P. Arnoux, V. Berthé, S. Ferenczi, S. Ito,
C. Mauduit, M. Mori, J. Peyrière, A. Siegel,
J.- I. Tamura, Z.- Y. Wen

Introduction to Finite Automata and Substitution Dynamical Systems

Editors: V. Berthé, S. Ferenczi, C. Mauduit, A. Siegel

Preface

There are two basic ways of constructing dynamical systems. One approach is to take an already existing system from the vast reserve arising in biology, physics, geometry, or probability; such systems are typically rather complex and equipped with rigid structures (a "natural" system is generally a "smooth" system). Alternatively one can build a system by hand using some basic tools like strings of letters on a finite alphabet; the latter system will have a simple structure. Nevertheless, these "simple" systems prove to be very useful in many mathematical fields (number theory, harmonic analysis, combinatorics, ergodic theory, and so on), as well as in theoretical computer science and physics. This category also includes various "classical" systems.

To make more precise this intuitive concept of "simple" systems, we can use the combinatorial notion of complexity of a sequence of letters with values in a finite alphabet, which counts the number of factors of given length. This gives an indication of the degree of randomness of the sequence: a periodic sequence has a bounded complexity, while the g-adic expansion of a normal number in base g has an exponential complexity. There are many examples of sequences having a reasonably low complexity function, the most famous being automatic sequences and Sturmian sequences, and the "reasonably simple" dynamical systems we like to consider are those which are canonically associated to these kind of sequences.

The overall structure of this book reflects our purpose which is twofold. We first want to provide an introduction to the theory of finite automata and discrete dynamical systems by focusing on several topics including various aspects of mathematics (as for instance geometry, combinatorics, ergodic theory and spectral analysis, number theory, numeration systems, fractals and tilings) but also computer science and theoretical physics, with the structural properties of quasicrystals. Secondly, we want to give a state of the art on this field, spotlighting representative aspects of the theory. The early chapters include basic introductory material. Then we focus on the following themes:

An introduction to elementary properties of **combinatorics on words** is given in the first four chapters. In particular Chap. 2 and Chap. 4 provide an introduction to automatic sequences, which are produced by very natural algorithms coming from theoretical computer science. Chap. 6 and Chap. 9 give an analysis of the combinatorial properties of Sturmian words: these are the sequences (or infinite words) which have the smallest complexity function among non-ultimately periodic sequences.

Tools from **ergodic theory and spectral analysis** are introduced in Chap. 3, via a detailed study of systems associated with sequences having a low complexity function. An elementary introduction to correlation properties, and some examples of computations of correlation measures, are given in Chap. 2. Chapter 7 surveys the latest results in the spectral study of substitutive dynamical systems. Chapter 11 gives a special account of the ergodic properties of the Perron–Frobenius transfer operator.

Numeration systems appear in a natural way in the study of low complexity sequences. Beatty sequences and in particular the Fibonacci numeration system are introduced in Chap. 2 and dealt with in detail in Chap. 5. Chapter 6 introduces Ostrowski's numeration system, while the numeration systems defined in Chap. 8 are linked to the geometric properties of Pisot substitutive sequences.

The question of the **geometric representation** of substitutive sequences and more generally of low complexity sequences has given birth to a great amount of work. We give an account of the development and current state of this problem in Chap. 7. The study of the Sturmian case is particularly instructive (Chap. 6); it provides a well-known and fundamental interaction between ergodic theory, number theory, and symbolic dynamics, which comes from the study of irrational rotations on the one-dimensional torus $\mathbb{T}^1$. To an irrational real number α we associate a geometric dynamical system, the rotation $R : x \mapsto x + \alpha \bmod 1$, an arithmetic algorithm, the usual continued fraction approximation, and a set of Sturmian sequences which are codings of trajectories under R by a canonical partition; the continued fraction algorithm arises naturally as the link between the dynamical system and the symbolic sequences, and the study of the arithmetic and symbolic objects is very useful for the study of the dynamical system. Chapter 2 and 5 allude to Beatty sequences, whereas Chap. 9 studies in detail the connections between Sturmian and invertible substitutions over a two-letter alphabet. Chapter 8 gives a detailed introduction to the tools and technique used in this framework, in order to generalize this interaction to further systems and sequences.

The notion of **self-similarity** is illustrated by symbolic objects, namely substitutions, and by geometric objects, the fractal sets. Chapter 7 and 8 give examples of interactions between these two notions, as the most natural geometric representations of a substitutive sequence are sets with fractal boundaries. See also Chap. 11 for the study of fractal sets associated in a natural way with piecewise linear transformations of the unit interval.

An interesting field of application of automatic sequences deals with the **transcendence** of formal power series. Indeed, the study of automata provides a very fruitful transcendence criterion (Christol's theorem) for formal power series with coefficients in a finite field, which can be considered as a natural translation into algebraic terms of the combinatorial properties of automatic sequences. Chapter 4 presents this criterion and surveys the most recent transcendence results obtained via automata theory. Note that a real number whose g-adic expansion is an automatic sequence is conjectured to be either transcendental or rational, and has been proved to be transcendental when the sequence is Sturmian. Hence this presentation emphasizes the following philosophy: algebraicity strongly depends on the (generalized) "base" in which one works. Some connected results of transcendence are studied in Chap. 5; Chapter 8 reviews some Diophantine approximation properties issued from the study of substitutive systems.

We illustrate the fascinating connections between **physics** and low complexity sequences through the study of trace maps in Chap. 8 and 9. Indeed free monoids, groups and their morphisms occur in a natural way in physics: finite automata and substitutive sequences are very useful to model and describe certain situations in solid state physics. In particular, one important question in quasicrystal theory is to compute the traces of matrices defined inductively according to a substitutive process. Trace maps are effective algorithms for constructing the recursion relations that the traces satisfy.

The book is organized as follows. The introductory chapter unifies the notation and contains the necessary background for later chapters. Indeed we introduce in this chapter the most fundamental notions we will use throughout the book: words, languages, complexity function, substitutions, finite automata, and so on.

The aim of the second chapter is to focus on the notions of finite automata and substitutions by showing some typical examples of situations in which they occur in a quite natural way. For example, the Morse sequence was first introduced to answer a question in combinatorial number theory and rediscovered by many people in various other circumstances (including geometry, group theory, logic). The Rudin-Shapiro sequence, first introduced to answer a question in harmonic analysis asked by Salem, is nowadays a basic construction in number theory and ergodic theory. The Fibonacci sequence, introduced as a natural example of a generator of a symbolic dynamical system, is deeply connected with the continued fraction algorithm and gives rise to many applications in theoretical computer science (for example to obtain good algorithms for the drawing of a straight line on a computer screen).

In the third chapter, we introduce the fundamental notions of ergodic theory, also through the study of a few examples of symbolic dynamical systems, both in the topological and measure-theoretic framework. These include substitutions, which are maps on a finite alphabet, and which define naturally a class of infinite sequences, and the shift on this set. We study four examples of substitutions (Morse, Rudin-Shapiro, Fibonacci, already introduced in Chap. 2, and Chacon) and use them, together with symbolic notions (language, frequencies, complexity) to define basic ergodic notions (minimality, ergodicity), and begin the study of elementary spectral properties, geometric representation of symbolic systems, and the vast problem of joinings between systems.

The aim of Chap. 4 is to investigate the connections between automata theory and transcendence in fields of positive characteristic. We first recall some definitions concerning the transcendence of formal power series and then state the main theorem of transcendence of this chapter (Christol's theorem). We also consider some generalizations of this theorem, first to the multidimensional case and then, to general fields of positive characteristic. We then introduce some functions defined by Carlitz (exponential, logarithm, zeta) which are analogous to the corresponding real functions. We review the results and methods of transcendence concerning these functions and emphasize the results involving automata.

Chapter 5 is devoted to various partitions of the set of positive integers: Beatty sequences and connections with Sturmian sequences, partitions generated by substitutions, and similis partitions illustrated by linguistic properties of the Hungarian and Japanese languages. Special attention is devoted to non-periodic words which are shown to be fixed points of some combinatorial processes: above all the notion of log-fixed points is introduced. This study is illustrated by the Kolakoski word, which is shown to be not only a log-fixed point but also the unique fixed point of several maps: a map based on the Minkowski question-mark map, or maps defined by using the continued fraction expansion, the base-3 expansion, and so on. We then generalize these situations and present some open problems.

The main idea of Chap. 6 is to show how it is possible to recover all the classical properties of rotations and continued fractions in a purely combinatorial way; we only use the combinatorial definition of Sturmian sequences, and obtain the existence of a combinatorial continued fraction acting on the set of Sturmian sequences. Some proofs, and also some geometric interpretations, become simpler and more natural in this setting; it also suggests non trivial ways to generalize the usual continued fraction.

Chapter 7 presents an overview of the general spectral theory of substitutive dynamical systems. After recalling the tools and concepts required (including properties of translations on abelian compact groups, subshifts of finite type and adic systems, maximal equicontinuous factors, the notions of recognizability, and Markov partitions), a complete description of the related literature is given, including very recent work and the main conjectures in this subject.

The problem of the geometric representation of substitutive dynamical systems is studied in detail in Chap. 8. Some basic tools and notions are introduced: stepped surface, generation by generalized substitutions, renormalization, study of the fractal boundary, and so on. Special attention is devoted to some important applications, as the quasi-periodicity of the tiling related to the stepped surface, the existence of Markov partitions of group automorphisms on $\mathbb{T}^3$ or Diophantine approximation properties in connection with the modified Jacobi-Perron algorithm.

The purpose of Chap. 9 is to study the properties of factors of sequences generated by invertible substitutions over a two-letter alphabet. We first discuss the local isomorphism between two invertible substitutions, reducing the problem to the study of some special classes of invertible substitutions. We then study some elementary properties of factors (including palindrome properties, powers of factors), which generalize some classical results for the Fibonacci sequence. By introducing the notion of singular factors, we establish a decomposition of the fixed points of invertible substitutions and we discuss the factor properties associated with this decomposition. As a simple application, we easily determine the special factors of a fixed point of an invertible substitution. Some other applications, such as trace maps are also discussed.

Chapter 10 investigates more deeply trace maps. Trace maps are dynamical systems attached to automorphisms of the free group of two generators, which occur in a wide range of physical situations. By exploiting polynomial identities in rings of matrices, recursion formulas are produced between the traces of matrices defined by an induction using substitutions. We then study from an algebraic point of view endomorphisms of free monoids and free groups. Such endomorphisms are shown to induce a trace map of the ring of Fricke characters. Particular emphasis is given to the group structure of trace maps and the Fricke-Vogt invariant.

Chapter 11 deals with Cantor sets generated by expanding piecewise linear maps. The main tool is the α-Fredholm matrix. This is the extension of the Fredholm matrix which is introduced to study the spectral properties of the Perron–Frobenius operator associated with one-dimensional maps. Using this α-Fredholm matrix, the Hausdorff dimension of the Cantor sets we consider is studied, as well as the ergodic properties of the dynamical system on it.

This book ends with a survey in Chap. 12 of some of the big open problems in the subject. The first is the S-adic conjecture, about the equivalence for a minimal sequence of having sub-linear complexity and being generated by a finite number of substitutions. Then we look at possible generalizations of the interaction between rotations and Sturmian sequences through the usual continued fraction algorithm; these involve Arnoux-Rauzy sequences, interval exchanges and codings of two rotations. This approach induces various open questions about geometric representations of substitutions, arithmetics in $SL(3,\mathbb{Z})$ and $SL(3,\mathbb{N})$, and about the definition in the two-dimensional case of some fundamental combinatorial objects as the complexity function or the notion of substitution.

The idea for this book stemmed from the collaboration between various groups of mathematicians mainly from France, Japan and China. It is largely based on courses performed by the authors in several universities and given during several Summer Schools in the past five years. The reader is not assumed to have a detailed knowledge in each of the fields covered by this book; we rather try to provide the necessary information, allowing it to be be used by undergraduates and graduate students alike.

We give hereafter a list of contributors with their addresses and their corresponding chapter.

Acknowledgements. We would like to express our gratitude to Jean-Paul Allouche, Guy Melançon, Luca Zamboni and Michel Koskas for their careful reading, and for their numerous suggestions, including pointing out mathematical and stylistic Al errors.

We are especially grateful to Madelena Cropley-Gonzalez for helping us to refine our use of the English language. We are also greatly indebted to Éric Lozingot for his contribution to the technical tasks related to the bibliography.

The editors

V. Berthé S. Ferenczi C. Mauduit A. Siegel

List of contributors [1]

P. Arnoux (Chapter 6, 12)
Institut de Math. de Luminy
Case 907, 163 avenue de Luminy
F-13288 Marseille cedex 09
FRANCE
arnoux@iml.univ-mrs.fr

V. Berthé (Chapter 1, 4, 12)
Institut de Math. de Luminy
Case 907, 163 avenue de Luminy
F-13288 Marseille cedex 09
FRANCE
berthe@iml.univ-mrs.fr

S. Ferenczi (Chapter 3, 12)
Institut de Math. de Luminy
Case 907, 163 avenue de Luminy
F-13288 Marseille cedex 09
FRANCE
ferenczi@iml.univ-mrs.fr

S. Ito (Chapter 8)
Department of Mathematics
Tsuda College
Tsuda Machi
Kodaira, Tokyo 187
JAPON
ito@tsuda.ac.jp

C. Mauduit (Chapter 2)
Institut de Math. de Luminy
Case 907, 163 avenue de Luminy
F-13288 Marseille cedex 09
FRANCE
mauduit@iml.univ-mrs.fr

M. Mori (Chapter 11)
Department of Mathematics
College of Humanities and Sciences
Nihon University
3-25-40 Sakura Jyousui
Setagayaku Tokyo
JAPAN 156-8550
mori@math.chs.nihon-u.ac.jp,
mori@tsuda.ac.jp

J. Peyrière (Chapter 10)
Université Paris-Sud
Mathématiques, bât. 425
91405 Orsay Cedex
FRANCE
Jacques.PEYRIERE@math.u-psud.fr

A. Siegel (Chapter 7)
Institut de Math. de Luminy
Case 907, 163 avenue de Luminy
F-13288 Marseille cedex 09
FRANCE
siegel@iml.univ-mrs.fr

J.- I. Tamura (Chapter 5)
3-3-7-307 Azamino
Aoba-ku Yokohama
225-0011 JAPAN
tamura@rkmath.rikkyo.ac.jp

Z.- Y. Wen (Chapter 9)
Tsinghua University
Department of Mathematics
Beijing 100084
P. R CHINA
wenzy@tsinghua.edu.cn

[1] Some chapters have been modified by V. Berthé and A. Siegel so as to make them more reader-friendly and to fit in more harmoniously into the study as a whole.

Contents