



HAL
open science

Bifix codes and interval exchanges

Valérie Berthé, Clelia de Felice, Francesco Dolce, Julien Leroy, Dominique Perrin, Christophe Reutenauer, Giuseppina Rindone

► **To cite this version:**

Valérie Berthé, Clelia de Felice, Francesco Dolce, Julien Leroy, Dominique Perrin, et al.. Bifix codes and interval exchanges. *Journal of Pure and Applied Algebra*, 2015, 10.1016/j.jpaa.2014.09.028 . hal-01367685

HAL Id: hal-01367685

<https://hal.science/hal-01367685v1>

Submitted on 16 Sep 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Bifix codes and interval exchanges

Valérie Berthé¹, Clelia De Felice², Francesco Dolce³, Julien Leroy⁴,
Dominique Perrin³, Christophe Reutenauer⁵, Giuseppina Rindone³

¹CNRS, Université Paris 7, ²Università degli Studi di Salerno,
³Université Paris Est, LIGM, ⁴Université du Luxembourg,
⁵Université du Québec à Montréal

July 22, 2014 18 h 45

Abstract

We investigate the relation between bifix codes and interval exchange transformations. We prove that the class of natural codings of regular interval exchange transformations is closed under maximal bifix decoding.

Contents

1	Introduction	2
2	Interval exchange transformations	3
2.1	Regular interval exchange transformations	4
2.2	Natural coding	5
2.3	Uniformly recurrent sets	6
2.4	Interval exchange sets	7
3	Bifix codes and interval exchange	9
3.1	Prefix codes and bifix codes	9
3.2	Maximal bifix codes	10
3.3	Bifix codes and regular transformations	13
4	Tree sets	16
4.1	Extension graphs	16
4.2	Interval exchange sets and planar tree sets	17
4.3	Exchange of pieces	18
4.4	Subgroups of finite index	21

24 1 Introduction

25 This paper is part of a research initiated in [2] which studies the connections
26 between the three subjects formed by symbolic dynamics, the theory of codes
27 and combinatorial group theory. The initial focus was placed on the classical
28 case of Sturmian systems and progressively extended to more general cases.

29 The starting point of the present research is the observation that the family
30 of Sturmian sets is not closed under decoding by a maximal bifix code, even in
31 the more simple case of the code formed of all words of fixed length n . Actually,
32 the decoding of the Fibonacci word (which corresponds to a rotation of angle
33 $\alpha = (3 - \sqrt{5})/2$) by blocks of length n is an interval exchange transformation
34 corresponding to a rotation of angle $n\alpha$ coded on $n + 1$ intervals. This has lead
35 us to consider the set of factors of interval exchange transformations, called
36 interval exchange sets. Interval exchange transformations were introduced by
37 Oseledec [15] following an earlier idea of Arnold [1]. These transformations form
38 a generalization of rotations of the circle.

39 The main result in this paper is that the family of regular interval exchange
40 sets is closed under decoding by a maximal bifix code (Theorem 3.13). This
41 result invited us to try to extend to regular interval exchange transformations
42 the results relating bifix codes and Sturmian words. This lead us to generalize
43 in [5] to a large class of sets the main result of [2], namely the Finite Index Basis
44 Theorem relating maximal bifix codes and bases of subgroups of finite index of
45 the free group.

46 Theorem 3.13 reveals a close connection between maximal bifix codes and
47 interval exchange transformations. Indeed, given an interval exchange trans-
48 formation T each maximal bifix code X defines a new interval exchange trans-
49 formation T_X . We show at the end of the paper, using the Finite Index Basis
50 Theorem, that this transformation is actually an interval exchange transforma-
51 tion on a stack, as defined in [7] (see also [19]).

52 The paper is organized as follows.

53 In Section 2, we recall some notions concerning interval exchange transfor-
54 mations. We state the result of Keane [12] which proves that regularity is a
55 sufficient condition for the minimality of such a transformation (Theorem 2.3).

56 We study in Section 3 the relation between interval exchange transforma-
57 tions and bifix codes. We prove that the transformation associated with a finite
58 S -maximal bifix code is an interval exchange transformation (Proposition 3.8).
59 We also prove a result concerning the regularity of this transformation (Theo-
60 rem 3.12).

61 We discuss the relation with bifix codes and we show that the class of regular
62 interval exchange sets is closed under decoding by a maximal bifix code, that
63 is, under inverse images by coding morphisms of finite maximal bifix codes
64 (Theorem 3.13).

65 In Section 4 we introduce tree sets and planar tree sets. We show, reformu-
66 lating a theorem of [9], that uniformly recurrent planar tree sets are the regular
67 interval exchange sets (Theorem 4.3). We show in another paper [4] that, in
68 the same way as regular interval exchange sets, the class of uniformly recurrent

69 tree sets is closed under maximal bifix decoding.

70 In Section 4.3, we explore a new direction, extending the results of this
71 paper to a more general case. We introduce exchange of pieces, a notable
72 example being given by the Rauzy fractal. We indicate how the decoding of the
73 natural codings of exchange of pieces by maximal bifix codes are again natural
74 codings of exchange of pieces. We finally give in Section 4.4 an alternative
75 proof of Theorem 3.13 using a skew product of a regular interval exchange
76 transformation with a finite permutation group.

77 **Acknowledgements** This work was supported by grants from Région Ile-de-
78 France, the ANR projects Eqinocs and Dyna3S, the Labex Bezout, the FARB
79 Project “Aspetti algebrici e computazionali nella teoria dei codici, degli automi
80 e dei linguaggi formali” (University of Salerno, 2013) and the MIUR PRIN 2010-
81 2011 grant “Automata and Formal Languages: Mathematical and Applicative
82 Aspects”. We thank the referee for his useful remarks on the first version of the
83 paper which lead us to separate it in two parts.

84 2 Interval exchange transformations

85 Let us recall the definition of an interval exchange transformation (see [8] or [6]).

86 A *semi-interval* is a nonempty subset of the real line of the form $[\alpha, \beta[=$
87 $\{z \in \mathbb{R} \mid \alpha \leq z < \beta\}$. Thus it is a left-closed and right-open interval. For two
88 semi-intervals Δ, Γ , we denote $\Delta < \Gamma$ if $x < y$ for any $x \in \Delta$ and $y \in \Gamma$.

89 Let $(A, <)$ be an ordered set. A partition $(I_a)_{a \in A}$ of $[0, 1[$ in semi-intervals
90 is *ordered* if $a < b$ implies $I_a < I_b$.

91 Let A be a finite set ordered by two total orders $<_1$ and $<_2$. Let $(I_a)_{a \in A}$
92 be a partition of $[0, 1[$ in semi-intervals ordered for $<_1$. Let λ_a be the length of
93 I_a . Let $\mu_a = \sum_{b \leq_1 a} \lambda_b$ and $\nu_a = \sum_{b \leq_2 a} \lambda_b$. Set $\alpha_a = \nu_a - \mu_a$. The *interval*
94 *exchange transformation* relative to $(I_a)_{a \in A}$ is the map $T : [0, 1[\rightarrow [0, 1[$ defined
95 by

$$T(z) = z + \alpha_a \quad \text{if } z \in I_a.$$

96 Observe that the restriction of T to I_a is a translation onto $J_a = T(I_a)$, that
97 μ_a is the right boundary of I_a and that ν_a is the right boundary of J_a . We
98 additionally denote by γ_a the left boundary of I_a and by δ_a the left boundary
99 of J_a . Thus

$$I_a = [\gamma_a, \mu_a[, \quad J_a = [\delta_a, \nu_a[.$$

100 Note that $a <_2 b$ implies $\nu_a < \nu_b$ and thus $J_a < J_b$. This shows that
101 the family $(J_a)_{a \in A}$ is a partition of $[0, 1[$ ordered for $<_2$. In particular, the
102 transformation T defines a bijection from $[0, 1[$ onto itself.

103 An interval exchange transformation relative to $(I_a)_{a \in A}$ is also said to be
104 on the alphabet A . The values $(\alpha_a)_{a \in A}$ are called the *translation values* of the
105 transformation T .

106 **Example 2.1** Let R be the interval exchange transformation corresponding to
 107 $A = \{a, b\}$, $a <_1 b$, $b <_2 a$, $I_a = [0, 1 - \alpha[$, $I_b = [1 - \alpha, 1[$. The transformation R is
 108 the rotation of angle α on the semi-interval $[0, 1[$ defined by $R(z) = z + \alpha \bmod 1$.

109 Since $<_1$ and $<_2$ are total orders, there exists a unique permutation π of A such
 110 that $a <_1 b$ if and only if $\pi(a) <_2 \pi(b)$. Conversely, $<_2$ is determined by $<_1$
 111 and π and $<_1$ is determined by $<_2$ and π . The permutation π is said to be
 112 *associated* with T .

113 If we set $A = \{a_1, a_2, \dots, a_s\}$ with $a_1 <_1 a_2 <_1 \dots <_1 a_s$, the pair (λ, π)
 114 formed by the family $\lambda = (\lambda_a)_{a \in A}$ and the permutation π determines the map
 115 T . We will also denote T as $T_{\lambda, \pi}$. The transformation T is also said to be an
 116 s -interval exchange transformation.

117 It is easy to verify that if T is an interval exchange transformation, then T^n
 118 is also an interval exchange transformation for any $n \in \mathbb{Z}$.

119 **Example 2.2** A 3-interval exchange transformation is represented in Figure 3.2.
 120 One has $A = \{a, b, c\}$ with $a <_1 b <_1 c$ and $b <_2 c <_2 a$. The associated permu-
 tation is the cycle $\pi = (abc)$.

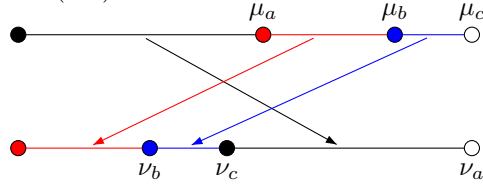


Figure 2.1: A 3-interval exchange transformation

121

122 2.1 Regular interval exchange transformations

123 The *orbit* of a point $z \in [0, 1[$ is the set $\{T^n(z) \mid n \in \mathbb{Z}\}$. The transformation T
 124 is said to be *minimal* if, for any $z \in [0, 1[$, the orbit of z is dense in $[0, 1[$.

125 Set $A = \{a_1, a_2, \dots, a_s\}$ with $a_1 <_1 a_2 <_1 \dots <_1 a_s$, $\mu_i = \mu_{a_i}$ and $\delta_i =$
 126 δ_{a_i} . The points $0, \mu_1, \dots, \mu_{s-1}$ form the set of *separation points* of T , denoted
 127 $\text{Sep}(T)$. Note that the singular points of the transformation T (that is the points
 128 $z \in [0, 1[$ at which T is not continuous) are among the separation points but
 129 that the converse is not true in general (see Example 3.9).

130 An interval exchange transformation $T_{\lambda, \pi}$ is called *regular* if the orbits of
 131 the nonzero separation points μ_1, \dots, μ_{s-1} are infinite and disjoint. Note that
 132 the orbit of 0 cannot be disjoint of the others since one has $T(\mu_i) = 0$ for some
 133 i with $1 \leq i \leq s - 1$. The term regular was introduced by Rauzy in [17]. A
 134 regular interval exchange transformation is also said to be *without connections*
 135 or satisfying the *idoc* condition (where idoc stands for infinite disjoint orbit
 136 condition).

137 Note that since $\delta_2 = T(\mu_1), \dots, \delta_s = T(\mu_{s-1})$, T is regular if and only if the
 138 orbits of $\delta_2, \dots, \delta_s$ are infinite and disjoint.

139 As an example, the 2-interval exchange transformation of Example 2.1 which
 140 is the rotation of angle α is regular if and only if α is irrational.

141 Note that if T is a regular s -interval exchange transformation, then for any
 142 $n \geq 1$, the transformation T^n is an $n(s-1)+1$ -interval exchange transformation.
 143 Indeed, the points $T^i(\mu_j)$ for $0 \leq i \leq n-1$ and $1 \leq j \leq s-1$ are distinct and
 144 define a partition in $n(s-1)+1$ intervals.

145 The following result is due to Keane [12].

146 **Theorem 2.3 (Keane)** *A regular interval exchange transformation is mini-*
 147 *mal.*

148 The converse is not true. Indeed, consider the rotation of angle α with α
 149 irrational, as a 3-interval exchange transformation with $\lambda = (1-2\alpha, \alpha, \alpha)$ and
 150 $\pi = (132)$. The transformation is minimal as any rotation of irrational angle
 151 but it is not regular since $\mu_1 = 1-2\alpha$, $\mu_2 = 1-\alpha$ and thus $\mu_2 = T(\mu_1)$.

152 The following necessary condition for minimality of an interval exchange
 153 transformation is useful. A permutation π of an ordered set A is called *de-*
 154 *composable* if there exists an element $b \in A$ such that the set B of elements
 155 strictly less than b is nonempty and such that $\pi(B) = B$. Otherwise it is called
 156 *indecomposable*. If an interval exchange transformation $T = T_{\lambda, \pi}$ is minimal,
 157 the permutation π is indecomposable. Indeed, if B is a set as above, the set
 158 $S = \cup_{a \in B} I_a$ is closed under T and strictly included in $[0, 1[$.

159 The following example shows that the indecomposability of π is not sufficient
 160 for T to be minimal.

161 **Example 2.4** Let $A = \{a, b, c\}$ and λ be such that $\lambda_a = \lambda_c$. Let π be the
 162 transposition (ac) . Then π is indecomposable but $T_{\lambda, \pi}$ is not minimal since it
 163 is the identity on I_b .

164 2.2 Natural coding

165 Let A be a finite nonempty alphabet. All words considered below, unless stated
 166 explicitly, are supposed to be on the alphabet A . We denote by A^* the set of
 167 all words on A . We denote by 1 or by ε the empty word. We refer to [3] for the
 168 notions of prefix, suffix, factor of a word.

169 Let T be an interval exchange transformation relative to $(I_a)_{a \in A}$. For a
 170 given real number $z \in [0, 1[$, the *natural coding* of T relative to z is the infinite
 171 word $\Sigma_T(z) = a_0 a_1 \dots$ on the alphabet A defined by

$$a_n = a \quad \text{if} \quad T^n(z) \in I_a.$$

172 For a word $w = b_0 b_1 \dots b_{m-1}$, let I_w be the set

$$I_w = I_{b_0} \cap T^{-1}(I_{b_1}) \cap \dots \cap T^{-m+1}(I_{b_{m-1}}). \quad (2.1)$$

173 Note that each I_w is a semi-interval. Indeed, this is true if w is a letter. Next,
 174 assume that I_w is a semi-interval. Then for any $a \in A$, $T(I_{aw}) = T(I_a) \cap I_w$ is a

175 semi-interval since $T(I_a)$ is a semi-interval by definition of an interval exchange
 176 transformation. Since $I_{aw} \subset I_a$, $T(I_{aw})$ is a translate of I_{aw} , which is therefore
 177 also a semi-interval. This proves the property by induction on the length.

178 Set $J_w = T^m(I_w)$. Thus

$$J_w = T^m(I_{b_0}) \cap T^{m-1}(I_{b_1}) \cap \dots \cap T(I_{b_{m-1}}). \quad (2.2)$$

179 In particular, we have $J_a = T(I_a)$ for $a \in A$. Note that each J_w is a semi-
 180 interval. Indeed, this is true if w is a letter. Next, for any $a \in A$, we have
 181 $T^{-1}(J_{wa}) = J_w \cap I_a$. This implies as above that J_{wa} is a semi-interval and
 182 proves the property by induction. We set by convention $I_\varepsilon = J_\varepsilon = [0, 1[$. Then
 183 one has for any $n \geq 0$

$$a_n a_{n+1} \cdots a_{n+m-1} = w \iff T^n(z) \in I_w \quad (2.3)$$

184 and

$$a_{n-m} a_{n-m+1} \cdots a_{n-1} = w \iff T^n(z) \in J_w. \quad (2.4)$$

185 Let $(\alpha_a)_{a \in A}$ be the translation values of T . Note that for any word w ,

$$J_w = I_w + \alpha_w \quad (2.5)$$

186 with $\alpha_w = \sum_{j=0}^{m-1} \alpha_{b_j}$ as one may verify by induction on $|w| = m$. Indeed
 187 it is true for $m = 1$. For $m \geq 2$, set $w = ua$ with $a = b_{m-1}$. One has
 188 $T^m(I_w) = T^{m-1}(I_w) + \alpha_a$ and $T^{m-1}(I_w) = I_w + \alpha_u$ by the induction hypothesis
 189 and the fact that I_w is included in I_u . Thus $J_w = T^m(I_w) = I_w + \alpha_u + \alpha_a =$
 190 $I_w + \alpha_w$. Equation (2.5) shows in particular that the restriction of $T^{|w|}$ to I_w
 191 is a translation.

192 2.3 Uniformly recurrent sets

193 A set S of words on the alphabet A is said to be *factorial* if it contains the
 194 factors of its elements.

195 A factorial set is said to be *right-extendable* if for every $w \in S$ there is some
 196 $a \in A$ such that $wa \in S$. It is *biextendable* if for any $w \in S$, there are $a, b \in A$
 197 such that $awb \in S$.

198 A set of words $S \neq \{\varepsilon\}$ is *recurrent* if it is factorial and if for every $u, w \in S$
 199 there is a $v \in S$ such that $uvw \in S$. A recurrent set is biextendable. It is said
 200 to be *uniformly recurrent* if it is right-extendable and if, for any word $u \in S$,
 201 there exists an integer $n \geq 1$ such that u is a factor of every word of S of length
 202 n . A uniformly recurrent set is recurrent.

203 We denote by $A^{\mathbb{N}}$ the set of infinite words on the alphabet A . For a set
 204 $X \subset A^{\mathbb{N}}$, we denote by $F(X)$ the set of factors of the words of X .

205 Let S be a set of words on the alphabet A . For $w \in S$, set $R(w) = \{a \in$
 206 $A \mid wa \in S\}$ and $L(w) = \{a \in A \mid aw \in S\}$. A word w is called *right-special* if
 207 $\text{Card}(R(w)) \geq 2$ and *left-special* if $\text{Card}(L(w)) \geq 2$. It is *bispecial* if it is both
 208 right and left-special.

209 An infinite word on a binary alphabet is *Sturmian* if its set of factors is
 210 closed under reversal and if for each n there is exactly one right-special word of
 211 length n .

212 An infinite word is a *strict episturmian* word if its set of factors is closed
 213 under reversal and for each n there is exactly one right-special word w of length
 214 n , which is moreover such that $\text{Card}(R(w)) = \text{Card}(A)$.

215 A morphism $f : A^* \rightarrow A^*$ is called *primitive* if there is an integer k such that
 216 for all $a, b \in A$, the letter b appears in $f^k(a)$. If f is a primitive morphism, the
 217 set of factors of any fixpoint of f is uniformly recurrent (see [10], Proposition
 218 1.2.3 for example).

219 **Example 2.5** Let $A = \{a, b\}$. The Fibonacci word is the fixpoint $x = f^\omega(a) =$
 220 $abaababa \dots$ of the morphism $f : A^* \rightarrow A^*$ defined by $f(a) = ab$ and $f(b) = a$.
 221 It is a Sturmian word (see [13]). The set $F(x)$ of factors of x is the *Fibonacci*
 222 *set*.

223 **Example 2.6** Let $A = \{a, b, c\}$. The Tribonacci word is the fixpoint $x =$
 224 $f^\omega(a) = abacaba \dots$ of the morphism $f : A^* \rightarrow A^*$ defined by $f(a) = ab$,
 225 $f(b) = ac$, $f(c) = a$. It is a strict episturmian word (see [11]). The set $F(x)$ of
 226 factors of x is the *Tribonacci set*.

227 2.4 Interval exchange sets

228 Let T be an interval exchange set. The set $F(\Sigma_T(z))$ is called an *interval*
 229 *exchange set*. It is biextendable.

230 If T is a minimal interval exchange transformation, one has $w \in F(\Sigma_T(z))$
 231 if and only if $I_w \neq \emptyset$. Thus the set $F(\Sigma_T(z))$ does not depend on z . Since it
 232 depends only on T , we denote it by $F(T)$. When T is regular (resp. minimal),
 233 such a set is called a *regular interval exchange set* (resp. a minimal interval
 234 exchange set).

235 Let T be an interval exchange transformation. Let M be the closure in $A^{\mathbb{N}}$
 236 of the set of all $\Sigma_T(z)$ for $z \in [0, 1[$ and let σ be the shift on M . The pair
 237 (M, σ) is a *symbolic dynamical system*, formed of a topological space M and a
 238 continuous transformation σ . Such a system is said to be minimal if the only
 239 closed subsets invariant by σ are \emptyset or M (that is, every orbit is dense). It is
 240 well-known that (M, σ) is minimal if and only if $F(T)$ is uniformly recurrent
 241 (see for example [13] Theorem 1.5.9).

We have the following commutative diagram (Figure 2.2).

$$\begin{array}{ccc}
 [0, 1[& \xrightarrow{T} & [0, 1[\\
 \downarrow \Sigma_T & & \downarrow \Sigma_T \\
 M & \xrightarrow{\sigma} & M
 \end{array}$$

Figure 2.2: The transformations T and σ .

242 The map Σ_T is neither continuous nor surjective. This can be corrected by
 243 embedding the interval $[0, 1[$ into a larger space on which T is a homeomorphism
 244 (see [12] or [6] page 349). However, if the transformation T is minimal, the
 245 symbolic dynamical system (M, S) is minimal (see [6] page 392). Thus, we
 246 obtain the following statement.
 247

248 **Proposition 2.7** *For any minimal interval exchange transformation T , the set*
 249 *$F(T)$ is uniformly recurrent.*

250 Note that for a minimal interval exchange transformation T , the map Σ_T is
 251 injective (see [12] page 30).

252 The following is an elementary property of the intervals I_u which will be
 253 used below. We denote by $<_1$ the lexicographic order on A^* induced by the
 254 order $<_1$ on A .

255 **Proposition 2.8** *One has $I_u < I_v$ if and only if $u <_1 v$ and u is not a prefix*
 256 *of v .*

257 *Proof.* For a word u and a letter a , it results from (2.1) that $I_{ua} = I_u \cap T^{-|u|}(I_a)$.
 258 Since $(I_a)_{a \in A}$ is an ordered partition, this implies that $(T^{|u|}(I_u) \cap I_a)_{a \in A}$ is an
 259 ordered partition of $T^{|u|}(I_u)$. Since the restriction of $T^{|u|}$ to I_u is a translation,
 260 this implies that $(I_{ua})_{a \in A}$ is an ordered partition of I_u . Moreover, for two words
 261 u, v , it results also from (2.1) that $I_{uv} = I_u \cap T^{-|u|}(I_v)$. Thus $I_{uv} \subset I_u$.

262 Assume that $u <_1 v$ and that u is not a prefix of v . Then $u = las$ and
 263 $v = lbt$ with a, b two letters such that $a <_1 b$. Then we have $I_{la} < I_{lb}$, with
 264 $I_u \subset I_{la}$ and $I_v \subset I_{lb}$ whence $I_u < I_v$.

265 Conversely, assume that $I_u < I_v$. Since $I_u \cap I_v = \emptyset$, the words u, v cannot be
 266 comparable for the prefix order. Set $u = las$ and $v = lbt$ with a, b two distinct
 267 letters. If $b <_1 a$, then $I_v < I_u$ as we have shown above. Thus $a <_1 b$ which
 268 implies $u <_1 v$. ■

269 We denote by $<_2$ the order on A^* defined by $u <_2 v$ if u is a proper suffix
 270 of v or if $u = waz$ and $v = tbz$ with $a <_2 b$. Thus $<_2$ is the lexicographic order
 271 on the reversal of the words induced by the order $<_2$ on the alphabet.

272 We denote by π the morphism from A^* onto itself which extends to A^* the
 273 permutation π on A . Then $u <_2 v$ if and only if $\pi^{-1}(\tilde{u}) <_1 \pi^{-1}(\tilde{v})$, where \tilde{u}
 274 denotes the reversal of the word u .

275 The following statement is the analogue of Proposition 2.8.

276 **Proposition 2.9** *Let $T_{\lambda, \pi}$ be an interval exchange transformation. One has*
 277 *$J_u < J_v$ if and only if $u <_2 v$ and u is not a suffix of v .*

278 *Proof.* Let $(I'_a)_{a \in A}$ be the family of semi-intervals defined by $I'_a = J_{\pi(a)}$. Then
 279 the interval exchange transformation T' relative to (I'_a) with translation values
 280 $-\alpha_a$ is the inverse of the transformation T . The semi-intervals I'_w defined by
 281 Equation (2.1) with respect to T' satisfy $I'_w = J_{\pi(\tilde{w})}$ or equivalently $J_w =$

282 $I'_{\pi^{-1}(\tilde{w})}$. Thus, $J_u < J_v$ if and only if $I'_{\pi^{-1}(\tilde{u})} < I'_{\pi^{-1}(\tilde{v})}$ if and only if (by
 283 Proposition 2.8) $\pi^{-1}(\tilde{u}) <_1 \pi^{-1}(\tilde{v})$ or equivalently $u <_2 v$. ■

284 3 Bifix codes and interval exchange

285 In this section, we first introduce prefix codes and bifix codes. For a more de-
 286 tailed exposition, see [3]. We describe the link between maximal bifix codes and
 287 interval exchange transformations and we prove our main result (Theorem 3.13).

288 3.1 Prefix codes and bifix codes

289 A *prefix code* is a set of nonempty words which does not contain any proper
 290 prefix of its elements. A *suffix code* is defined symmetrically. A *bifix code* is a
 291 set which is both a prefix code and a suffix code.

292 A *coding morphism* for a prefix code $X \subset A^+$ is a morphism $f : B^* \rightarrow A^*$
 293 which maps bijectively B onto X .

294 Let S be a set of words. A prefix code $X \subset S$ is S -maximal if it is not
 295 properly contained in any prefix code $Y \subset S$. Note that if $X \subset S$ is an S -
 296 maximal prefix code, any word of S is comparable for the prefix order with a
 297 word of X .

298 A map $\lambda : A^* \rightarrow [0, 1]$ such that $\lambda(\varepsilon) = 1$ and, for any word w

$$\sum_{a \in A} \lambda(aw) = \sum_{a \in A} \lambda(wa) = \lambda(w), \quad (3.1)$$

299 is called an *invariant probability distribution* on A^* .

300 Let $T_{\lambda, \pi}$ be an interval exchange transformation. For any word $w \in A^*$,
 301 denote by $|I_w|$ the length of the semi-interval I_w defined by Equation (2.1). Set
 302 $\lambda(w) = |I_w|$. Then $\lambda(\varepsilon) = 1$ and for any word w , Equation (3.1) holds and thus
 303 λ is an invariant probability distribution.

304 The fact that λ is an invariant probability measure is equivalent to the fact
 305 that the Lebesgue measure on $[0, 1[$ is invariant by T . It is known that almost
 306 all regular interval exchange transformations have no other invariant probability
 307 measure (and thus are uniquely ergodic, see [6] for references).

308 **Example 3.1** Let S be the set of factors of the Fibonacci word (see Exam-
 309 ple 2.5). It is the natural coding of the rotation of angle $\alpha = (3 - \sqrt{5})/2$ with
 310 respect to α (see [13], Chapter 2). The values of the map λ on the words of
 311 length at most 4 in S are indicated in Figure 3.1.

312 The following result is a particular case of a result from [2] (Proposition
 313 3.3.4).

314 **Proposition 3.2** *Let T be a minimal interval exchange transformation, let $S =$
 315 $F(T)$ and let λ be an invariant probability distribution on S . For any finite S -
 316 maximal prefix code X , one has $\sum_{x \in X} \lambda(x) = 1$.*

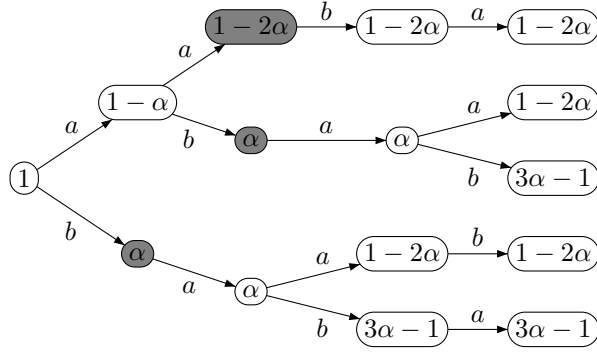


Figure 3.1: The invariant probability distribution on the Fibonacci set.

317 The following statement is connected with Proposition 3.2.

318 **Proposition 3.3** *Let T be a minimal interval exchange transformation relative*
 319 *to $(I_a)_{a \in A}$, let $S = F(T)$ and let X be a finite S -maximal prefix code ordered*
 320 *by $<_1$. The family $(I_w)_{w \in X}$ is an ordered partition of $[0, 1[$.*

321 *Proof.* By Proposition 2.8, the sets (I_w) for $w \in X$ are pairwise disjoint. Let
 322 π be the invariant probability distribution on S defined by $\pi(w) = |I_w|$. By
 323 Proposition 3.2, we have $\sum_{w \in X} \pi(w) = 1$. Thus the family $(I_w)_{w \in X}$ is a parti-
 324 tion of $[0, 1[$. By Proposition 2.8 it is an ordered partition. ■

325 **Example 3.4** Let T be the rotation of angle $\alpha = (3 - \sqrt{5})/2$. The set $S = F(T)$
 326 is the Fibonacci set. The set $X = \{aa, ab, b\}$ is an S -maximal prefix code (see
 327 the grey nodes in Figure 3.1). The partition of $[0, 1[$ corresponding to X is

$$I_{aa} = [0, 1 - 2\alpha[, \quad I_{ab} = [1 - 2\alpha, 1 - \alpha[, \quad I_b = [1 - \alpha, 1[.$$

328 The values of the lengths of the semi-intervals (the invariant probability distri-
 329 bution) can also be read on Figure 3.1.

330 A symmetric statement holds for an S -maximal suffix code, namely that the
 331 family $(J_w)_{w \in X}$ is an ordered partition of $[0, 1[$ for the order $<_2$ on X .

332 3.2 Maximal bifix codes

333 Let S be a set of words. A bifix code $X \subset S$ is S -maximal if it is not properly
 334 contained in a bifix code $Y \subset S$. For a recurrent set S , a finite bifix code is
 335 S -maximal as a bifix code if and only if it is an S -maximal prefix code (see [2],
 336 Theorem 4.2.2).

337 A *parse* of a word w with respect to a bifix code X is a triple (v, x, u) such
 338 that $w = vxu$ where v has no suffix in X , u has no prefix in X and $x \in X^*$. We
 339 denote by $\delta_X(w)$ the number of parses of w with respect to X .

340 The number of parses of a word w is also equal to the number of suffixes of
 341 w which have no prefix in X and the number of prefixes of w which have no
 342 suffix in X (see Proposition 6.1.6 in [3]).

343 By definition, the S -degree of a bifix code X , denoted $d_X(S)$, is the maximal
 344 number of parses of a word in S . It can be finite or infinite.

345 The set of *internal factors* of a set of words X , denoted $I(X)$, is the set of
 346 words w such that there exist nonempty words u, v with $uwv \in X$.

347 Let S be a recurrent set and let X be a finite S -maximal bifix code of S -
 348 degree d . A word $w \in S$ is such that $\delta_X(w) < d$ if and only if it is an internal
 349 factor of X , that is

$$I(X) = \{w \in S \mid \delta_X(w) < d\}$$

350 (Theorem 4.2.8 in [2]). Thus any word of S which is not a factor of X has d
 351 parses. This implies that the S -degree d is finite.

352 **Example 3.5** Let S be a recurrent set. For any integer $n \geq 1$, the set $S \cap A^n$
 353 is an S -maximal bifix code of S -degree n .

354 The *kernel* of a bifix code X is the set $K(X) = I(X) \cap X$. Thus it is the set
 355 of words of X which are also internal factors of X . By Theorem 4.3.11 of [2], a
 356 finite S -maximal bifix code is determined by its S -degree and its kernel.

357 **Example 3.6** Let S be the Fibonacci set. The set $X = \{a, baab, bab\}$ is the
 358 unique S -maximal bifix code of S -degree 2 with kernel $\{a\}$. Indeed, the word
 359 bab is not an internal factor and has two parses, namely $(1, bab, 1)$ and (b, a, b) .

360 The following result shows that bifix codes have a natural connection with
 361 interval exchange transformations.

362 **Proposition 3.7** *If X is a finite S -maximal bifix code, with S as in Propo-*
 363 *sition 3.3, the families $(I_w)_{w \in X}$ and $(J_w)_{w \in X}$ are ordered partitions of $[0, 1[$,*
 364 *relatively to the orders $<_1$ and $<_2$ respectively.*

365 *Proof.* This results from Proposition 3.3 and its symmetric and from the fact
 366 that, since S is recurrent, a finite S -maximal bifix code is both an S -maximal
 367 prefix code and an S -maximal suffix code. ■

368 Let T be a regular interval exchange transformation relative to $(I_a)_{a \in A}$. Let
 369 $(\alpha_a)_{a \in A}$ be the translation values of T . Set $S = F(T)$. Let X be a finite
 370 S -maximal bifix code on the alphabet A .

371 Let T_X be the transformation on $[0, 1[$ defined by

$$T_X(z) = T^{|u|}(z) \quad \text{if } z \in I_u$$

372 with $u \in X$. The transformation is well-defined since, by Proposition 3.7, the
 373 family $(I_u)_{u \in X}$ is a partition of $[0, 1[$.

374 Let $f : B^* \rightarrow A^*$ be a coding morphism for X . Let $(K_b)_{b \in B}$ be the family
 375 of semi-intervals indexed by the alphabet B with $K_b = I_{f(b)}$. We consider B as

376 ordered by the orders $<_1$ and $<_2$ induced by f . Let T_f be the interval exchange
 377 transformation relative to $(K_b)_{b \in B}$. Its translation values are $\beta_b = \sum_{j=0}^{m-1} \alpha_{a_j}$
 378 for $f(b) = a_0 a_1 \cdots a_{m-1}$. The transformation T_f is called the *transformation*
 379 *associated* with f .

380 **Proposition 3.8** *Let T be a regular interval exchange transformation relative*
 381 *to $(I_a)_{a \in A}$ and let $S = F(T)$. If $f : B^* \rightarrow A^*$ is a coding morphism for a finite*
 382 *S -maximal bifix code X , one has $T_f = T_X$.*

383 *Proof.* By Proposition 3.7, the family $(K_b)_{b \in B}$ is a partition of $[0, 1[$ ordered
 384 by $<_1$. For any $w \in X$, we have by Equation (2.5) $J_w = I_w + \alpha_w$ and thus
 385 T_X is the interval exchange transformation relative to $(K_b)_{b \in B}$ with translation
 386 values β_b . ■

387 In the sequel, under the hypotheses of Proposition 3.8, we consider T_f as an
 388 interval exchange transformation. In particular, the natural coding of T_f relative
 389 to $z \in [0, 1[$ is well-defined.

390 **Example 3.9** Let S be the Fibonacci set. It is the set of factors of the Fi-
 391 bonacci word, which is a natural coding of the rotation of angle $\alpha = (3 - \sqrt{5})/2$
 392 relative to α (see Example 3.1). Let $X = \{aa, ab, ba\}$ and let f be the coding
 393 morphism defined by $f(u) = aa$, $f(v) = ab$, $f(w) = ba$. The two partitions of
 394 $[0, 1[$ corresponding to T_f are

$$I_u = [0, 1 - 2\alpha[, \quad I_v = [1 - 2\alpha, 1 - \alpha[, \quad I_w = [1 - \alpha, 1[$$

395 and

$$J_v = [0, \alpha[, \quad J_w = [\alpha, 2\alpha[, \quad J_u = [2\alpha, 1[$$

The transformation T_f is represented in Figure 3.2. It is actually a representa-

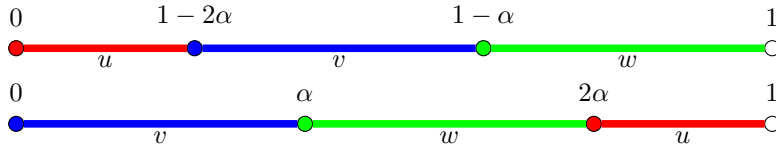


Figure 3.2: The transformation T_f .

396 tion on 3 intervals of the rotation of angle 2α . Note that the point $z = 1 - \alpha$ is
 397 a separation point which is not a singularity of T_f . The first row of Table 3.1
 398 gives the two orders on X . The next two rows give the two orders for each of
 399 the two other S -maximal bifix codes of S -degree 2 (there are actually exactly
 400 three S -maximal bifix codes of S -degree 2 in the Fibonacci set, see [2]).
 401

402 Let T be a minimal interval exchange transformation on the alphabet A .
 403 Let x be the natural coding of T relative to some $z \in [0, 1[$. Set $S = F(x)$. Let
 404 X be a finite S -maximal bifix code. Let $f : B^* \rightarrow A^*$ be a morphism which
 405 maps bijectively B onto X . Since S is recurrent, the set X is an S -maximal

$(X, <_1)$	$(X, <_2)$
aa, ab, ba	ab, ba, aa
$a, baab, bab$	$bab, baab, a$
aa, aba, b	b, aba, aa

Table 3.1: The two orders on the three S -maximal bifix codes of S -degree 2.

406 prefix code. Thus x has a prefix $x_0 \in X$. Set $x = x_0x'$. In the same way x'
407 has a prefix x_1 in X . Iterating this argument, we see that $x = x_0x_1 \cdots$ with
408 $x_i \in X$. Consequently, there exists an infinite word y on the alphabet B such
409 that $x = f(y)$. The word y is the *decoding* of the infinite word x with respect
410 to f .

411 **Proposition 3.10** *The decoding of x with respect to f is the natural coding of*
412 *the transformation associated with f relative to z : $\Sigma_T(z) = f(\Sigma_{T_f}(z))$.*

413 *Proof.* Let $y = b_0b_1 \cdots$ be the decoding of x with respect to f . Set $x_i = f(b_i)$
414 for $i \geq 0$. Then, for any $n \geq 0$, we have

$$T_f^n(z) = T^{|u_n|}(z) \quad (3.2)$$

415 with $u_n = x_0 \cdots x_{n-1}$ (note that $|u_n|$ denotes the length of u_n with respect to
416 the alphabet A). Indeed, this is true for $n = 0$. Next $T_f^{n+1}(z) = T_f(t)$ with $t =$
417 $T_f^n(z)$. Arguing by induction, we have $t = T^{|u_n|}(z)$. Since $x = u_nx_nx_{n+1} \cdots$,
418 t is in I_{x_n} by (2.3). Thus by Proposition 3.8, $T_f(t) = T^{|x_n|}(t)$ and we obtain
419 $T_f^{n+1}(z) = T^{|x_n|}(T^{|u_n|}(z)) = T^{|u_{n+1}|}(z)$ proving (3.2). Finally, for $u = f(b)$
420 with $b \in B$,

$$b_n = b \iff x_n = u \iff T^{|u_n|}(z) \in I_u \iff T_f^n(z) \in I_u = K_b$$

421 showing that y is the natural coding of T_f relative to z . ■

423 **Example 3.11** Let T, α, X and f be as in Example 3.9. Let $x = abaababa \cdots$
424 be the Fibonacci word. We have $x = \Sigma_T(\alpha)$. The decoding of x with respect to
425 f is $y = vuwv \cdots$.

426 3.3 Bifix codes and regular transformations

427 The following result shows that for the coding morphism f of a finite S -maximal
428 bifix code, the map $T \mapsto T_f$ preserves the regularity of the transformation.

429 **Theorem 3.12** *Let T be a regular interval exchange transformation and let*
430 *$S = F(T)$. For any finite S -maximal bifix code X with coding morphism f , the*
431 *transformation T_f is regular.*

432 *Proof.* Set $A = \{a_1, a_2, \dots, a_s\}$ with $a_1 <_1 a_2 <_1 \dots <_1 a_s$. We denote
 433 $\delta_i = \delta_{a_i}$. By hypothesis, the orbits of $\delta_2, \dots, \delta_s$ are infinite and disjoint. Set
 434 $X = \{x_1, x_2, \dots, x_t\}$ with $x_1 <_1 x_2 <_1 \dots <_1 x_t$. Let d be the S -degree of X .

435 For $x \in X$, denote by δ_x the left boundary of the semi-interval J_x . For each
 436 $x \in X$, it follows from Equation (2.2) that there is an $i \in \{1, \dots, s\}$ such that
 437 $\delta_x = T^k(\delta_i)$ with $0 \leq k < |x|$. Moreover, we have $i = 1$ if and only if $x = x_1$.
 438 Since T is regular, the index $i \neq 1$ and the integer k are unique for each $x \neq x_1$.
 439 And for such x and i , by (2.4), we have $\Sigma_T(\delta_i) = u\Sigma_T(\delta_x)$ with u a proper suffix
 440 of x .

441 We now show that the orbits of $\delta_{x_2}, \dots, \delta_{x_t}$ for the transformation T_f are
 442 infinite and disjoint. Assume that $\delta_{x_p} = T_f^n(\delta_{x_q})$ for some $p, q \in \{2, \dots, t\}$ and
 443 $n \in \mathbb{Z}$. Interchanging p, q if necessary, we may assume that $n \geq 0$. Let $i, j \in$
 444 $\{2, \dots, s\}$ be such that $\delta_{x_p} = T^k(\delta_i)$ with $0 \leq k < |x_p|$ and $\delta_{x_q} = T^\ell(\delta_j)$ with
 445 $0 \leq \ell < |x_q|$. Since $T^k(\delta_i) = T_f^n(T^\ell(\delta_j)) = T^{m+\ell}(\delta_j)$ for some $m \geq 0$, we cannot
 446 have $i \neq j$ since otherwise the orbits of δ_i, δ_j for the transformation T intersect.
 447 Thus $i = j$. Since $\delta_{x_p} = T^k(\delta_i)$, we have $\Sigma_T(\delta_i) = u\Sigma_T(\delta_{x_p})$ with $|u| = k$, u
 448 proper suffix of x_p . And since $\delta_{x_p} = T_f^n(\delta_{x_q})$, we have $\Sigma_T(\delta_{x_q}) = x\Sigma_T(\delta_{x_p})$ with
 449 $x \in X^*$. Since on the other hand $\delta_{x_q} = T^\ell(\delta_i)$, we have $\Sigma_T(\delta_i) = v\Sigma_T(\delta_{x_q})$ with
 450 $|v| = \ell$ and v a proper suffix of x_q . We obtain

$$\begin{aligned} \Sigma_T(\delta_i) &= u\Sigma_T(\delta_{x_p}) \\ &= v\Sigma_T(\delta_{x_q}) = vx\Sigma_T(\delta_{x_p}) \end{aligned}$$

451 Since $|u| = |vx|$, this implies $u = vx$. But since u cannot have a suffix in X ,
 452 $u = vx$ implies $x = 1$ and thus $n = 0$ and $p = q$. This concludes the proof. ■

453 Let f be a coding morphism for a finite S -maximal bifix code $X \subset S$. The set
 454 $f^{-1}(S)$ is called a *maximal bifix decoding* of S .

455 **Theorem 3.13** *The family of regular interval exchange sets is closed under*
 456 *maximal bifix decoding.*

457 *Proof.* Let T be a regular interval exchange transformation such that $S = F(T)$.
 458 By Theorem 3.12, T_f is a regular interval exchange transformation. We show
 459 that $f^{-1}(S) = F(T_f)$, which implies the conclusion.

460 Let $x = \Sigma_T(z)$ for some $z \in [0, 1[$ and let $y = f^{-1}(x)$. Then $S = F(x)$ and
 461 $F(T_f) = F(y)$. For any $w \in F(y)$, we have $f(w) \in F(x)$ and thus $w \in f^{-1}(S)$.
 462 This shows that $F(T_f) \subset f^{-1}(S)$. Conversely, let $w \in f^{-1}(S)$ and let $v = f(w)$.
 463 Since $S = F(x)$, there is a word u such that uv is a prefix of x . Set $z' = T^{|u|}(z)$
 464 and $x' = \Sigma_T(z')$. Then v is a prefix of x' and w is a prefix of $y' = f^{-1}(x')$.
 465 Since T_f is regular, it is minimal and thus $F(y') = F(T_f)$. This implies that
 466 $w \in F(T_f)$. ■

467 Since a regular interval exchange set is uniformly recurrent, Theorem 3.13
 468 implies in particular that if S is a regular interval exchange set and f a coding
 469 morphism of a finite S -maximal bifix code, then $f^{-1}(S)$ is uniformly recurrent.

470 This is not true for an arbitrary uniformly recurrent set S , as shown by the
 471 following example.

472 **Example 3.14** Set $A = \{a, b\}$ and $B = \{u, v\}$. Let S be the set of factors
 473 of $(ab)^*$ and let $f : B^* \rightarrow A^*$ be defined by $f(u) = ab$ and $f(v) = ba$. Then
 474 $f^{-1}(S) = u^* \cup v^*$ which is not recurrent.

475 We illustrate the proof of Theorem 3.12 in the following example.

476 **Example 3.15** Let T be the rotation of angle $\alpha = (3 - \sqrt{5})/2$. The set $S =$
 477 $F(T)$ is the Fibonacci set. Let $X = \{a, baab, babaabaabab, babaabab\}$. The set
 478 X is an S -maximal bifix code of S -degree 3 (see [2]). The values of the μ_{x_i}
 479 (which are the right boundaries of the intervals I_{x_i}) and δ_{x_i} are represented in
 Figure 3.3.

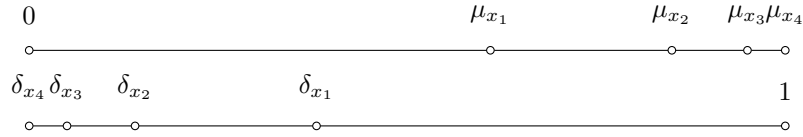


Figure 3.3: The transformation associated with a bifix code of S -degree 3.

480 The infinite word $\Sigma_T(0)$ is represented in Figure 3.4. The value indicated
 481 on the word $\Sigma_T(0)$ after a prefix u is $T^{|u|}(0)$. The three values $\delta_{x_4}, \delta_{x_2}, \delta_{x_3}$
 482 correspond to the three prefixes of $\Sigma_T(0)$ which are proper suffixes of X .

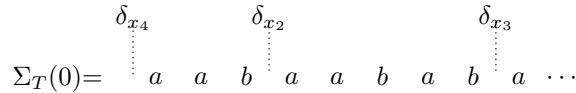


Figure 3.4: The infinite word $\Sigma_T(0)$.

483 The following example shows that Theorem 3.13 is not true when X is not bifix.
 484
 485

486 **Example 3.16** Let S be the Fibonacci set and let $X = \{aa, ab, b\}$. The set X is
 487 an S -maximal prefix code. Let $B = \{u, v, w\}$ and let f be the coding morphism
 488 for X defined by $f(u) = aa$, $f(v) = ab$, $f(w) = b$. The set $W = f^{-1}(S)$ is
 489 not an interval exchange set. Indeed, we have $vu, vv, wu, vw \in W$. This implies
 490 that both J_v and J_w meet I_u and I_v , which is impossible in an interval exchange
 491 transformation.

492 **4 Tree sets**

493 We introduce in this section the notions of tree sets and planar tree sets. We first
 494 introduce the notion of extension graph which describes the possible two-sided
 495 extensions of a word.

496 **4.1 Extension graphs**

497 Let S be a biextendable set of words. For $w \in S$, we denote

$$L(w) = \{a \in A \mid aw \in S\}, \quad R(w) = \{a \in A \mid wa \in S\}$$

498 and

$$E(w) = \{(a, b) \in A \times A \mid awb \in S\}.$$

499 For $w \in S$, the *extension graph* of w is the undirected bipartite graph $G(w)$ on
 500 the set of vertices which is the disjoint union of two copies of $L(w)$ and $R(w)$
 501 with edges the pairs $(a, b) \in E(w)$.

502 Recall that an undirected graph is a tree if it is connected and acyclic.

503 Let S be a biextendable set. We say that S is a *tree set* if the graph $G(w)$
 504 is a tree for all $w \in S$.

505 Let $<_1$ and $<_2$ be two orders on A . For a set S and a word $w \in S$, we
 506 say that the graph $G(w)$ is *compatible* with the orders $<_1$ and $<_2$ if for any
 507 $(a, b), (c, d) \in E(w)$, one has

$$a <_1 c \implies b \leq_2 d.$$

508 Thus, placing the vertices of $L(w)$ ordered by $<_1$ on a line and those of $R(w)$
 509 ordered by $<_2$ on a parallel line, the edges of the graph may be drawn as straight
 510 noncrossing segments, resulting in a planar graph.

511 We say that a biextendable set S is a *planar tree set* with respect to two
 512 orders $<_1$ and $<_2$ on A if for any $w \in S$, the graph $G(w)$ is a tree compatible
 513 with $<_1, <_2$. Obviously, a planar tree set is a tree set.

514 The following example shows that the Tribonacci set is not a planar tree set.
 515

516 **Example 4.1** Let S be the Tribonacci set (see example 2.6). The words a, aba
 517 and $abacaba$ are bispecial. Thus the words $ba, caba$ are right-special and the
 518 words $ab, abac$ are left-special. The graphs $G(\varepsilon), G(a)$ and $G(aba)$ are shown in
 Figure 4.1. One sees easily that it not possible to find two orders on A making

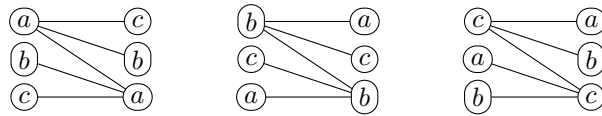


Figure 4.1: The graphs $G(\varepsilon), G(a)$ and $G(aba)$ in the Tribonacci set.

519 the three graphs planar.
 520

521 **4.2 Interval exchange sets and planar tree sets**

522 The following result is proved in [9] with a converse (see below).

523 **Proposition 4.2** *Let T be an interval exchange transformation on A ordered*
 524 *by $<_1$ and $<_2$. If T is regular, the set $F(T)$ is a planar tree set with respect to*
 525 *$<_2$ and $<_1$.*

526 *Proof.* Assume that T is a regular interval exchange transformation relative to
 527 $(I_a, \alpha_a)_{a \in A}$ and let $S = F(T)$.

528 Since T is minimal, w is in S if and only if $I_w \neq \emptyset$. Thus, one has

- 529 (i) $b \in R(w)$ if and only if $I_w \cap T^{-|w|}(I_b) \neq \emptyset$ and
 530 (ii) $a \in L(w)$ if and only if $J_a \cap I_w \neq \emptyset$.

531 Condition (i) holds because $I_{wb} = I_w \cap T^{-|w|}(I_b)$ and condition (ii) because
 532 $I_{aw} = I_a \cap T^{-1}(I_w)$, which implies $T(I_{aw}) = J_a \cap I_w$. In particular, (i) implies
 533 that $(I_{wb})_{b \in R(w)}$ is an ordered partition of I_w with respect to $<_1$.

534 We say that a path in a graph is reduced if does not use consecutively the
 535 same edge. For $a, a' \in L(w)$ with $a <_2 a'$, there is a unique reduced path in
 536 $G(w)$ from a to a' which is the sequence a_1, b_1, \dots, a_n with $a_1 = a$ and $a_n = a'$
 537 with $a_1 <_2 a_2 <_2 \dots <_2 a_n$, $b_1 <_1 b_2 <_1 \dots <_1 b_{n-1}$ and $J_{a_i} \cap I_{wb_i} \neq \emptyset$,
 538 $J_{a_{i+1}} \cap I_{wb_i} \neq \emptyset$ for $1 \leq i \leq n-1$ (see Figure 4.2). Note that the hypothesis
 539 that T is regular is needed here since otherwise the right boundary of J_{a_i} could
 540 be the left boundary of I_{wb_i} . Thus $G(w)$ is a tree. It is compatible with $<_2, <_1$
 541 since the above shows that $a <_2 a'$ implies that the letters b_1, b_{n-1} such that
 542 $(a, b_1), (a', b_{n-1}) \in E(w)$ satisfy $b_1 \leq_1 b_{n-1}$. ■

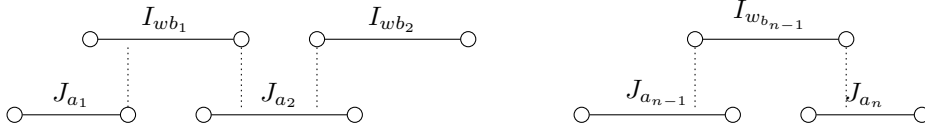


Figure 4.2: A path from a_1 to a_n in $G(w)$.

543 By Proposition 4.2, a regular interval exchange set is a planar tree set, and
 544 thus in particular a tree set. Note that the analogue of Theorem 3.13 holds for
 545 the class of uniformly recurrent tree sets [4].

546 The main result of [9] states that a uniformly recurrent set S on an alphabet
 547 A is a regular interval exchange set if and only if $A \subset S$ and there exist two
 548 orders $<_1$ and $<_2$ on A such that the following conditions are satisfied for any
 549 word $w \in S$.

- 550 (i) The set $L(w)$ (resp. $R(w)$) is formed of consecutive elements for the order
 551 $<_2$ (resp. $<_1$).
 552 (ii) For $(a, b), (c, d) \in E(w)$, if $a <_2 c$, then $b \leq_1 d$.
 553 (iii) If $a, b \in L(w)$ are consecutive for the order $<_2$, then the set $R(aw) \cap R(bw)$
 554 is a singleton.

555 It is easy to see that a biextendable set S containing A satisfies (ii) and (iii)
556 if and only if it is a planar tree set. Actually, in this case, it automatically
557 satisfies also condition (i). Indeed, let us consider a word w and $a, b, c \in A$ with
558 $a <_1 b <_1 c$ such that $wa, wc \in S$ but $wb \notin S$. Since $b \in S$ there is a (possibly
559 empty) suffix v of w such that $vb \in S$. We choose v of maximal length. Since
560 $wb \notin S$, we have $w = uv$ with u nonempty. Let d be the last letter of u . Then
561 we have $dva, dvc \in S$ and $dwb \notin S$. Since $G(v)$ is a tree and $b \in R(v)$, there is a
562 letter $e \in L(v)$ such that $evb \in S$. But $e <_2 d$ and $d <_2 e$ are both impossible
563 since $G(v)$ is compatible with $<_2$ and $<_1$. Thus we reach a contradiction.

564 This shows that the following reformulation of the main result of [9] is equiv-
565 alent to the original one.

566 **Theorem 4.3 (Ferenczi, Zamboni)** *A set S is a regular interval exchange*
567 *set on the alphabet A if and only if it is a uniformly recurrent planar tree set*
568 *containing A .*

569 We have already seen that the Tribonacci set is a tree set which is not a
570 planar tree set (Example 4.1). The next example shows that there are uniformly
571 recurrent tree sets which are neither Sturmian nor regular interval exchange sets.
572

573 **Example 4.4** Let S be the Tribonacci set on the alphabet $A = \{a, b, c\}$ and
574 let $f : \{x, y, z, t, u\}^* \rightarrow A^*$ be the coding morphism for $X = S \cap A^2$ defined by
575 $f(x) = aa, f(y) = ab, f(z) = ac, f(t) = ba, f(u) = ca$. By Theorem 7.1 in [4],
576 the set $W = f^{-1}(S)$ is a uniformly recurrent tree set. It is not Sturmian since
577 y and t are two right-special words. It is not either a regular interval exchange
578 set. Indeed, for any right-special word w of W , one has $\text{Card}(R(w)) = 3$. This
579 is not possible in a regular interval exchange set T since, Σ_T being injective,
580 the length of the interval J_w tends to 0 as $|w|$ tends to infinity and it cannot
581 contain several separation points. It can of course also be verified directly that
582 W is not a planar tree set.

583 4.3 Exchange of pieces

584 In this section, we show how one can define a generalization of interval exchange
585 transformations called exchange of pieces. In the same way as interval exchange
586 is a generalization of rotations on the circle, exchange of pieces is a generalization
587 of rotations of the torus. We begin by studying this direction starting from the
588 Tribonacci word. For more on the Tribonacci word, see [17] and also [14, Chap.
589 10].

590 **The Tribonacci shift** The Tribonacci set S is not an interval exchange set
591 but it is however the natural coding of another type of geometric transformation,
592 namely an exchange of pieces in the plane, which is also a translation acting on
593 the two-dimensional torus \mathbb{T}^2 . This will allow us to show that the decoding of

594 the Tribonacci word with respect to a coding morphism for a finite S -maximal
 595 bifix code is again a natural coding of an exchange of pieces.

596 The *Tribonacci shift* is the symbolic dynamical system (M_x, σ) , where $M_x =$
 597 $\overline{\{\sigma^n(x) : n \in \mathbb{N}\}}$ is the closure of the σ -orbit of x where x is the Tribonacci word.
 598 By uniform recurrence of the Tribonacci word, (M_x, σ) is minimal and $M_x = M_y$
 599 for each $y \in M_x$ ([16, Proposition 4.7]). The Tribonacci set is the set of factors
 600 of the Tribonacci shift (M_x, σ) .

601 **Natural coding** Let Λ be a full-rank lattice in \mathbb{R}^d . We say that an infinite
 602 word x is a *natural coding* of a toral translation $T_{\mathbf{t}} : \mathbb{R}^d/\Lambda \rightarrow \mathbb{R}^d/\Lambda$, $\mathbf{x} \mapsto \mathbf{x} + \mathbf{t}$
 603 if there exists a fundamental domain R for Λ together with a partition $R =$
 604 $R_1 \cup \dots \cup R_k$ such that on each R_i ($1 \leq i \leq k$), there exists a vector \mathbf{t}_i such that
 605 the map $T_{\mathbf{t}}$ is given by the translation along \mathbf{t}_i , and x is the coding of a point
 606 $\mathbf{x} \in R$ with respect to this partition. A symbolic dynamical system (M, σ) is
 607 a *natural coding* of $(\mathbb{R}^d/\Lambda, T_{\mathbf{t}})$ if every element of M is a natural coding of the
 608 orbit of some point of the d -dimensional torus \mathbb{R}^d/Λ (with respect to the same
 609 partition) and if, furthermore, (M, σ) and $(\mathbb{R}^d/\Lambda, T_{\mathbf{t}})$ are measurably conjugate.

610 **Definition of the Rauzy fractal** Let β stand for the Perron-Frobenius eigen-
 611 value of the Tribonacci substitution. It is the largest root of $z^3 - z^2 - z - 1$.
 612 Consider the translation $R_\beta : \mathbb{T}^2 \rightarrow \mathbb{T}^2$, $x \mapsto x + (1/\beta, 1/\beta^2)$. Rauzy introduces
 613 in [18] a fundamental domain for a two-dimensional lattice, called the Rauzy
 614 fractal (it has indeed fractal boundary), which provides a partition for the sym-
 615 bolic dynamical system (M_x, σ) to be a natural coding for R_β . The Tribonacci
 616 word is a natural coding of the orbit of the point 0 under the action of the
 617 toral translation in \mathbb{T}^2 : $x \mapsto x + (\frac{1}{\beta}, \frac{1}{\beta^2})$. Similarly as in the case of interval
 618 exchanges, we have the following commutative diagram

$$\begin{array}{ccc} \mathbb{T}^2 & \xrightarrow{R_\beta} & \mathbb{T}^2 \\ \downarrow & & \downarrow \\ M_x & \xrightarrow{\sigma} & M_x \end{array}$$

619 The *Abelianization map* \mathbf{f} of the free monoid $\{1, 2, 3\}^*$ is defined by $\mathbf{f} :$
 620 $\{1, 2, 3\}^* \rightarrow \mathbb{Z}^3$, $\mathbf{f}(w) = |w|_1 e_1 + |w|_2 e_2 + |w|_3 e_3$, where $|w|_i$ denotes the number
 621 of occurrences of the letter i in the word w , and (e_1, e_2, e_3) stands for the
 622 canonical basis of \mathbb{R}^3 .

623 Let f be the morphism $a \mapsto ab, b \mapsto ac, c \mapsto a$ such that the Tribonacci word
 624 is the fixpoint of f (see Example 2.6). The incidence matrix F of f is defined
 625 by $F = (|f(j)|_i)_{(i,j) \in \mathcal{A}^2}$, where $|f(j)|_i$ counts the number of occurrences of the

626 letter i in $f(j)$. One has $F = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. The incidence matrix F admits as

627 eigenspaces in \mathbb{R}^3 one *expanding eigenline* (generated by the eigenvector with
 628 positive coordinates $v_\beta = (1/\beta, 1/\beta^2, 1/\beta^3)$ associated with the eigenvalue β).

629 We consider the projection π onto the antidiagonal plane $x + y + z = 0$ along
 630 the expanding direction of the matrix F .

631 One associates with the Tribonacci word $x = (x_n)_{n \geq 0}$ a broken line starting
 632 from 0 in \mathbb{Z}^3 and approximating the expanding line v_β as follows. The *Tribonacci*
 633 *broken line* is defined as the broken line which joins with segments of length 1
 634 the points $\mathbf{f}(x_0 x_1 \cdots x_{n-1})$, $n \in \mathbb{N}$. In other words we describe this broken line
 635 by starting from the origin, and then by reading successively the letters of the
 636 Tribonacci word x , going one step in direction e_i if one reads the letter i . The
 637 vectors $\mathbf{f}(x_0 x_1 \cdots x_n)$, $n \in \mathbb{N}$, stay within bounded distance of the expanding
 638 line (this comes from the fact that β is a Pisot number). The closure of the
 639 set of projected vertices of the broken line is called the *Rauzy fractal* and is
 640 represented on Figure 4.3. We thus define the Rauzy fractal \mathcal{R} as

$$\mathcal{R} := \overline{\{\pi(\mathbf{f}(x_0 \cdots x_{n-1})); n \in \mathbb{N}\}},$$

641 where $x_0 \dots x_{n-1}$ stands for the empty word when $n = 0$.

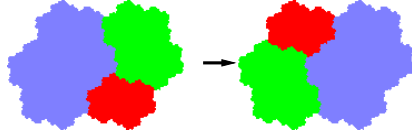


Figure 4.3: The Rauzy fractal

642 The Rauzy fractal is divided into three pieces, for $i = \{1, 2, 3\}$

$$\begin{aligned} \mathcal{R}(i) &:= \overline{\{\pi(\mathbf{f}(x_0 \cdots x_{n-1})); x_n = i, n \in \mathbb{N}\}}, \\ \mathcal{R}'(i) &:= \overline{\{\pi(\mathbf{f}(x_0 \cdots x_n)); x_n = i, n \in \mathbb{N}\}}. \end{aligned}$$

643 It has been proved in [18] that these pieces have non-empty interior and are
 644 disjoint up to a set of zero measure. The following exchange of pieces E is thus
 645 well-defined

$$E : \text{Int } \mathcal{R}_1 \cup \text{Int } \mathcal{R}_2 \cup \text{Int } \mathcal{R}_3 \rightarrow \mathcal{R}, \quad x \mapsto x + \pi(e_i), \quad \text{when } x \in \text{Int } \mathcal{R}_i.$$

646 One has $E(\mathcal{R}_i) = \mathcal{R}'_i$, for all i .

647 We consider the lattice Λ generated by the vectors $\pi(e_i) - \pi(e_j)$, for $i \neq j$.
 648 The Rauzy fractal tiles periodically the plane, that is, $\cup_{\gamma \in \Lambda} \gamma + \mathcal{R}$ is equal to
 649 the plane $x + y + z = 0$, and for $\gamma \neq \gamma' \in \Lambda$, $\gamma + \mathcal{R}$ and $\gamma' + \mathcal{R}$ do not intersect
 650 (except on a set of zero measure). This is why the exchange of pieces is in fact
 651 measurably conjugate to the translation R_β . Indeed the vector of coordinates
 652 of $\pi(\mathbf{f}(x_0 x_1 \cdots x_{n-1}))$ in the basis $(\pi(e_3) - \pi(e_1), \pi(e_3) - \pi(e_2))$ of the plane
 653 $x + y + z = 0$ is $n \cdot (1/\beta, 1/\beta^2) - (|x_0 x_1 \cdots x_{n-1}|_1, |x_0 x_1 \cdots x_{n-1}|_2)$. Hence the
 654 coordinates of $E^n(0)$ in the basis $(e_3 - e_1, e_3 - e_2)$ are equal to $R_\beta^n(0)$ modulo
 655 \mathbb{Z}^2 .

656 **Bifix codes and exchange of pieces** Let $(\mathcal{R}_a)_{a \in A}$ and $(\mathcal{R}'_a)_{a \in A}$ be two
657 families of subsets of a compact set \mathcal{R} included in \mathbb{R}^d . We assume that the
658 families $(\mathcal{R}_a)_{a \in A}$ and the $(\mathcal{R}'_a)_{a \in A}$ both form a partition of \mathcal{R} up to a set of
659 zero measure. We assume that there exist vectors e_a such that $\mathcal{R}'_a = \mathcal{R}_a + e_a$
660 for any $a \in A$. The exchange of pieces associated with these data is the map E
661 defined on \mathcal{R} (except a set of measure zero) by $E(z) = z + e_a$ if $z \in \mathcal{R}_a$. The
662 notion of natural coding of an exchange of pieces extends here in a natural way.

663 Assume that E is an exchange of pieces as defined above. Let S be the set
664 of factors of the natural codings of E . We assume that S is uniformly recurrent.

665 By analogy with the case of interval exchanges, let $I_a = \mathcal{R}_a$ and let $J_a =$
666 $E(\mathcal{R}_a)$. For a word $w \in A^*$, one defines similarly as for interval exchanges I_w
667 and J_w .

668 Let X be a finite S -maximal prefix code. The family I_w , $w \in X$, is a
669 partition (up to sets of zero measure) of \mathcal{R} . If X is a finite S -maximal suffix
670 code, then the family J_w is a partition (up to sets of zero measure) of \mathcal{R} . Let
671 f be a coding morphism for X . If X is a finite S -maximal bifix code, then
672 E_X is the exchange of pieces E_f (defined as for interval exchanges), hence the
673 decoding of x with respect to f is the natural coding of the exchange of pieces
674 associated with f . In particular, S being the Tribonacci set, the decoding of S
675 by a finite S -maximal bifix code is again the natural coding of an exchange of
676 pieces. If X is the set of factors of length n of S , then E_f is in fact equal to R_β^n
677 (otherwise, there is no reason for this exchange of pieces to be a translation).
678 The analogues of Proposition 3.8 and 3.10 thus hold here also.

679 4.4 Subgroups of finite index

680 We denote by $FG(A)$ the free group on the set A .

681 Let S be a recurrent set containing the alphabet A . We say that S has the
682 *finite index basis property* if the following holds: a finite bifix code $X \subset S$ is an
683 S -maximal bifix code of S -degree d if and only if it is a basis of a subgroup of
684 index d of $FG(A)$.

685 The following is a consequence of the main result of [5].

686 **Theorem 4.5** *A regular interval exchange set has the finite index basis prop-*
687 *erty.*

688 *Proof.* Let T be a regular interval exchange transformation and let $S = F(T)$.
689 Since T is regular, S is uniformly recurrent and by Proposition 4.2, it is a tree
690 set. By Theorem 4.4 in [5], a uniformly recurrent tree set has the finite index
691 basis property, and thus the conclusion follows. ■

692 Note that Theorem 4.5 implies in particular that if T is a regular s -interval
693 exchange set and if X is a finite S -maximal bifix code of S -degree d , then
694 $\text{Card}(X) = d(s - 1) + 1$. Indeed, by Schreier's Formula a basis of a subgroup of
695 index d in a free group of rank s has $d(s - 1) + 1$ elements.

696 We use Theorem 4.5 to give another proof of Theorem 3.12. For this, we
 697 recall the following notion.

698 Let T be an interval exchange transformation on $I = [0, 1[$ relative to
 699 $(I_a)_{a \in A}$. Let G be a transitive permutation group on a finite set Q . Let
 700 $\varphi : A^* \rightarrow G$ be a morphism and let ψ be the map from I into G defined
 701 by $\psi(z) = \varphi(a)$ if $z \in I_a$. The *skew product* of T and G is the transformation
 702 U on $I \times Q$ defined by

$$U(z, q) = (T(z), q\psi(z))$$

703 (where $q\psi(z)$ is the result of the action of the permutation $\psi(z)$ on $q \in Q$).
 704 Such a transformation is equivalent to an interval exchange transformation via
 705 the identification of $I \times Q$ with an interval obtained by placing the $d = \text{Card}(Q)$
 706 copies of I in sequence. This is called an *interval exchange transformation on a*
 707 *stack* in [7] (see also [19]). If T is regular, then U is also regular.

708 Let T be a regular interval exchange transformation and let $S = F(T)$. Let
 709 X be a finite S -maximal bifix code of S -degree $d = d_X(S)$. By Theorem 4.5, X
 710 is a basis of a subgroup H of index d of $FG(A)$. Let G be the representation of
 711 $FG(A)$ on the right cosets of H and let φ be the natural morphism from $FG(A)$
 712 onto G . We identify the right cosets of H with the set $Q = \{1, 2, \dots, d\}$ with 1
 713 identified to H . Thus G is a transitive permutation group on Q and H is the
 714 inverse image by φ of the permutations fixing 1.

715 The transformation induced by the skew product U on $I \times \{1\}$ is clearly
 716 equivalent to the transformation $T_f = T_X$ where f is a coding morphism for the
 717 S -maximal bifix code X . Thus T_X is a regular interval exchange transformation.

718 **Example 4.6** Let T be the rotation of Example 3.1. Let $Q = \{1, 2, 3\}$ and let φ
 719 be the morphism from A^* into the symmetric group on Q defined by $\varphi(a) = (23)$
 720 and $\varphi(b) = (12)$. The transformation induced by the skew product of T and G
 721 on $I \times \{1\}$ corresponds to the bifix code X of Example 3.15. For example, we
 722 have $U : (1 - \alpha, 1) \rightarrow (0, 2) \rightarrow (\alpha, 3) \rightarrow (2\alpha, 2) \rightarrow (3\alpha - 1, 1)$ (see Figure 4.4)
 and the corresponding word of X is $baab$.

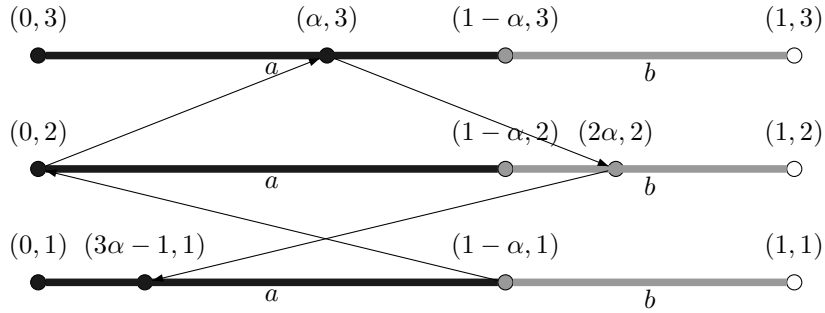


Figure 4.4: The transformation U .

723

724 **References**

- 725 [1] Vladimir I. Arnold. Small denominators and problems of stability of motion
726 in classical and celestial mechanics. *Uspehi Mat. Nauk*, 18(6 (114)):91–192,
727 1963. 2
- 728 [2] Jean Berstel, Clelia De Felice, Dominique Perrin, Christophe Reutenauer,
729 and Giuseppina Rindone. Bifix codes and Sturmian words. *J. Algebra*,
730 369:146–202, 2012. 2, 9, 10, 11, 12, 15
- 731 [3] Jean Berstel, Dominique Perrin, and Christophe Reutenauer. *Codes and*
732 *Automata*. Cambridge University Press, 2009. 5, 9, 11
- 733 [4] Valérie Berthé, Clelia De Felice, Francesco Dolce, Julien Leroy, Dominique
734 Perrin, Christophe Reutenauer, and Giuseppina Rindone. Maximal bifix
735 decoding. 2013. <http://arxiv.org/abs/1308.5396>. 2, 17, 18
- 736 [5] Valérie Berthé, Clelia De Felice, Francesco Dolce, Julien Leroy, Dominique
737 Perrin, Christophe Reutenauer, and Giuseppina Rindone. The finite index
738 basis property. 2014. <http://arxiv.org/abs/1305.0127>. 2, 21
- 739 [6] Valérie Berthé and Michel Rigo, editors. *Combinatorics, automata and*
740 *number theory*, volume 135 of *Encyclopedia of Mathematics and its Appli-*
741 *cations*. Cambridge University Press, Cambridge, 2010. 3, 8, 9
- 742 [7] Michael D. Boshernitzan and C. R. Carroll. An extension of Lagrange’s
743 theorem to interval exchange transformations over quadratic fields. *J. Anal.*
744 *Math.*, 72:21–44, 1997. 2, 22
- 745 [8] Isaac P. Cornfeld, Sergei V. Fomin, and Yakov G. Sinai. *Ergodic theory*,
746 volume 245 of *Grundlehren der Mathematischen Wissenschaften [Fundam-*
747 *ental Principles of Mathematical Sciences]*. Springer-Verlag, New York,
748 1982. Translated from the Russian by A. B. Sosinskiĭ. 3
- 749 [9] Sébastien Ferenczi and Luca Q. Zamboni. Languages of k -interval exchange
750 transformations. *Bull. Lond. Math. Soc.*, 40(4):705–714, 2008. 2, 17, 18
- 751 [10] N. Pytheas Fogg. *Substitutions in dynamics, arithmetics and combina-*
752 *torics*, volume 1794 of *Lecture Notes in Mathematics*. Springer-Verlag,
753 Berlin, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel.
754 7
- 755 [11] Jacques Justin and Laurent Vuillon. Return words in Sturmian and epis-
756 turmian words. *Theor. Inform. Appl.*, 34(5):343–356, 2000. 7
- 757 [12] Michael Keane. Interval exchange transformations. *Math. Z.*, 141:25–31,
758 1975. 2, 5, 8
- 759 [13] M. Lothaire. *Algebraic Combinatorics on Words*. Cambridge University
760 Press, 2002. 7, 9
- 761 [14] M. Lothaire. *Applied combinatorics on words*, volume 105 of *Encyclopedia*
762 *of Mathematics and its Applications*. Cambridge University Press, Cam-
763 bridge, 2005. 18
- 764 [15] V. I. Oseledec. The spectrum of ergodic automorphisms. *Dokl. Akad. Nauk*
765 *SSSR*, 168:1009–1011, 1966. 2
- 766 [16] Martine Queffélec. *Substitution dynamical systems—spectral analysis*, vol-
767 *ume 1294 of Lecture Notes in Mathematics*. Springer-Verlag, Berlin, second
768 edition, 2010. 19

- 769 [17] Gérard Rauzy. Échanges d'intervalles et transformations induites. *Acta*
770 *Arith.*, 34(4):315–328, 1979. 4, 18
- 771 [18] Gérard Rauzy. Nombres algébriques et substitutions. *Bull. Soc. Math.*
772 *France*, 110(2):147–178, 1982. 19, 20
- 773 [19] William A. Veech. Finite group extensions of irrational rotations. *Israel*
774 *J. Math.*, 21(2-3):240–259, 1975. Conference on Ergodic Theory and Topo-
775 logical Dynamics (Kibbutz Lavi, 1974). 2, 22