2. Introduction to finite automata and substitutions

The aim of this chapter is to introduce finite automata and substitutions by showing some typical and important examples of situations where these concepts appear. We will allude in this chapter to several themes that will be handled in detail in the next chapters. Special stress will be given to the statistical properties of these sequences. We first recall some properties of the Morse sequence, then we introduce the Rudin-Shapiro sequence and focus on its spectral properties. We also evoke the Baum-Sweet sequence, the Cantor sequence and the Fibonacci sequence.

For all the definitions related to words, substitutions and automata, we refer the reader to the first paragraphs of Chap. 1 (see also [135] and [285]).

2.1 The Morse sequence

As mentioned in Chap. 1, M. Morse introduced his famous sequence in the 1920's to show the existence of non periodic recurrent geodesics on some surfaces with negative curvature. This problem led him to the construction of a non periodic sequence u over the alphabet $\{a, b\}$ which is also uniformly recurrent, i.e., every word occurring in u occurs infinitely often with bounded gaps. For more details, see Chap. 3.

Let us recall that there are several ways to construct the Morse sequence u. One of them is to say that u is preserved if we substitute the word ab to the letter a and the word ba to the letter b in any place. The solution of this problem

$$u = abbabaabbaababbabaababbaabbabaab\ldots$$

is unique, up to the exchange of the letters a and b (without mentioning the trivial empty solution).

Definition 2.1.1. *The* Morse sequence u *is defined as the fixed point beginning by* a *of the* Morse substitution σ *defined over the alphabet* $\{a, b\}$ *by* $\sigma(a) = ab$ *and* $\sigma(b) = ba$.

Remark. Let us insist on the importance of the order of the letters in the definition of a substitution. For example the (non-empty) fixed points of the substitution over the alphabet $\{a, b\}$ defined by $\sigma(a) = ab$ and $\sigma(b) = ab$ are the periodic sequences $abababababab\ldots$ and $babababababa\ldots$.

Proposition 2.1.1. *Let* $(U_r)_{r\in\mathbb{N}}$ *and* $(U'_r)_{r\in\mathbb{N}}$ *be the sequences of words over the alphabet* $\{a,b\}$ *defined by* $U_0 = a$, $U'_0 = b$ *and for* $r \geq 0$ $\begin{cases} U_{r+1} = U_r U'_r \\ U'_{r+1} = U'_r U_r \end{cases}$. *Then for any positive integer* r, *one has* $|U_r| = |U'_r| = 2^r$ *and the word* U_r *is equal to the word* U'_r *modulo the exchange of the letters* a *and* b. *Moreover* $u = \lim_{r\to+\infty} U_r$.

Proof. It is easy to prove by induction over r that $U_r = \sigma^r(a)$ and $U'_r = \sigma^r(b)$ satisfy the properties. ■

2.1.1 Arithmetic definition of the Morse sequence

The Morse sequence can be defined by a property of the dyadic development of the integers. As shown in Chap. 1, this is a special case of a general property of sequences obtained as a fixed point of a constant length substitution: if the sequence v is a fixed point of a substitution of constant length q, then v can be defined by a property of the q-adic development of the integers (Proposition 1.5.1). Moreover this property is simple enough to be recognizable by a finite automaton.

Consider the Morse sequence. Let us denote by $\mathbb{N}_a$ (resp. $\mathbb{N}_b$) the set of integers n such that the $(n+1)$-th letter of the Morse sequence is a (respectively b):

$$\begin{aligned} \mathbb{N}_a &= \{0,3,5,6,9,10,12,15,\ldots\} \\ \mathbb{N}_b &= \{1,2,4,7,8,11,13,14,\ldots\}. \end{aligned}$$

For any integer n, we denote by $S_2(n)$ the sum of the dyadic digits of n i.e.,

$$S_2(n) = \sum_{i\geq 0} n_i \text{ if } n = \sum_{i\geq 0} n_i 2^i,\ n_i \in \{0,1\}$$

is the dyadic development of n.

Proposition 2.1.2. *We have*

$$\mathbb{N}_a = \{n \in \mathbb{N}, \quad S_2(n) \textit{ is even }\}$$

$$\mathbb{N}_b = \{n \in \mathbb{N}, \quad S_2(n) \textit{ is odd }\}.$$

Proof. Obviously $\mathbb{N}_a$ and $\mathbb{N}_b$ form a partition of $\mathbb{N}$. It is thus enough to show by induction over r the following property:

$$(\mathcal{P}_r) \begin{cases} n \in \mathbb{N}_a \text{ and } n < 2^r \Longrightarrow S_2(n) \text{ is even} \\ n \in \mathbb{N}_b \text{ and } n < 2^r \Longrightarrow S_2(n) \text{ is odd.} \end{cases}$$

Note that $(\mathcal{P}_0)$ is obvious. Suppose that $(\mathcal{P}_r)$ is true and consider $n \in \mathbb{N}_a$ (respectively $\mathbb{N}_b$) such that $2^r \leq n < 2^{r+1}$. Then $n = 2^r + n'$ with $n' \in \mathbb{N}_b$

(respectively $\mathbb{N}_a$) and $n' < 2^r$ by Proposition 2.1.1. As $S_2(n) = S_2(2^r + n') = S_2(n') + 1$, $(\mathcal{P}_{r+1})$ follows. ■

Let us recall that if we consider integers as words over the alphabet $\{0,1\}$ (via their dyadic development), then it is clear that the following 2-automaton with initial state a and exit map $Id_{\{a,b\}}$ recognizes the Morse sequence (either in direct reading or in reverse reading):

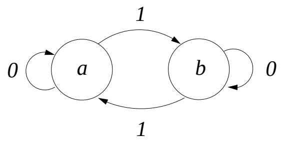


Fig. 2.1. Automaton associated with the Morse sequence

2.1.2 The problem of Prouhet

In fact, it seems that the first mathematician who introduced the Morse sequence was E. Prouhet. In 1851 he solved in [284] the following problem: given the positive integers q and r, find an infinite number of sequences of q^r numbers that can be cut in q sets of q^{r-1} elements such that, for any $k < r$, the sum of all the k-th powers of the elements of each set is the same.

If $q = 2$, the solution to Prouhet's problem is given by the Morse sequence in the following sense:

Proposition 2.1.3. *For any nonnegative integers k, r and n_0 such that $k < r$, we have*

$$\sum_{\substack{n \in n_0 + \mathbb{N}_a \\ n < n_0 + 2^r}} n^k = \sum_{\substack{n \in n_0 + \mathbb{N}_b \\ n < n_0 + 2^r}} n^k .$$

Proof. As for $\alpha \in \{a,b\}$

$$\begin{aligned} \sum_{\substack{n \in n_0 + \mathbb{N}_\alpha \\ n < n_0 + 2^r}} n^k &= \sum_{\substack{n \in \mathbb{N}_\alpha \\ n < 2^r}} (n_0 + n)^k \\ &= \sum_{i \le k} \binom{k}{i} n_0^{k-i} \sum_{\substack{n \in \mathbb{N}_\alpha \\ n < 2^r}} n^i , \end{aligned}$$

it is enough to prove the property in the case $n_0 = 0$. This is equivalent to saying that for any nonnegative integers k and r such that $k < r$, we have

$\sum_{n<2^r} (-1)^{S_2(n)} n^k = 0$. Let us introduce the polynomial $F_r \in \mathbf{Z}[X]$ defined by $F_r(X) = \sum_{n<2^r} (-1)^{S_2(n)} X^n = \prod_{k<r} \left(1 - X^{2^k}\right)$. One checks that $F_r(X) = (1 - X)^r G_r(X)$ with $G_r \in \mathbf{Z}[X]$, so that $F_r^{(k)}(1) = \sum_{n<2^r} (-1)^{S_2(n)} n(n-1)\dots(n-k+1) = 0$ for $k < r$, from which we deduce that $\sum_{n<2^r} (-1)^{S_2(n)} n^k = 0$ for $k < r$. ∎

For more details, see the survey [72].

2.1.3 A combinatorial property of the Morse sequence

The Morse sequence has been also studied at the beginning of this century by A. Thue. In 1906 and 1912 he wrote two papers that he considered as an attempt to open new ways to number theory the purpose of which was to construct sequences over finite alphabets containing no square or no cube (see [353] and [352]). In 1938, G. A. Hedlund and M. Morse used the same property to link the Morse sequence to the problem of unending chess ([257]). We will prove in the next chapter that the Morse sequence does not contain any cube and, more precisely, does not contain any word of the form WWw where W is a non-empty word with first letter w (see also Corollary 2.2.4 from [228]).

Exercise 2.1.1. Use Morse sequence to construct a square-free sequence over a three-letter alphabet containing no square (see also Sec. 2.3 from [228]).

In 1902 W. Burnside submitted the following problem: if G is a group of finite type such that there is an integer n such that $g^n = 1$ for any g in G, is G finite? W. Burnside gave a positive answer to this question in the case $n \leq 3$, Sanov in the case $n = 4$ in 1940, followed by M. Hall for the case $n = 6$ in 1957. In 1968, S. I. Adian and P. S. Novikov showed that the general answer to Burnside's problem is negative: there are counter-examples at least for any odd integer greater than or equal to 665. It is interesting to point out that their method uses fundamentally the construction of square-free sequences (see [2]).

2.1.4 A statistical property of the Morse sequence

By a theorem of Fréchet, any monotone function f can be decomposed as $f = f_1 + f_2 + f_3$, where f_1 is a monotone step-function, f_2 a monotone function which is the integral of its derivative, and f_3 a monotone continuous function which has almost everywhere a derivative equal to zero.

In [379] Wiener extended the spectrum theory to the harmonic analysis of functions defined for a countable set of arguments (that he called arrays).

As an application of some theorems proved in [379], Mahler gives in [235] a construction based on the array $(-1)^{S_2(n)}$ for which $f_3 \neq 0$ in the Fréchet decomposition. The crucial point is the following property:

Proposition 2.1.4. *For any positive integer k and any positive integer N, if*

$$\gamma_N(k) = \frac{1}{N} \sum_{n<N} (-1)^{S_2(n)}(-1)^{S_2(n+k)} ,$$

then for any k the sequence $(\gamma_N(k))_{N>0}$ converges and its limit is non-zero for infinitely many k.

The convergence of the sequence $(\gamma_N(k))_{N\geq 1}$ can be understood as a consequence of the unique ergodicity of the symbolic dynamical system associated with the infinite word $\mathcal{W}$ (see [285] for this approach). It follows also from [197] that $\lim_{N\to+\infty} \gamma_N(k)$ is equal to the $k-th$ Fourier coefficient of the correlation measure associated with the Morse dynamical system, which is the Riesz product $\prod_{n\geq 0}(1 - cos\ 2^n t)$.

Let us present here a proof of Proposition 2.1.4 independent of these results.

Proof. For $k=0$, we have for any positive integer $N, \gamma_N(0) = 1$.

For $k=1$, we have for any positive integer N,

$$\begin{aligned}
\gamma_{2N}(1) &= \frac{1}{2N} \sum_{n<2N} (-1)^{S_2(n)}(-1)^{S_2(n+1)} \\
&= \frac{1}{2N} \sum_{2n<2N} (-1)^{S_2(2n)}(-1)^{S_2(2n+1)} \\
&\qquad\qquad + \frac{1}{2N} \sum_{2n+1<2N} (-1)^{S_2(2n+1)}(-1)^{S_2(2n+2)} \\
&= \frac{1}{2N} \sum_{n<N} (-1)^{S_2(n)}(-1)^{S_2(n)} + \frac{1}{2N} \sum_{n<N} -(-1)^{S_2(n)}(-1)^{S_2(n+1)} \\
&= -\frac{1}{2} - \frac{1}{2}\gamma_N(1),
\end{aligned}$$

and for any nonnegative integer N

$$\begin{aligned}
\gamma_{2N+1}(1) &= \frac{1}{2N+1}\sum_{n<2N+1}(-1)^{S_2(n)}(-1)^{S_2(n+1)} \\
&= \frac{1}{2N+1}\sum_{2n<2N+1}(-1)^{S_2(2n)}(-1)^{S_2(2n+1)} \\
&\qquad\qquad + \frac{1}{2N+1}\sum_{2n+1<2N+1}(-1)^{S_2(2n+1)}(-1)^{S_2(2n+2)} \\
&= \frac{1}{2N+1}\sum_{n<N+1} -(-1)^{S_2(n)}(-1)^{S_2(n)} \\
&\qquad\qquad + \frac{1}{2N+1}\sum_{n<N} -(-1)^{S_2(n)}(-1)^{S_2(n+1)} \\
&= -\frac{N+1}{2N+1} - \frac{N}{2N+1}\gamma_N(1).
\end{aligned}$$

If we put $\delta_N = \gamma_N(1) + \frac{1}{3}$ for any strictly positive integer N, these two relations become

$$\begin{cases} \delta_{2N} = -\frac{1}{2}\delta_N \\ \delta_{2N+1} = -\frac{N}{2N+1}\delta_N - \frac{2}{3(2N+1)} \end{cases},$$

with $\delta_1 = \gamma_1(1) + \frac{1}{3} = -\frac{2}{3}$.

It is easy to check by induction over N that if $N \geq 2$, then $\mid \delta_N \mid \leq \frac{2}{3}\frac{\log_2 N}{N}$. Indeed, this is true for $N = 2$ and $N = 3$ ($\delta_2 = \frac{1}{3}$ and $\delta_3 = 0$); for $2N \geq 4$ we have

$$\mid \delta_{2N} \mid = \frac{1}{2} \mid \delta_N \mid \leq \frac{1}{2}\ \frac{2}{3}\ \frac{\log_2 N}{N} \frac{2}{3}\ \frac{\log_2 2N}{2N}$$

and for $2N+1 \geq 5$ we have

$$\begin{aligned}
\mid \delta_{2N+1} \mid &\leq \frac{N}{2N+1}\ \frac{2}{3}\ \frac{\log_2 N}{N} \quad + \quad \frac{2}{3(2N+1)} \\
&\leq \frac{2}{3(2N+1)}\quad(\log_2 N + 1) \\
&\leq \frac{2}{3}\quad\frac{\log_2(2N+1)}{2N+1}.
\end{aligned}$$

This computation shows that $\gamma_N(1) = -\frac{1}{3} + O(\log N/N)$ so that the sequence $(\gamma_N(1))_{N>0}$ converges to $-\frac{1}{3}$. It is now easy to deduce by induction over k the convergence of the sequences $(\gamma_N(k))_{N>0}$ for $k \geq 2$ from the convergence of the sequences $(\gamma_N(0))_{N>0}$ and $(\gamma_N(1))_{N>0}$:

- If $2k \geq 2$, we have for any positive integer N,

$$
\begin{aligned}
\gamma_{2N}(2k) &= \frac{1}{2N}\sum_{n<2N}(-1)^{S_2(n)}(-1)^{S_2(n+2k)}\\
&= \frac{1}{2N}\sum_{2n<2N}(-1)^{S_2(2n)}(-1)^{S_2(2n+2k)}\\
&\qquad\qquad +\frac{1}{2N}\sum_{2n+1<2N}(-1)^{S_2(2n+1)}(-1)^{S_2(2n+2k+1)}\\
&= \gamma_N(k),
\end{aligned}
$$

and for any nonnegative integer N,

$$
\begin{aligned}
\gamma_{2N+1}(2k) &= \frac{1}{2N+1}\sum_{n<2N+1}(-1)^{S_2(n)}(-1)^{S_2(n+2k)}\\
&= \sum_{2n<2N+1}(-1)^{S_2(2n)}(-1)^{S_2(2n+2k)}\\
&\qquad\qquad +\frac{1}{2N+1}\sum_{2n+1<2N+1}(-1)^{S_2(2n+1)}(-1)^{S_2(2n+2k+1)}\\
&= \frac{N+1}{2N+1}\gamma_{N+1}(k)+\frac{N}{2N+1}\gamma_N(k).
\end{aligned}
$$

It follows that if $\gamma(k)=\lim_{N\to+\infty}\gamma_N(k)$, both sequences $(\,\gamma_{2N}(2k)\,)_{N>0}$ and $(\,\gamma_{2N+1}(2k)\,)_{N\geq 0}$ converge to $\gamma(k)$, so that the sequence $(\,\gamma_N(2k)\,)_{N>0}$ converges and $\gamma(2k)=\lim_{N\to+\infty}\gamma_N(2k)=\gamma(k)$.

• If $2k+1\geq 3$, we have for any positive integer N,

$$
\begin{aligned}
\gamma_{2N}(2k+1) &= \frac{1}{2N}\sum_{n<2N}(-1)^{S_2(n)}(-1)^{S_2(n+2k+1)}\\
&= \frac{1}{2N}\sum_{2n<2N}(-1)^{S_2(2n)}(-1)^{S_2(2n+2k+1)}\\
&\qquad\qquad +\frac{1}{2N}\sum_{2n+1<2N}(-1)^{S_2(2n+1)}(-1)^{S_2(2n+2k+2)}\\
&= -\frac{1}{2}\gamma_N(k)-\frac{1}{2}\gamma_N(k+1),
\end{aligned}
$$

and for any nonnegative integer N,

$$\begin{aligned}
\gamma_{2N+1}(2k+1) &= \frac{1}{2N+1} \sum_{n<2N+1} (-1)^{S_2(n)} (-1)^{S_2(n+2k+1)} \\
&= \frac{1}{2N+1} \sum_{2n<2N+1} (-1)^{S_2(2n)} (-1)^{S_2(2n+2k+1)} \\
&\qquad + \frac{1}{2N+1} \sum_{2n+1<2N+1} (-1)^{S_2(2n+1)} (-1)^{S_2(2n+2k+2)} \\
&= \frac{N+1}{2N+1} \gamma_{N+1}(k) - \frac{N}{2N+1} \gamma_N(k+1).
\end{aligned}$$

It follows that if $\gamma(k) = \lim_{N\to\infty} \gamma_N(k)$ and $\gamma(k+1) = \lim_{N\to\infty} \gamma_N(k+1)$, both sequences $(\,\gamma_{2N}(2k+1)\,)_{N>0}$ and $(\,\gamma_{2N+1}(2k+1)\,)_{N\geq 0}$ converge to $-\frac{1}{2}\gamma(k) - \frac{1}{2}\gamma(k+1)$, so that the sequence $(\,\gamma_N(2k+1)\,)_{N>0}$ converges and $\gamma(2k+1) = \lim_{N\to\infty} \gamma_N(2k+1) = -\frac{1}{2}\gamma(k) - \frac{1}{2}\gamma(k+1)$.

It is now easy to deduce from the relations

$$\begin{cases} \gamma(0) = 1 \\ \gamma(2k) = \gamma(k) \\ \gamma(2k+1) = -\frac{1}{2}\gamma(k) - \frac{1}{2}\gamma(k+1) \end{cases}$$

the values of $\gamma(k)$ for any nonnegative integer k.

In particular for any integer i we have $\gamma(2^i) = -\frac{1}{3}$, which ends the proof of Proposition 2.1.4. ∎

2.2 The Rudin-Shapiro sequence

2.2.1 Definition

For any sequence $\varepsilon = (\varepsilon_n)_{n\in\mathbb{N}} \in \{-1,+1\}^{\mathbb{N}}$ we have

$$\begin{aligned}
&\int_0^1 \left| \sum_{n<N} \varepsilon_n e\,(n\theta) \right|^2 d\theta \\
&= \int_0^1 \sum_{\substack{n<N \\ n'<N}} \varepsilon_n \varepsilon_{n'} e\,((n-n')\theta)\, d\theta \\
&= N,
\end{aligned}$$

because

$$\begin{aligned}
\int_0^1 e\,(k\theta)\, d\theta &= 1 \quad \text{if} \quad k = 0 \\
&= 0 \quad \text{if} \quad k \in \mathbb{Z} \setminus \{0\}
\end{aligned}$$

(with the notation $e(x) = e^{2i\pi x}$).

We hence have

$$\sup_{\theta\in[0,1]} \left|\sum_{n<N} \varepsilon_n e\left(n\theta\right)\right| \geq \left(\int_0^1 \left|\sum_{n<N} \varepsilon_n e\left(n\theta\right)\right|^2 d\theta\right)^{\frac{1}{2}}$$
$$= \sqrt{N}\,.$$

R. Salem asked in 1950 the following question, linked to several problems in harmonic analysis (see Chap. X from [190]) : is it possible to find a sequence $\varepsilon \in \{-1,+1\}^{\mathbb{N}}$ such that there is a constant $c > 0$ for which

$$\sqrt{N} \leq \sup_{\theta\in[0,1]} \left|\sum_{n<N} \varepsilon_n e\left(n\theta\right)\right| \leq c\sqrt{N}$$

holds for any positive integer N?

H.S. Shapiro in 1951 and W. Rudin in 1959 gave independently a positive answer to this question in [326] and [304].

Definition 2.2.1. *The Rudin-Shapiro sequence ε over the alphabet $\{-1,+1\}$ is defined by the relation $\varepsilon_0 = 1$ and for any nonnegative integer n*

$$\begin{cases} \varepsilon_{2n} &= \varepsilon_n \\ \varepsilon_{2n+1} &= (-1)^n \varepsilon_n. \end{cases}$$

Remark.

The relations $u_0 = 1$ and $\begin{cases} u_{2n} &= u_n \\ u_{2n+1} &= -u_n \end{cases}$ for any nonnegative integer n define the Morse sequence over the alphabet $\{+1,-1\}$.

By analogy with the fact that u_n gives the parity of the sum of digits of n, i.e., the parity of the number of one's in the dyadic development of n, it is easy to verify that ε_n gives the parity of the number of words 11 in the dyadic development of n.

Proposition 2.2.1. *For any nonnegative integer n with dyadic development $n = \sum_{i\geq 0} n_i 2^i$ $(n_i \in \{0,1\})$ we have*

$$\varepsilon_n = (-1)^{\sum_{i\geq 0} n_i n_{i+1}}\,.$$

Furthermore for any nonnegative integers a, b and n such that $b < 2^n$, we have

$$\varepsilon_{2^{n+1}a+b} = \varepsilon_a \varepsilon_b.$$

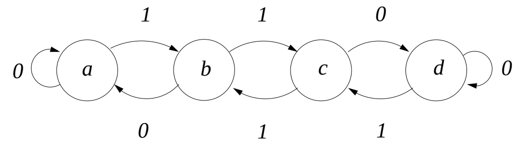


Fig. 2.2. Automaton associated with the the Rudin-Shapiro substitution

Proof. The proof can be deduced by induction from the relations $\varepsilon_{2n} = \varepsilon_n$ and $\varepsilon_{2n+1} = (-1)^n \varepsilon_n$. ∎

If we consider again integers as words over the alphabet $\{0,1\}$, it is clear that the 2-automaton on Figure 2.2 (in reverse reading or direct reading) with initial state a and exit map φ defined by $\varphi(a) = \varphi(b) = +1$ and $\varphi(c) = \varphi(d) = -1$ recognizes the Rudin-Shapiro sequence:

Proposition 2.2.2. *For any nonnegative integer N*

$$\sup_{\theta\in[0,1]} \left| \sum_{n<N} \varepsilon_n e\,(n\theta) \right| \le (2+\sqrt{2})\sqrt{N}\ .$$

Proof. For any nonnegative integer n, put

$$S_n(\theta) = \sum_{k<2^n} \varepsilon_k e\,(k\theta)\,,$$

$$S'_n(\theta) = \sum_{k<2^n} (-1)^k \varepsilon_k e\,(k\theta)\,,$$

$$M_n(\theta) = \begin{pmatrix} 1 & e\,(2^n\theta) \\ 1 & -e\,(2^n\theta) \end{pmatrix}\ .$$

As

$$S_{n+1}(\theta) = S_n(2\theta) + e(\theta) S'_n(2\theta) \text{ and } S'_{n+1}(\theta) = S_n(2\theta) - e(\theta) S'_n(2\theta),$$

we have

$$\begin{pmatrix} S_n(\theta) \\ S'_n(\theta) \end{pmatrix} = \prod_{k<n} M\,(2^k\theta) \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and

$$\begin{aligned} |S_n(\theta)| &\le \sqrt{|S_n(\theta)|^2 + |S'_n(\theta)|^2} \\ &\le \sqrt{2} \prod_{k<n} \left\| M\,(2^k\theta) \right\|_2 \\ &= \sqrt{2}\, 2^{n/2}\,. \end{aligned}$$

Now, if $N = 2^{n_1} + 2^{n_2} + \cdots + 2^{n_k}$, with $n_1 > n_2 > \cdots > n_k$, we have

$$\sum_{k<N} \varepsilon_k e\,(k\theta) = S_{n_1}(\theta) + e\,(2^{n_1}\theta) S_{n_2}(\theta) + \dots$$
$$+ \, e\,\left((2^{n_1} + \dots + 2^{n_{k-1}})\,\theta\right)\, S_{n_k}(\theta),$$

since it follows from Proposition 2.2.1 that for any nonnegative integers k and n, then $\varepsilon_{2^n+k} = \varepsilon_k$, if $k < 2^{n-1}$.

We deduce from this equality that

$$\left|\sum_{k<N} \varepsilon_k e\,(k\theta)\right| \leq \sqrt{2}\,\left(2^{n_1/2} + 2^{n_2/2} + \dots + 2^{n_k/2}\right)$$

and Proposition 2.2.2 follows from the following lemma:

Lemma 2.2.1. *For any nonnegative integers* $n_1, n_2, \cdots, n_k$ *with* $n_1 > n_2 > \cdots > n_k$, *we have*

$$2^{n_1/2} + 2^{n_2/2} + \dots + 2^{n_k/2} \;\leq\; (1+\sqrt{2})\sqrt{2^{n_1} + \dots + 2^{n_k}}\ .$$

Proof. The proof works by induction over k. The property is obvious in the case $k = 1$. Suppose that for any nonnegative integers $n_1,\, n_2, \cdots, n_k$ with $n_1 > n_2 > \cdots > n_k$, we have

$$2^{n_1/2} + 2^{n_2/2} + \dots + 2^{n_k/2} \;\leq\; (1+\sqrt{2})\sqrt{2^{n_1} + \dots + 2^{n_k}}\ .$$

Then for any nonnegative integers $n_1,\, n_2, \cdots, n_{k+1}$ with $n_1 > n_2 > \cdots > n_{k+1}$, we get

$$2^{n_1/2} + 2^{n_2/2} + \dots + 2^{n_{k+1}/2} \;\leq\; 2^{n_1/2} + (1+\sqrt{2})\sqrt{2^{n_2} + \dots + 2^{n_{k+1}}}$$

by induction.

But as $2^{n_2} + \dots + 2^{n_{k+1}} \;\leq\; 2^{n_1}$, we have

$$2^{n_1/2}\sqrt{2^{n_2} + \dots + 2^{n_{k+1}}} \;\leq\; 2^{n_1}$$

and

$$2^{n_1} + 2(1+\sqrt{2})\, 2^{n_1/2}\sqrt{2^{n_2} + \dots + 2^{n_{k+1}}} + (3+2\sqrt{2})\,(2^{n_2} + \dots + 2^{n_{k+1}})$$
$$\leq (3+2\sqrt{2})\,(2^{n_1} + 2^{n_2} + \dots + 2^{n_{k+1}})$$

i.e.,

$$2^{n_1/2} + (1+\sqrt{2})\sqrt{2^{n_2} + \dots + 2^{n_{k+1}}}$$
$$\leq (1+\sqrt{2})\sqrt{2^{n_1} + 2^{n_2} + \dots + 2^{n_{k+1}}}\ ,$$

which ends the proof. ∎

2.2.2 Statistical properties of the Rudin-Shapiro sequence

As the Rudin-Shapiro sequence is defined by a very simple algorithm, one should expect its behavior to be very far from that of a "random" sequence. In particular, one should expect, as in the case of the Morse sequence (see Proposition 2.1.4) that the Rudin-Shapiro sequence has positive correlations.

Nevertheless the following proposition shows a result in the opposite direction:

Proposition 2.2.3. *For any nonnegative integers k and N, define $\gamma_N(k) = \frac{1}{N}\sum_{n<N} \varepsilon_n\varepsilon_{n+k}$. Then for any nonnegative integer k the sequence $(\gamma_N(k))_{N>0}$ converges and*

$$\lim_{N\to\infty} \gamma_N(k) = 0 \quad \textit{for every} \quad k \geq 1 .$$

Proposition 2.2.3 is a corollary of the more precise following result, due to C. Mauduit and A. Sárközy:

Proposition 2.2.4. *For any positive integers k and N, we have*

$$\mid \sum_{k<N} \varepsilon_n\varepsilon_{n+k} \mid < 2k + 4k\log_2 \frac{2N}{k}.$$

Proof. Let us first consider the case where N is a power of 2, $N = 2^M$ with M a nonnegative integer.

If we put for any positive integers k and M

$$S_M(k) = \sum_{n<2^M} \varepsilon_n\varepsilon_{n+k} \text{ and } S'_M(k) = \sum_{n<2^M} = (-1)^n\varepsilon_n\varepsilon_{n+k},$$

then we have

$$\begin{aligned} S_{M+1}(2k) &= \sum_{n<2^{M+1}} \varepsilon_n\varepsilon_{n+2k} \\ &= \sum_{n<2^M} \varepsilon_{2n}\varepsilon_{2n+2k} + \sum_{n<2^M} \varepsilon_{2n+1}\varepsilon_{2n+2k+1} \\ &= \sum_{n<2^M} \varepsilon_n\varepsilon_{n+k} + (-1)^k \sum_{n<2^M} \varepsilon_n\varepsilon_{n+k} \\ &= (1+(-1)^k)S_M(k), \end{aligned}$$

and

$$\begin{aligned} S_{M+1}(2k+1) &= \sum_{n<2^{M+1}} \varepsilon_n\varepsilon_{n+2k+1} \\ &= \sum_{n<2^M} \varepsilon_{2n}\varepsilon_{2n+2k+1} + \sum_{n<2^M} \varepsilon_{2n+1}\varepsilon_{2n+2k+2} \\ &= \sum_{n<2^M} (-1)^{n+k}\varepsilon_n\varepsilon_{n+k} + \sum_{n<2^M} (-1)^n\varepsilon_n\varepsilon_{n+k+1} \\ &= (-1)^k)S'_M(k) + S'_M(k+1). \end{aligned}$$

By a similar computation, we get

$$S'_{M+1}(2k) = (1-(-1)^k)S_M(k)$$

and

$$S'_{M+1}(2k+1) = (-1)^k S'_M(k) - S'_M(k+1).$$

In particular we have $S'_M(0) = S'_M(1) = \delta_0(M)$ and $S_M(1) = \delta_0(M) + 2\delta_1(M)$, where δ_0 and δ_1 are Dirac measures.

It is now easy to deduce by induction over k from these four relations that for any positive integer k and any nonnegative integer M, we have

$$|S_M(k)| \leq 2k \text{ and } |S'_M(k)| \leq 2k.$$

Now, if $N = \sum_{i=1}^{l} 2^{M_i}$ is the dyadic expansion of the positive integer N, it is easy to check that

$$\sum_{n<N} \varepsilon_n \varepsilon_{n+k} = \sum_{i=1}^{l} \sum_{n<2^{M_i}} \varepsilon_{(n+\sum_{1\leq j<i} 2^{M_j})} \cdot \varepsilon_{(n+k+\sum_{1\leq j<i} 2^{M_j})}.$$

Let us define the integer ν by

$$2^\nu \leq k < 2^{\nu+1}$$

If $i \in I^+ = \{j \in \{1,\ldots,l\}, M_j > \nu\}$ we have

$$\sum_{n<2^{M_i}} \varepsilon_{(n+\sum_{1\leq j<i} 2^{M_j})} \cdot \varepsilon_{(n+k+\sum_{1\leq j<i} 2^{M_j})}$$
$$= S_{M_i}(k) + \sum_{n<2^{M_i}} \left(\varepsilon_{(n+\sum_{1\leq j<i} 2^{M_j})} \cdot \varepsilon_{(n+k+\sum_{1\leq j<i} 2^{M_j})} - \varepsilon_n \varepsilon_{n+k} \right),$$

because it follows from Proposition 2.2.2 that

$$\sum_{n<2^{Mi}-k} \varepsilon_{(n+\sum_{1\leq j<i} 2^{M_j})} \varepsilon_{(n+k+\sum_{1\leq j<i} 2^{M_j})}$$
$$= \sum_{n<2^{Mi}-k} \left(\varepsilon_{(\sum_{1\leq j<i} 2^{M_j})} \right)^2 \varepsilon_n \varepsilon_{n+k} = \sum_{n<2^{Mi}-k} \varepsilon_n \varepsilon_{n+k}.$$

This shows that

$$|\sum_{i\in I^+}\sum_{n<2^{M_i}} \varepsilon_{(n+\sum_{1\le j<i} M_j)} \cdot \varepsilon_{(n+k+\sum_{1\le j<i} 2^{M_j})}| \le \sum_{i\in I+}\left(|S_{M_i}(k)| + 2k\right)$$
$$\le 4k \text{ Card } I^+.$$

Now

$$\begin{aligned}\text{Card}I^+ &= M_1 - \nu \\ &= [\log_2 N] - [\log_2 k] \\ &\le log_2 N - log_2 k + 1,\end{aligned}$$

which gives

$$|\sum_{i\in I^+}\sum_{n<2^{M_i}} \varepsilon_{(n+\sum_{1\le j<i} 2^{M_j})} \cdot \varepsilon_{(+k+\sum_{1\le j<i} 2^{M_j})}| \le 4k\log_2 \frac{2N}{k}.$$

If we put $I^- = \{i \in \{1,\dots,l\}, M_i \le \nu\}$, then we have

$$|\sum_{i\in I^-}\sum_{n<2^{M_i}} \varepsilon_{\left(n+\sum_{1\le j<i} 2^{M_j}\right)} \varepsilon_{\left(n+k+\sum_{1\le j<i} 2^{M_j}\right)}| \le \sum_{i\in \dot{I}^-} 2^{M_i}$$
$$\le 2^{\nu+1} - 1 \ < \ 2k.$$

Finally we proved that

$$|\sum_{n<N} \varepsilon_n \varepsilon_{n+k}| \le 2k + 4k\log_2 \frac{2N}{k}.$$

■

Exercise 2.2.1. What can be said about the multiple correlations of the Rudin-Shapiro sequence defined for nonnegative integers $k_1, k_2, \cdots, k_p$ and N by

$$\gamma_N(k_1, k_2, \dots, k_p) = \frac{1}{N}\sum_{n<N} \varepsilon_n \varepsilon_{n+k_1} \varepsilon_{n+k_2} \dots \varepsilon_{n+k_p}.$$

2.3 The Baum-Sweet sequence

It is well known that the continued fraction expansion of an irrational algebraic number x is ultimately periodic if and only if x is quadratic (this result is known as *Lagrange's Theorem*). But we know almost nothing about the expansion of nonquadratic irrational algebraic numbers. In particular, the following question seems to be of the highest difficulty: is there any algebraic number of degree at least 3 with a bounded continued fraction expansion (i.e., such that the partial quotients in the expansion are bounded)? If we replace $\mathbb{R}$ by the field $\mathbb{F}_2((X^{-1}))$ of formal power series in X^{-1} on $\mathbb{F}_2$ (where $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ denotes the finite field with two elements), the analogous problem was solved in 1976 by L. E. Baum and M. M. Sweet. Indeed they give in [48] an example of algebraic element of degree 3 with a bounded continued fraction expansion (i.e., with partial quotients in $\mathbb{F}_2[X]$ of bounded degree).

Definition 2.3.1. *The* Baum-Sweet sequence $(f_n)_{n\in\mathbb{N}}$ *with values in the alphabet* $\mathbb{F}_2$ *is defined by*

$$\begin{aligned} f_n &= 0 \ \textit{ if the dyadic development of } n \textit{ contains} \\ &\qquad \textit{at least one odd string of } 0's\,, \\ &= 1 \ \textit{ if not}\,. \end{aligned}$$

It is clear from this definition that the Baum-Sweet sequence is recognized by the following 2-automaton (in reverse and direct reading) with initial state a and exit map φ defined by $\varphi(a) = \varphi(b) = 1$ and $\varphi(c) = \varphi(d) = 0$.

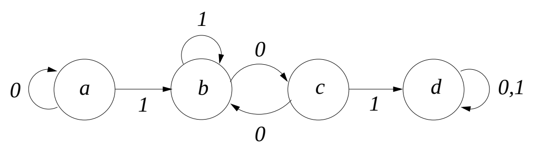


Fig. 2.3. Automaton associated with the the Baum-Sweet sequence

The following proposition illustrates a typical property of sequences generated by an automaton. It is an illustration of a fundamental theorem on q-automata (for q power of a prime) due to G. Christol, T. Kamae, M. Mendes-France and G. Rauzy that will be discussed in Chap. 4 (see also [99]).

A formal power series f is said to be *algebraic* over $\mathbb{F}_2(X)$ if there exists a non-trivial polynomial P with coefficients in $\mathbb{F}_2(X)$ such that $P(f) = 0$.

Proposition 2.3.1. *Let* $f = \sum_{n\geq 0} f_n X^{-n}$ *be the formal power series with coefficients given by the Baum-Sweet sequence. Then*

- f *is an algebraic element of degree 3 over* $\mathbb{F}_2(X)$*;*
- *furthermore, the continued fraction expansion of* f *is bounded and consists of elements of the set* $\{1, X, X+1, X^2, X^2+1\}$.

Proof. Let us give a proof of part i). We do not give here a proof of part ii) which is rather a special and surprising property of the Baum-Sweet sequence than a general property of automata. Note that there is no characterization of algebraic elements over $\mathbb{F}_q(X)$ (q power of a prime) the continued fraction expansion of which is bounded, and this problem seems to be hopelessly difficult.

It follows from the definition that for any nonnegative integer n we have

$$\begin{cases} f_{2n+1} = f_n \\ f_{4n} \quad = f_n \\ f_{4n+2} = 0 \end{cases}, \text{ with } f_0 = 1.$$

Now if we write

$$\begin{aligned} f &= \sum_{n\geq 0} f_n X^{-n} \\ &= \sum_{n\geq 0} f_{2n+1} X^{-(2n+1)} + \sum_{n\geq 0} f_{4n+2} X^{-(4n+2)} \\ &+ \sum_{n\geq 0} f_{4n} X^{-4n} , \end{aligned}$$

we see that

$$\begin{aligned} f &= X^{-1} \sum_{n\geq 0} f_n X^{-2n} + \sum_{n\geq 0} f_n X^{-4n} \\ &= X^{-1} f^2 + f^4 \end{aligned}$$

and f is solution of the algebraic equation of degree 3

$$X f^3 + f + X = 0 .$$

■

For more details on continued fraction expansions of formal power series, see the surveys [64, 217, 314].

2.4 The Cantor sequence

The *Cantor substitution* τ is defined over the alphabet $\{a, b\}$ by

$$\begin{aligned} \tau(a) &= aba \\ \tau(b) &= bbb . \end{aligned}$$

Definition 2.4.1. *The sequence of words* $(\tau^n(a))_{n\in\mathbb{N}}$ *converges to the fixed point* C *of the Cantor substitution beginning by* a*. This sequence is called the* Cantor sequence*. Its first terms are:*

$$aba\,bbb\,aba\,bbb\,bbb\,bbb\,aba\,bbb\,aba \ldots$$

Let us denote by C_a *the set of integers* n *such that the* $(n+1)$*-th letter of the Cantor sequence is* a*.*

It follows from Proposition 1.5.1 that the Cantor sequence v is recognized by the following 3-automaton in direct reading with initial state a and exit map $Id_{\{a,b\}}$:

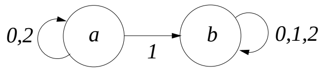


Fig. 2.4. Automaton associated with the Cantor sequence

This proves the following proposition:

Proposition 2.4.1.

$$C_a = \left\{ n \in \mathbb{N},\ n = \sum_{i \geq 0} n_i 3^i,\ \textit{with}\ \forall i \geq 0\ \ n_i \in \{0, 2\} \right\} .$$

2.5 An application of substitutions to criterions of divisibility

The four examples of sequences we have seen (Morse, Rudin-Shapiro, Baum-Sweet and Cantor) are fixed points of substitutions of constant length 2 or 3. As we pointed out several times, this means that they can be defined by an automatic property of the q-adic development of the integers.

We want to give now an application of Proposition 1.5.1 to the resolution of the following problem: given integers d and q greater or equal to 2, is it possible to decide only from its q-adic development whether a given positive integer is divisible by d or not?

Let v be the periodic sequence with values in the alphabet $\{0, 1, \ldots, d-1\}$:

$$01 \ldots (d-1) 01 \ldots (d-1) 01 \ldots (d-1) \ldots .$$

To solve our problem, we need to construct a q-automaton which recognizes the sequence v.

To obtain this q-automaton, it is enough to find a substitution ϱ of constant length q such that v is a fixed point of ϱ.

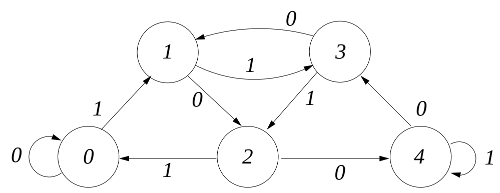
This can be easily done by cutting v into words of length q and rewriting v as $v = \varrho(0)\varrho(1) \ldots \varrho(d-1)\, \varrho(0)\varrho(1) \ldots \varrho(d-1) \ldots$.

To avoid heavy notations we will give explicitly the final automaton only in the case $q = 2$ and $d = 5$:

$$v = 01234012340123401234012340123401234012340123401234 \ldots .$$

$$\begin{aligned}\varrho(0) &= 01\\ \varrho(1) &= 23\\ \varrho(2) &= 40\\ \varrho(3) &= 12\\ \varrho(4) &= 34\end{aligned}$$

Clearly v is fixed point of ϱ, and it follows from Proposition 1.5.1 that v is recognized by the following 2-automaton with initial state 0 and exit map $Id_{\{0,1,2,3,4\}}$.



2.6 The Fibonacci sequence

2.6.1 Definition

After all these examples of sequences obtained as a fixed point of a constant length substitution and their arithmetic description, it is natural to look for non-constant length substitutions and to ask whether their fixed point can be defined by some simple arithmetic property of the integers.

This lead us to associate some generalized systems of numeration to substitutions. In the most general context, this problem is still linked to several open questions : we do not yet understand all the connections between the arithmetics and the geometry of the substitution dynamical systems (see [285]). A detailed survey of the literature on this subject is given in Chap. 7.

This fifth and last introductory example shows what happens in a case where both the arithmetics and the geometry are well understood. It is the "simplest" sequence among Sturmian sequences introduced in the 1940's by G. A. Hedlund and M. Morse in [255] and [256]. More details on the Fibonacci sequence will be found in Chap. 3, 6, and 9.

Definition 2.6.1. *The* Fibonacci sequence w *is the unique non-empty fixed point* w *of the* Fibonacci substitution φ *defined over the alphabet* $\{a,b\}$ *by*

$$\begin{cases}\varphi(a) = ab\\ \varphi(b) = a\end{cases}$$

The first terms of w are

$$abaababaabaababaababaabaababaabaab\dots.$$

2.6.2 The Fibonacci system of numeration

Let $(F)_{n\in\mathbb{N}}$ be the sequence of integers defined by $F_0 = 1$, $F_1 = 2$ and for any integer $n \geq 1$, $F_{n+1} = F_n + F_{n-1}$.

Proposition 2.6.1. *Every nonnegative integer n can be written in a unique way as $n = \sum_{i\geq 0} n_i F_i$ with $n_i \in \{0,1\}$ and $n_i n_{i+1} = 0$ for any $i \geq 0$.*

Proof.

- Let us first prove the existence of the decomposition. It is true for $n = 0$, $n = 1$ and $n = 2$. Let us suppose that the property is true for any integer $n < F_k$ (with $k \geq 2$).
  Then if $F_k \leq n < F_{k+1} = F_k + F_{k+1}$, we have $n - F_k < F_{k-1}$ and by hypothesis, $n - F_k = \sum_{i=0}^{k-2} n_i F_i$, hence $n = F_k + \sum_{i=0}^{k-2} n_i F_i$, and the property is true by induction over k.
- Let us prove now the unicity of the expansion. Let us remark that if $(n_0, n_1, \dots, n_k) \in \{0,1\}^{k+1}$ and $n_i n_{i+1} = 0$ for any $i < k$, then $\sum_{i=0}^{k} n_i F_i < F_{k+1}$. This can be easily be proved by induction. Indeed if $n_k = 1$, then $n_{k-1} = 0$ and $\sum_{i=0}^{k} n_i F_i = F_k + \sum_{i=0}^{k-2} n_i F_i$. Hence if we suppose that $\sum_{i=0}^{k-2} n_i F_i < F_{k-1}$ we get that $\sum_{i=0}^{k} n_i F_i < F_k + F_{k-1} = F_{k+1}$.
  It is now easy to prove the unicity of the decomposition : let us suppose that $\sum_{i\geq 0} n_i F_i = \sum_{i\geq 0} n'_i F_i$ with $n_i, n'_i \in \{0,1\}$ and $n_i n_{i+1} = n'_i n'_{i+1} = 0$ for any $i \geq 0$.
  Only to simplify the terms of higher indices, we can suppose that $n_k = 1$, $n'_k = 0$ and $n_i = n'_i = 0$ for $i > k$. This would imply that $\sum_{i\geq 0} n_i F_i \geq F_k$ and, because of our previous remark, that $\sum_{i\geq 0} n'_i F_i < F_k$, which is a contradiction.

■

Definition 2.6.2. *If $n = \sum_{i=0}^{k} n_i F_i$ with $n_k = 1$, $n_i \in \{0,1\}$ and $n_i n_{i+1} = 0$ for any $i < k$, we say that $Fib(n) = n_k n_{k-1} \dots n_0 \in \{0,1\}^{k+1}$ is the* Fibonacci expansion *of the positive integer n.*

Let us denote by $\mathbb{F}_a$ (respectively $\mathbb{F}_b$) the sets of integers n such that the $(n+1)$-th letter of the Fibonacci sequence is a (respectively b):

$$\mathbb{F}_a = \{0, 2, 3, 5, 7, 8, 10, 11, 13, 15, \ldots\}$$

$$\mathbb{F}_b = \{1, 4, 6, 9, 12, 14, 17, 19, 22, 25, \ldots\}$$

Proposition 2.6.2. *We have*

$$\mathbb{F}_a = \{n \in \mathbb{N},\ Fib(n) \in \{0,1\}^\star 0\}$$

$$\mathbb{F}_b = \{n \in \mathbb{N},\ Fib(n) \in \{0,1\}^\star 1\}.$$

Proof. It follows from Definition 2.6.1 that $F = \lim_{n\to\infty} \varphi^r(a)$, $|\varphi^r(a)| = F_r$, and $\varphi^{r+1}(a) = \varphi^r(a)\varphi^{r-1}(a)$, for any $r \geq 1$.

This last relation shows that if $\mathrm{Fib}(n) = n_k n_{k-1} \ldots n_0$, then $n \in \mathbb{F}_a$ if and only if $n - F_k \in \mathbb{F}_a$. As $0 \in \mathbb{F}_a$ and $1 \in \mathbb{F}_b$ we deduce from this that $n \in \mathbb{F}_a$ (respectively $n \in \mathbb{F}_b$) if and only if $n_0 = 0$ (respectively $n_0 = 1$). ■

Remark. Chapter 5 develops this approach in a more systematic way by associating partitions of the set of positive integers first to Sturmian sequences (these are the so-called *Beatty partitions*) and second, to fixed points of substitutions.