Block variables for deterministic aperiodic sequences



1997 J. Phys. A: Math. Gen. 30 7057

(http://iopscience.iop.org/0305-4470/30/20/011)

J. Phys. A: Math. Gen. **30** (1997) 7057–7065. Printed in the UK PII: S0305-4470(97)81999-3

# Block variables for deterministic aperiodic sequences

Michael Hörnquist†

Department of Physics and Measurement Technology, Linköping University, S-581 83 Linköping, Sweden

Received 19 February 1997, in final form 22 May 1997

**Abstract.** We use the concept of block variables to obtain a measure of order/disorder for some one-dimensional deterministic aperiodic sequences. For the Thue–Morse sequence, the Rudin–Shapiro sequence and the period-doubling sequence it is possible to obtain analytical expressions in the limit of infinite sequences. For the Fibonacci sequence, we present some analytical results which can be supported by numerical arguments. It turns out that the block variables show a wide range of different behaviour, some of them indicating that some of the considered sequences are more 'random' than other. However, the method does not give any definite answer to the question of which sequence is more disordered than the other and, in this sense, the results obtained are negative. We compare this with some other ways of measuring the amount of order/disorder in such systems, and there seems to be no direct correspondence between the measures.

## 1. Introduction

There has been much interest in the classification of different aperiodic sequences during the last decades. The reason for this comes from both physical as well as non-physical problems. The main question has been how 'random' a certain sequence is, a question which turns out to have no definite answer, but instead depends upon which measure we choose to use.

A brief historical outline might look like the following: Both the experimental discovery of incommensurate crystals in the early 1960s [1] and of quasicrystals in 1984 [2] as well as the fabrication of a Fibonacci superlattice in 1985 [3] have inspired much theoretical and experimental work concerning systems with order between periodic and random. Due to mathematical constraints, much theoretical work has been restricted to one-dimensional models, such as the Fibonacci model, the Thue–Morse model and the Rudin–Shapiro model. In particular, which systems cause localization for the wavefunction governed by the Schrödinger equation in the tight-binding formalism with only nearest-neighbour interaction has been studied. It is well known that in one dimension, a randomly ordered system has all eigenstates localized (Anderson localization) and a spectrum which is pure point (this statement is not absolutely true, e.g. a system built from random dimers has extended states [4], and if we consider a specific transfer-model with hopping integrals of different signs, all states are extended [5]), while a periodic system has extended (Bloch) states and an absolutely continuous spectrum. It has been rigorously proven that the systems generated by, for example, the Fibonacci sequence and the Thue–Morse sequence have singular continuous spectra and wavefunctions which are neither localized nor extended in

† E-mail address: mih@ifm.liu.se

 

the ordinary sense (they are often denoted as 'critical'). Other physical systems with an aperiodic order which have been studied are, for example, the Ising quantum chain [6] and vibrating lattices with anharmonic potentials [7]. In both these cases, a behaviour which differs from both a periodic structure and a random structure is found.

Another important area which deals with the classification of aperiodic sequences is the theory of communication [8]. Here it is important to be able to distinguish between the noise and the significant parts of a sequence. This is often connected with the more general study of symbolic dynamics [9], which has developed to become a research area in its own right. In particular, much attention has been paid to the period-doubling sequence, which describes the behaviour of any system at the Myrberg point (the accumulation point of the period-doubling cascade [9]). It has also been suggested that the concepts from symbolic dynamics might be useful in the analysis of DNA sequences [10].

There have been several attempts to find a simple measure for the amount of order for aperiodic sequences, e.g. an entropy concept from information theory [11], a quantity called 'log-entropy' [12], $\chi^2$-tests [13], spectral tests [10, 13], and energy spectrum [14] have been considered. Furthermore, the behaviour of the Fourier transform has been intensively studied [14], also with ordinary multifractal analysis, capturing the global behaviour [15] and with wavelets to obtain the local properties [16]. However, depending on which property we consider, different sequences qualify as 'most ordered' or 'most disordered'.

In this paper, we apply the formalism of *block variables* [17] in order to classify the different sequences mentioned above. This method has previously been used in the context of the protein folding problem to distinguish random sequences from non-random sequences in [18]. There it is claimed that it is indeed possible to distinguish between these two types, at least if real proteins are considered as the ordered sequences and the two different building blocks are hydrophobic and hydrophilic residues, respectively.

In the next section, we will introduce the formalism and define the block variables and some other related entities of interest. We will also state some already known results when averaging over all sequences with a fixed ratio between the numbers of different elements. Section 3 contains the explicit definitions of the deterministic sequences we will study and the calculations of their block variables. Finally in section 4, we draw some conclusions and make an outlook.

## 2. Formalism

Let $\{V_n\}_{n=1}^N$ be a sequence whose elements are $+1$ or $-1$. The block variables are then defined for all $s$ which divides $N$ as

$$\sigma_i^{(s)} = \sum_{n=1}^{s} V_{(i-1)s+n} \qquad i = 1, \ldots, \frac{N}{s}. \tag{1}$$

If the $V_n$s are independent random numbers drawn from the same distribution, there is no correlation between the different $V_n$s, and the quantity $\sigma_i^{(s)}$ will scale linearly with $s$. We will use this fact as one way of examining whether a given sequence is ordered or disordered.

An important quantity to study in this context is the normalized mean-square fluctuation of the block variables. To obtain this entity, we start by normalizing the block variables according to

$$\psi_i^{(s)} = \frac{1}{K}(\sigma_i^{(s)} - H_s(N))^2 \qquad i = 1, \ldots, \frac{N}{s} \tag{2}$$

where

$$H_s(N) = \frac{s}{N}\sum_{j=1}^{N/s}\sigma_j^{(s)} \qquad \text{and} \qquad K = \frac{4N_+N_-}{N(N-1)}\left(1-\frac{s}{N}\right). \tag{3}$$

Here $N_+$ ($N_-$) denotes the number of $+1$s ($-1$s) in the sequence. The number $K$ is chosen such that if we average over all sequences with a fixed length $N$ and a fixed number of positive elements $N_+$, we have $\langle\psi_i^{(s)}\rangle_{N,N_+} = s$ [18]. Note also that the sum $H_s(N)$ in (3) is easily calculated by the observation $\sum_{j=1}^{N/s}\sigma_j^{(s)} = N_+ - N_-$. The normalized mean-square fluctuation is now defined and calculated as

$$\psi^{(s)} = \frac{s}{N}\sum_{i=1}^{N/s}\psi_i^{(s)} \qquad i = 1, \ldots, \frac{N}{s}. \tag{4}$$

Due to the choice of the constant $K$, asserting that $\langle\psi_i^{(s)}\rangle_{N,N_+} = s$, we have as an immediate consequence that $\langle\psi^{(s)}\rangle_{N,N_+} = s$. We take this behaviour as an indication of a disordered sequence, i.e. if a specific sequence has a normalized mean-square fluctuation which behaves like the average of all sequences, we say it is disordered.

## 3. Results

We will focus our attention on the Thue–Morse sequence [19], the Rudin–Shapiro sequence [20], the period-doubling sequence [9] and the Fibonacci sequence [21]. For the first three sequences, we will analytically calculate both the block variables themselves and their normalized mean-square fluctuation. For the Fibonacci sequence, we will obtain some approximate analytical results for the block variables and confirm these results with numerical calculations. Also the normalized mean-square fluctuation will be obtained numerically.

### *3.1. Thue–Morse sequence*

If we consider the Thue–Morse sequence as composed by ones and zeros, it can be defined recursively as

$$m_0 = 0 \tag{5}$$

$$m_{2n} = m_n \tag{6}$$

$$m_{2n+1} = 1 - m_n. \tag{7}$$

(An alternative way of defining the Thue–Morse sequence is to count the parity of the sum of digits in base two of the integer $n$ corresponding to the element $m_n$ [22].) Before we apply the formalism described in the previous section, we change the numbering of the elements such that the first element has index one, and change all ones to minus ones and all zeros to ones, i.e. $V_n = 1 - 2m_{n-1}$, $n = 1, 2, 3, \ldots$. It has been shown that this sequence can also be obtained from the substitution rule A $\rightarrow$ AB, B $\rightarrow$ BA, see [22]. We start with a seed A and apply the substitution rule repeatedly. This yields the generations A, AB, ABBA, ABBABAAB etc. Finally, we replace each A with $+1$ and each B with $-1$.

We consider finite chains obtained by applying the substitution rule $k$ times, i.e. chains consisting of $2^k$ elements. From (1) above, $\sigma_i^{(1)} = V_i$ for all $i$. Otherwise, when $s \geqslant 2$ and even (because it has to be a divisor of $N = 2^k$) we have

$$\sigma_i^{(s)} = 0 \qquad \text{for all } i. \tag{8}$$

This is because the Thue–Morse sequence can be considered to be constructed from the two building blocks AB and BA. This also leads to $N_+ = N_- = N/2$, which for the only non-zero case, $s = 1$, yields $K = 1$, independent of the value of $N$ (as long as $N = 2^k$) and the sum $H_s(N)$ in (3) becomes zero. The normalized block variables then become

$$\psi_i^{(s)} = \frac{1}{K}(\sigma_i^{(s)})^2 = \begin{cases} 1 & \text{if } s = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Finally, because of this trivial behaviour of the block variables, the normalized mean-square fluctuation is easy to calculate and becomes

$$\psi^{(s)} = \frac{s}{N}\sum_{i=1}^{N/s}\psi_i^{(s)} = \frac{s}{N}\frac{N}{s}\psi_i^{(s)} = \begin{cases} 1 & \text{if } s = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

### *3.2. Rudin–Shapiro sequence*

The Rudin–Shapiro sequence was originally introduced [20] as an example of a sequence fulfilling the inequality

$$\sup_{\theta\in[0,1)}\left|\sum_{n=0}^{N} V_n e^{in2\pi\theta}\right| \leqslant (2+\sqrt{2})\sqrt{N}. \tag{11}$$

(Recall that a randomly picked sequence most probably has the magnitude $\sqrt{N\log N}$ [23].) It can be determined recursively as

$$r_0 = 1 \tag{12}$$

$$r_{2n} = r_n \tag{13}$$

$$r_{2n+1} = (-1)^n r_n. \tag{14}$$

This definition yields a sequence consisting of positive and negative units directly. (An alternative way of defining the Rudin–Shapiro sequence is to count the parity of the number of occurrences of the pattern 11 in base two of the integer $n$ corresponding to the element $r_n$ [22].) All we have to do to fit it into the formalism developed in the previous section is to change the indices one step according to $V_n = r_{n-1}$, $n = 1, 2, 3, \ldots$. As for the Thue–Morse sequence, it is possible to obtain the sequence from a substitution rule [22]. However, this time we need four different elements for the rule to work properly. The substitution rule is A $\to$ AB, B $\to$ AC, C $\to$ DB, D $\to$ DC. Here we use the letter A as a seed, and obtain the following generations: A, AB, ABAC, ABACABDB, ABACABDBABACDCAC, .... Finally we replace every A and B with $+1$ and every C and D with $-1$.

In order to obtain the explicit expressions for the block variables and their normalized mean-square fluctuations, we start by considering the *substitution matrix*, $M$. It is defined from the relation

$$n^{(k+1)} = Mn^{(k)} \tag{15}$$

where $n^{(k)}$ is a vector whose components are the number of different letters in the sequence when the substitution rule has been applied $k$ times. For the Rudin–Shapiro sequence, the matrix looks like

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}. \tag{16}$$

Since we start with the letter A as a seed, we have $n^{(0)} = (1, 0, 0, 0)^T$. We can easily calculate $n^{(k)}$ by a direct diagonalization of $M$ and obtain after some tedious, but elementary, work for all positive integers $k$

$$n^{(k)} = M^k n^{(0)} = \begin{pmatrix} 2^{k-2} + 2^{(k-5)/2}[\sqrt{2} + 1 + (-1)^k(\sqrt{2} - 1)] \\ 2^{k-2} + 2^{(k-5)/2}[1 - (-1)^k] \\ 2^{k-2} - 2^{(k-5)/2}[1 - (-1)^k] \\ 2^{k-2} - 2^{(k-5)/2}[\sqrt{2} + 1 + (-1)^k(\sqrt{2} - 1)] \end{pmatrix}. \tag{17}$$

From the result above it is easy to obtain the block variables $\sigma_1^{(s)}$ and $\sigma_2^{(s)}$ for the cases when $s$ can be written as two to the power of a positive integer, i.e. $s = 2^k$. If we denote the components of $n^{(k)}$ as $n_A^{(k)}$, $n_B^{(k)}$, $n_C^{(k)}$ and $n_D^{(k)}$, we have

$$\sigma_1^{(s)} = n_A^{(\log_2 s)} + n_B^{(\log_2 s)} - n_C^{(\log_2 s)} - n_D^{(\log_2 s)} = \begin{cases} \sqrt{2s} & \text{if } \log_2 s \text{ is odd} \\ \sqrt{s} & \text{if } \log_2 s \text{ is even.} \end{cases} \tag{18}$$

Now $\sigma_2^{(s)}$ can be obtained as

$$\sigma_2^{(s)} = \sigma_1^{(2s)} - \sigma_1^{(s)} = \begin{cases} \sqrt{2s} - \sqrt{2s} = 0 & \text{if } \log_2 s \text{ is odd} \\ \sqrt{4s} - \sqrt{s} = \sqrt{s} & \text{if } \log_2 s \text{ is even.} \end{cases} \tag{19}$$

Note that this result is valid for all substitutionable generable sequences with a substitution matrix as (16), i.e. the ordering of the letters in the substitution rule is of no importance when we calculate the first two block variables. When we consider $\sigma_i^{(s)}$ for $i \geqslant 3$, we find

$$\sigma_i^{(s)} = \begin{cases} V_i \sigma_1^{(s)} & \text{if } i \text{ is odd} \\ V_i \sigma_2^{(s)} & \text{if } i \text{ is even.} \end{cases} \tag{20}$$

This formula can be obtained from the composition of the sequence. Remember that the $V_i$s only take the values $\pm 1$ and therefore they only change the sign of the block variables.

When we calculate $K$ from (3), we use that when $N$ is a power of two, $\lim_{N\to\infty} N_+/N = \lim_{N\to\infty} N_-/N = \frac{1}{2}$. This result can be obtained either from (17) or from the eigenvector corresponding to the largest eigenvalue of the substitution matrix. (The largest eigenvalue is 2 with corresponding eigenvector $(1, 1, 1, 1)^T$, which shows that for the infinite sequence, all elements are equally common.) This does not contradict the result that the block variables grow as the square-root of the block size, we just have to be careful when we let the size of the system tend to infinity. For all $s$ which divide $N$, we then have $\lim_{N\to\infty} K = 1$. The sum $H_s(N)$, also from (3), can easily be determined in this case. Since we have the same ratio of positive and negative elements in the sequence, this entity disappears when we let $N$ tend to infinity, i.e. $\lim_{N\to\infty} H_s(N) = 0$.

The normalized mean-square fluctuation can now be calculated for all $N$ which are powers of two, $2^k$ say, and all $s$ which are divisors to $N$, as

$$\begin{aligned} \psi^{(s)} &= \frac{s}{N}\sum_{i=1}^{N/s} \psi_i^{(s)} = \frac{s}{KN}\sum_{i=1}^{N/s} (\sigma_i^{(s)} - H_s(N))^2 \\ &= \frac{s}{KN}\sum_{i=1}^{N/(2s)} [(V_{2i-1}\sigma_1^{(s)} - H_s(N))^2 + (V_{2i}\sigma_2^{(s)} - H_s(N))^2] \\ &= \begin{cases} \frac{s}{KN}\frac{N}{2s}[(\pm\sqrt{2s} - H_s(N))^2 + (0 - H_s(N))^2] & \text{if } \log_2 s \text{ is odd} \\ \frac{s}{KN}\frac{N}{2s}[(\pm\sqrt{s} - H_s(N))^2 + (\pm\sqrt{s} - H_s(N))^2] & \text{if } \log_2 s \text{ is even.} \end{cases} \end{aligned} \tag{21}$$

Hence

$$\psi^{(s)} \to s \qquad \text{when } N = 2^k \to \infty. \tag{22}$$

Note that this is the same result as we have when we average over all sequences with the same ratio between the number of positive and negative elements as the Rudin–Shapiro sequence, as stated above.

### *3.3. Period-doubling sequence*

As the other sequences considered here, the period-doubling sequence can be generated by a substitution rule. This time, it looks like A → AB, B → AA, and we still start with A as a seed and attribute to each A in the sequence the value $+1$ and to each B the value $-1$. The substitution matrix for this sequence is

$$M = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}. \tag{23}$$

We can now calculate the block variables in the same way as we did for the Rudin–Shapiro sequence. We calculate $n^{(k)}$ by diagonalization of $M$ and use as an initial value $n^{(0)} = (1, 0)^T$. This yields

$$n^{(k)} = \frac{1}{3}\begin{pmatrix} 2^{k+1} + (-1)^k \\ 2^k - (-1)^k \end{pmatrix}. \tag{24}$$

The first block variable is now obtained as

$$\sigma_1^{(s)} = n_A^{(\log_2 s)} - n_B^{(\log_2 s)} = \tfrac{1}{3}[s + 2(-1)^{\log_2 s}]. \tag{25}$$

The first term on the right-hand side, $s/3$, is exactly the result for a random sequence (or more appropriate, for the average over all sequences) with the same proportion between the number of As and Bs as for the period-doubling sequence. Also the second block variable depends only on the substitution matrix and can be calculated as

$$\sigma_2^{(s)} = \sigma_1^{(2s)} - \sigma_1^{(s)} = \tfrac{1}{3}[s - 4(-1)^{\log_2 s}]. \tag{26}$$

For larger indices, we get no new sequences $\{\sigma_i^{(s)}\}_s$, but alternate between $\{\sigma_1^{(s)}\}_s$ and $\{\sigma_2^{(s)}\}_s$ according to

$$\sigma_i^{(s)} = \begin{cases} \sigma_1^{(s)} & \text{if } V_i = +1 \\ \sigma_2^{(s)} & \text{if } V_i = -1. \end{cases} \tag{27}$$

When we calculate $K$ and $H_s(N)$ from (3), we use the fact that $\lim_{N\to\infty} N_+/N = \frac{2}{3}$ and $\lim_{N\to\infty} N_-/N = \frac{1}{3}$ for all $N$ which are powers of two, $2^k$ say. This can be obtained as for the Rudin–Shapiro sequence. This gives us $\lim_{N\to\infty} K = \frac{8}{9}$ and $\lim_{N\to\infty} H_s(N) = s/3$ for all $s$. Finally for the period-doubling sequence, we have the normalized mean-square fluctuation $\psi^{(s)}$. This is obtained from the normalized block variables

$$\psi_i^{(s)} = \frac{1}{K}[\sigma_i^{(s)} - H_s(N)]^2 \to \begin{cases} \frac{1}{2} & \text{if } V_i = +1 \\ 2 & \text{if } V_i = -1 \end{cases} \qquad \text{when } N = 2^k \to \infty \tag{28}$$

by averaging these values with their relative occurrence in the infinite sequence. This means

$$\psi^{(s)} = \frac{s}{N}\sum_{i=1}^{N/s} \psi_i^{(s)} \to \tfrac{2}{3} \times \tfrac{1}{2} + \tfrac{1}{3} \times 2 = 1 \qquad \text{when } N = 2^k \to \infty. \tag{29}$$

As before, this is for all $s$ which are divisors to $N$.

### *3.4. Fibonacci sequence*

The last sequence we will consider is the Fibonacci sequence. It can e.g. be obtained from the substitution rule A $\rightarrow$ AB, B $\rightarrow$ A, where, as usual, we use the letter A as a seed and put $V_n = +1(-1)$ if site $n$ is an A(B). Let us denote the sequence of Fibonacci numbers as $\{F_k\}$, where $F_1 = F_2 = 1$ and $F_k = F_{k-1} + F_{k-2}$, $k \geqslant 2$. The Fibonacci numbers can also be obtained as

$$F_k = \frac{1}{\sqrt{5}}[\tau^k - (-1)^k \tau^{-k}] \tag{30}$$

where $\tau$ denotes the golden mean $(\sqrt{5}+1)/2$. It is clear that after we have applied the substitution rule $k$ times, we will have $F_{k+2}$ letters in the chain. Of these letters, $F_{k+1}$ will be $A$s and $F_k$ will be $B$s. When $s$ is a Fibonacci number, $s = F_k$ say, the block variable $\sigma_1^{(s)}$ will obviously, for large $s$, behave as

$$\sigma_1^{(s)} = F_{k-1} - F_{k-2} = F_{k-3} \approx \tau^{-3} F_k = (\sqrt{5} - 2)s. \tag{31}$$

The $k$th generation of the Fibonacci sequence, i.e. the finite sequence we have when the substitution rule has been applied $k$ times, can also be obtained from a direct concatenation of the generations $k-1$ and $k-2$. Explicitly, let $w^{(k)}$ denote the $k$th generation of the sequence. It is obtained recursively as

$$w^{(k)} = w^{(k-1)} w^{(k-2)} \tag{32}$$

with $w^{(-1)} = B$ and $w^{(0)} = A$. As a consequence of this, it does not matter for the large behaviour where in the sequence we start the summation in (1). This implies that all other block variables, $\sigma_i^{(s)}$, $i \geqslant 2$, will behave in the same manner as $\sigma_1^{(s)}$ when $s$ is increasing. This has also been numerically checked to be true. As for the period-doubling sequence, this is exactly the behaviour we would get for a purely random sequence with the same proportion between the elements.

The normalized mean-square fluctuation, $\psi^{(s)}$, has very irregular behaviour, but numerical calculations indicate that the values are almost uniformly spread between zero and one (or slightly above). As far as the present author knows, there seems to be no other way to obtain these values but by directly performing the sum in (4).

## 4. Conclusions and outlook

In the previous section, we have seen how the behaviour of the block variables and their mean-square fluctuation vary very widely, depending on which sequence we consider.

None of the sequences treated above is of course truly random. After all, they could all be obtained from rather short substitution rules. Nevertheless, they show different degrees of 'randomness' dependent on which property we concentrate.

The block variables for the Thue–Morse sequence behave exactly as for a periodic sequence with period two (when $i \geqslant 2$ in (1)). This might look compelling when we notice that the electron wavefunction for a physical system generated by the Thue–Morse sequence can be Bloch-like [24]. On the other hand, both the Fibonacci sequence and the period-doubling sequence have block variables which behave as for random systems, despite both of them having a singular continuous spectrum for the same physical model as we mentioned for Thue–Morse.

The Rudin–Shapiro sequence has an absolutely continuous Fourier transform, just as a random sequence has (for the Rudin–Shapiro sequence it is actually a constant [14]), which seems to fit in well with the observation given above that the normalized mean-square fluctuation of the block variables behaves exactly as for a random sequence. On the other hand, the block variables themselves do not behave as for a random sequence, which should imply that the Rudin–Shapiro sequence is not random at all.

Furthermore, the Fibonacci sequence is quasiperiodic, while the period-doubling sequence is not. Nevertheless, their block variables show qualitatively the same behaviour. The Thue–Morse sequence has a singular continuous Fourier spectrum [25], but the block variables are, as mentioned above, identical to those of a periodic system with period two.

The results obtained above are, in some sense, only partial; we have only obtained results for some specific choices of deterministic aperiodic sequences. For instance, one can ask whether it possible to obtain a general result for fixed points of substitutions of constant length, or even for substitutions of non-constant length. Nevertheless, it is still possible to draw some conclusions from the cases treated in the present paper. The conclusion from this very brief survey must be that the concept of block variables is of limited value when we try to distinguish random sequences from the deterministic aperiodic sequences treated in this paper. Nevertheless, the method has proved to be useful when considering the protein folding problem [18], and therefore it might be interesting to study theoretically the properties of a protein constructed according to the sequences we have considered here.

## Acknowledgments

The author would like to thank Rolf Riklund for valuable discussions about physics in general, and aperiodic systems in particular. He would also like to thank the referees for valuable help in clarifying the article.

## References

[1] Janner A, Janssen T and de Wolff P M 1982 *Europhys. News* **13** 1 and references therein
[2] Shechtman D, Blech I, Gratias D and Cahn J W 1984 *Phys. Rev. Lett.* **53** 1951
[3] Merlin R, Bajema K, Clarke R, Juang F-Y and Bhattacharya P K 1985 *Phys. Rev. Lett.* **55** 1768
[4] Dunlap D H, Wu H-L and Phillips P W 1990 *Phys. Rev. Lett.* **65** 88
[5] Sil S, Karmakar S N and Moitra R K 1993 *Phys. Lett.* A **183** 344
[6] Grimm U and Baake M 1994 *J. Stat. Phys.* **74** 1233
[7] Hörnquist M and Riklund R 1996 *J. Phys. Soc. Japan* **65** 2872
Hörnquist M and Riklund R 1997 *Phys. Rev.* B **55** 875
[8] Shannon C E and Weaver W 1949 *The Mathematical Theory of Communication* (Urbana, IL: University of Illinois Press)
[9] Collet P and Eckmann J P 1980 *Iterated Maps on the Interval as Dynamical Systems* (Boston, MA: Birkhäuser)
Hao B-L 1989 *Elementary Symbolic Dynamics and Chaos in Dissipative Systems* (Singapore: World Scientific)
[10] Chechetkin V R and Yu Turygin A 1994 *J. Phys. A: Math. Gen.* **27** 4875
[11] Burrows B L and Sulston K W 1991 *J. Phys. A: Math. Gen.* **24** 3979
Berthé V 1994 *J. Phys. A: Math. Gen.* **27** 7993
Sulston K W, Burrows B L and Chishti A N 1995 *Physica* **217A** 146
[12] Levitov L S 1988 *Europhys. Lett.* **7** 343
[13] Brenner N and Fishman S 1992 *Nonlinearity* **4** 211
Knuth D E 1981 *The Art of Computer Programming* 2nd edn (Reading, MA: Addison-Wesley)
[14] Luck J M 1989 *Phys. Rev.* B **39** 5834
[15] Godrèche C and Luck J M 1990 *J. Phys. A: Math. Gen.* **23** 3769
[16] Liviotti E 1996 *J. Phys.: Condens. Matter* **8** 5007
[17] Parisi G 1988 *Statistical Field Theory (Frontiers in Physics 66)* (Redwood City, CA: Addison-Wesley)
[18] Irbäck A, Peterson C and Pothast F 1996 *Proc. Natl Acad. Sci., USA* **93** 9533
[19] Thue A 1906 *Kra. Vidensk. Selsk. Skrifter. I. Mat. Nat. Kl.* No 7 (Christiania)
Thue A 1912 *Kra. Vidensk. Selsk. Skrifter. I. Mat. Nat. Kl.* No 1 (reprinted 1977 *Selected Mathematical Papers of Axel Thue* (Oslo: Universitetsforlaget) p 139 resp. p 413

Morse H M 1921 *Trans. Am. Math. Soc.* **22** 48
[20] Shapiro H S 1951 Extremal problems for polynomials and power series *Master's Thesis* Massachusetts Institute of Technology
Rudin W 1959 *Proc. Am. Math. Soc.* **10** 855
[21] Ostlund S, Pandit R, Rand D, Schellnhuber H J, and Siggia E D 1983 *Phys. Rev. Lett.* **50** 1873
Ostlund S and Pandit R 1984 *Phys. Rev.* B **29** 1394
Kohmoto M, Kadanoff L P, and Tang C 1983 *Phys. Rev. Lett.* **50** 1870
Kohmoto M and Oono Y 1984 *Phys. Lett.* **A102** 145
[22] Queffélec M 1987 *Substitution Dynamical Systems—Spectral Analysis (Springer Lecture Notes in Mathematics 1294)* (Berlin: Springer)
[23] Salem R and Zygmund A 1954 *Acta Math.* **91** 245
[24] Riklund R, Severin M and Liu Y 1987 *Int. J. Mod. Phys.* B **1** 121
[25] Mahler K 1927 *J. Math. Phys.* **6** 158