* associated with $\mathbf{x}$ is then defined as the shift orbit closure $(\overline{\mathscr{O}(\mathbf{x})}, S)$, where $\overline{\mathscr{O}(\mathbf{x})} \subseteq \Sigma^\omega$ is the closure of the orbit $\mathscr{O}(\mathbf{x}) = \{S^n\mathbf{x} \mid n \in \mathbb{N}\}$ of x.

In the case of bi-infinite words, we similarly define $\mathscr{O}(\mathbf{x}) = \{S^n\mathbf{x} \mid n \in \mathbb{Z}\}$ where the (two-sided) shift map is defined on $\Sigma^{\mathbb{Z}}$. The set $X_{\mathbf{x}} := \overline{\mathscr{O}(\mathbf{x})}$ is a closed subset of the compact set Σ^ω; hence it is a compact space and S is a continuous map acting on it. One checks that, for every infinite word $\mathbf{y} \in \Sigma^\omega$, the word $\mathbf{y}$ belongs to $X_{\mathbf{x}}$ if and only if $L(\mathbf{y}) \subseteq L(\mathbf{x})$. For a proof, see [488] or Chapter 1 of [487]. Note that $\overline{\mathscr{O}(\mathbf{x})}$ is finite if and only if $\mathbf{x}$ is eventually periodic. Moreover, if $\mathbf{x}$ is an infinite word, $(X_{\mathbf{x}}, S)$ is minimal if and only if $\mathbf{x}$ is uniformly recurrent. Indeed, w is a factor of $\mathbf{x}$, we write

$$\overline{\mathscr{O}(\mathbf{x})} = \bigcup_{n\in\mathbb{N}} S^{-n}[w],$$

and we conclude by a compactness argument.

Generic examples of symbolic dynamical systems are provided by subshifts (also called shifts for short). Let Y be a closed subset of Σ^ω that is stable under the action of the shift S. The system (Y, S) is called a *subshift*. The *full shift* is defined as (Σ^ω, S). If Y is a subshift, there exists a set $\mathscr{F} \subset \Sigma^*$ of finite words such that an infinite word $\mathbf{x}$ belongs to X if and only if none of its factors belongs to $\mathscr{F}$. A subshift X is called a *subshift of finite type* if one can choose the set $\mathscr{F}$ to be finite. A subshift is said to be *sofic* if the set $\mathscr{F}$ is a regular language. A subshift (X, S) is said to be *periodic* if there exist $\mathbf{x} \in X$ and an integer k such that $X = \{\mathbf{x}, S\mathbf{x}, \ldots, S^k\mathbf{x} = \mathbf{x}\}$. Otherwise it is said to be *aperiodic*.

For the more general case of a group G acting on configurations in Σ^G, see Chapter 9. Elements of Σ^G can be considered as colorings of a group G by a finite alphabet Σ. The set of configurations Σ^G, endowed with the product topology, is a compact space on which we define the shift transformations: for every $g \in G$, the shift S^g translates a configuration $x \in \Sigma^G$ through $S^g(x)_h = x_{g^{-1}h}$ for every $h \in G$. In this framework, subshifts are exactly subsets of A^G that are both shift-invariant and closed for the product topology.

Example 1.7.5. The set of infinite words over $\{0, 1\}$ of Example 1.5.6 which do not contain the factor 11 is a subshift of finite type, whereas the set of infinite words over $\{0, 1\}$ having an even number of 1 between two occurrences of the letter 0 is a sofic subshift which is not of finite type.

Definition 1.7.6. Let Y be a subshift. For a word $w = w_0 \cdots w_r$, the *cylinder set* $[w]$ is the set $\{\mathbf{y} \in Y \mid y_0 = w_0, \ldots, y_r = w_r\}$.

The cylinder sets are *clopen* (open and closed) sets and form a basis of open sets for the topology of Y. Furthermore, one checks that a clopen set is a finite union of cylinders. In the bi-infinite case, the cylinders are the sets

$$[u.v]_Y = \{y \in Y \mid y_i = u_i, y_j = v_j,\ -|u| \leq i \leq -1,\ 0 \leq j \leq |v| - 1\}$$

and the same remark holds.

Then the *topological entropy* $h(X)$ of the symbolic dynamical system (X, S) measures the richness of its language L, defined as the set of factors of elements in X. It is defined as

$$h(X) = \lim_{n \to \infty} \frac{1}{n} \ln |L \cap \Sigma^n| .$$

It is closely related to the *growth rate* of the language L defined as $\limsup_{n \to \infty} |L \cap \Sigma^n|^{\frac{1}{n}}$ and considered in Chapter 5.