Journal de Théorie des Nombres de Bordeaux

Patrick Morton

# Connections between binary patterns and paperfolding

*Journal de Théorie des Nombres de Bordeaux*, tome 2, n° 1 (1990), p. 1-12
<http://www.numdam.org/item?id=JTNB_1990__2_1_1_0>

*Séminaire de Théorie des Nombres,*
*Bordeaux* **2** *(1990), 1–12*

# Connections between binary patterns and paperfolding.

par Patrick MORTON

**1. Introduction.** This talk will be an overview of some recent work on binary patterns and their relationship to paperfolding sequences. Much of this work was motivated by the Rudin-Shapiro sequence $\{a_{11}(n)\}$ defined by

$$a_{11}(n) = (-1)^{e_{11}(n)},$$

where $e_{11}(n)$ is the number of occurrences of the pattern 11 in the binary representation of $n$. On the one hand, this definition is easy to generalize. One can consider the analogous sequence

$$(1) \qquad a_P(n) = (-1)^{e_P(n)},$$

where $P$ is any pattern of 0's and 1's and $e_P(n)$ counts occurrences of $P$ in $n$.

On the other hand, there is the beautiful fact discovered by Mendès France that $a_{11}$ is exactly the direction sequence of the paperfolding sequence obtained by folding a rectangular piece of paper alternately under and over the left edge (which is held fixed). (See [2] or [4].)

This raises the question: is this result about $a_{11}$ isolated or does it point to some deeper connection between binary patterns and paperfolding?

**2. Properties of $a_P(n)$.** The results of this section represent joint work with David Boyd and Janice Cook. (See [1].)

Let me begin by recalling some of the properties of the sequence $a_P(n)$ defined by (1), where $P$ is any pattern of 0's and 1's. First define

$$(2) \qquad s_P(x) = \sum_{k \leq x} a_P(k).$$

Then the following theorem holds concerning the behavior of $s_P(x)$.

**Manuscrit reçu le 20 juin 1989**

THEOREM 1. *(See [1].) The partial sum function $s_P(x)$ has the representation*

$$s_P(x) = t(x) + s_0(x),$$

*where $t(x)$ is continuous and $s_0(x)$ is bounded. Further,*

$$t(x) = \sum_{|\xi_k|>1} \lambda_k(x) x^{\log \xi_k / \log 2},$$

*where the $\xi_k$ are the roots outside the unit circle of the polynomial*

$$(3) \qquad P(x) = (x-2)(x^{d-1} + 2\pi_1 x^{d-2} + \cdots + 2\pi_{d-1}) + 2,$$

*and*

$$\pi_k = \begin{cases} 1 \text{ if } k \text{ is a period of } P, \\ 0 \text{ if not.} \end{cases}$$

*(By definition $k$ is a period of $P = p_{d-1} \cdots p_0$ if and only if $p_i = p_{i+k}$ for $0 \leq i \leq d-1-k$.)*

*Finally, the $\lambda_k(x)$ are continuous functions of $x > 0$ satisfying $\lambda_k(2x) = \lambda_k(x)$.*

As a corollary we have

$$s_P(x) = O(x^\tau),$$

where $\tau = \log\xi / \log 2$ and $\xi$ is the maximum modulus of the roots of $P(x)$. For all but 14 patterns (in particular, if $d \geq 5$) $\xi = \xi_1$ is a root of $P(x)$.

From this follows the formula

$$\lim_{k\to\infty} \frac{s_P(2^k x)}{(2^k x)^\tau} = \lambda_1(x),$$

so in these cases $\lambda(x) = \lambda_1(x)$ (for $1 \leq x \leq 2$) represents the limiting behavior of $s_P(x)/x^\tau$ on the intervals $[2^k, 2^{k+1}]$, as $k \to \infty$. This limit function possesses two rather nice properties:

1) $\lambda(x)$ is nondifferentiable almost everywhere,

2) $x \in Q \Rightarrow x^\tau \lambda(x) \in Q(\xi)$.

Property 1) also implies that $s_P(x) = \Omega(x^\tau)$, so that $x^\tau$ is the correct order of magnitude of $s_P(x)$.

The polynomials $P(x)$ in theorem 1 also satisfy a remarkable set of recursions. In order to state these recursions, let $P$ and $P'$ be patterns of length $d$ and $d+1$ recpectively, and let $\Pi$ and $\Pi'$ denote their sets of periods, e.g.

$$k \in \Pi \text{ if and only if } k \text{ is a period of } P.$$

If $P(x)$ and $P'(x)$ denote the corresponding polynomials, as in (3), then $P(x)$ and $P'(x)$ depend only on the period sets $\Pi, \Pi'$, and for every period set $\Pi'$ there is a unique period set $\Pi$ for which

$$P'(x) = \begin{cases} xP(x) - 2, & \text{if } \Pi' = \Pi \cup \{d\} \\ xP(x) - 2x + 2 & \text{if } \Pi' = \Pi. \end{cases}$$

We called these formulae the red and blue rules, respectively. They show that the polynomials $P(x)$ form a tree, with red or blue branches depending on the rule which connects $P(x)$ to $P'(x)$. In [1] we show that this tree has a fractal-like structure. It has a sequence of periodic subtrees which are isomorphic to larger and larger initial pieces of the whole tree; thus the tree reproduces itself in its subtrees.

The sequences $a_P$ are very fundamental. They can be used to study arbitrary sequences of $\pm 1$'s. If $a = \{a(n)\}_{n \geq 0}$ is any sequence of $\pm 1$'s, then there is a unique sequence $\{P_k\}$ of binary patterns with increasing values (the value of a pattern is the integer it represents) and no leading zeros, for which

$$a = a(0) \prod_{k=1}^{\infty} a_{P_k}.$$

This infinite product is defined using the topology in which two sequences are close if a large number of their initial terms agree. Let us call the set $\{P_k\}$ the binary spectrum of the sequence $a$.

Now consider the question: is it possible to characterize the sequences $a$ which have a *finite* spectrum?

## 3. The arithmetic fractal groups $\Gamma_k(G)$.

The answer to the last question is in fact yes, and depends on the following definition. (See Morton and Mourant [4] for the results of this section.)

DEFINITION. *Let $G$ be an abelian group, $a = \{a(n)\}$ an infinite sequence of elements of $G$, and*

$$X_n^q = (a(k^q n), a(k^q n + 1), \ldots, a(k^q n + k^q - 1))$$

*the associated sequence of $k^q$-segments of $a$. These vectors partition the sequence $a$ into vectors of length $k^q$. We define*

$$(4) \qquad \Gamma_k(G) = \{a : a^{-1}(n)X_n^q \textit{ is periodic with period } M \textit{ for all } q \geq 0\}$$

*Here $a^{-1}(n)X_n^q$ is defined using scalar multiplication (or scalar addition if the operation in $G$ is addition). The least period $M$ of a sequence $a$ in $\Gamma_k(G)$ is called its* conductor.

The set $\Gamma_k(G)$ is an abelian group under componentwise multiplication.

As an example, let $P$ be a pattern of digits base $k$ and consider the sequence $e_p(n)$ which counts the number of occurrences of $P$ in the base-$k$ representation of $n$. Then $e_P \in \Gamma_k(Z)$ with conductor dividing $k^{d-1}$, where $d$ is the number of digits of $P$. (This holds as long as $P$ is not a pattern of all 0's). This fact follows from the equation

$$X_n^q - e_P(n) = X_m^q - e_P(m), \text{ if } n \equiv m \ (mod\ k^{d-1}).$$

If $k = 2$, reducing $e_1$ modulo 2 gives the Thue-Morse sequence $e_1^*$, which satisfies

$$X_n^q = X_0^q + e_1^*(n) \ (\text{mod } 2);$$

hence $e_1^* \in \Gamma_k(Z_2)$. This last formula is a restatement of the familiar fact that $e_1^*$ is invariant under the substitution

$$0 \to 01, \ 1 \to 10.$$

The corresponding equation for the sequence of *integers* $e_1$ also holds in characteristic 0:

$$X_n^q = X_0^q + e_1(n).$$

The related sequences $a_P = \zeta^{e_P}$, where $\zeta$ is a root of unity, generate a special subgroup of $\Gamma_k(<\zeta>)$, namely the subgroup

$$\Lambda_k(<\zeta>) = \{a \in \Gamma_k(<\zeta>) \text{ with conductor dividing a power of } k\},$$

which provides an answer to the question we raised in section 2.

THEOREM 2. *(See [4].) The sequences in $\Lambda_2(\pm 1)$ are exactly the sequences which have a finite binary spectrum.*

Of course an analogous result characterizes sequences in $\Lambda_k(<\zeta>)$ in terms of their base $k$ spectra.

The sequences in $\Gamma_k(G)$ can be thought of as arithmetic analogues of fractals. As an example consider the Rudin-Shapiro sequence $a_{11}$, whose first four terms are

$$1,1,1,-1.$$

We have

$$X_2^0 = X_0^0 \text{ and } X_3^0 = -X_1^0. \tag{5}$$

Since $a_{11} \in \Gamma_2(\pm 1)$ with $M = 2$, the definition (4) implies that the relations (5) persist for $2^q$-segments at all levels:

$$X_2^q = X_0^q \text{ and } X_3^q = -X_1^q \text{ for } q \geq 0. \tag{6}$$

This can be used to generate $a_{11}$ by considering the first four terms as the elements of $X_0^1$ and $X_1^1$. Using (6) with $q = 1$ gives the next four terms of the sequence:

$$11, 1-1, 11, -11 = X_0^1, X_1^1, X_2^1, X_3^1.$$

For $q = 2$ we then get 8 more terms :

$$111-1, 11-11, 111-1, -1-11-1,$$

etc. The same rule connects $2^q$ segments at ever increasing levels, in analogy to the familiar constructions used to generate certain types of geometric fractals. The same remarks hold for any sequence in $\Gamma_k(G)$ for all $k$ and $G$.

**4. Properties of $\Gamma_k(G)$.** Among the properties of these groups I want to particularly mention three. For proofs of 1. and 2. see [4].

1. In the definition of $\Gamma_k(G)$ we only need to require that $a^{-1}(n)X_n^1$ be periodic of period $M$. The fact that $a^{-1}(n)X_n^q$ has period $M$ for all $q \geq 0$ follows.

2. If $G = U$ is the group of complex roots of unity, $\Gamma_k(U)$ contains an isomorphic copy of every finite abelian group. Every finite abelian group is a homomorphic image of $\Gamma_k(Z)$, for any $k$.

3. If $G$ is finite, the sequences in $\Gamma_k(G)$ are all $k$-automatic.

I will prove this last property using the following maps $T_i$ on sequences:

$$(T_i a)(n) = a(kn+i), \ 0 \leq i \leq k-1.$$

From property 1. above we have the following characterization of sequences in $\Gamma_k(G)$.

LEMMA. *The sequence $a$ lies in $\Gamma_k(G)$ if and only if $a^{-1}T_ia$ is periodic for all $i = 0, 1, \cdots, k-1$.*

**Proof.** The sequence $a^{-1}(n)X_n^1$ is periodic if and only if all its component sequences are periodic, and its $i$-th component is just $a^{-1}T_ia$.

If $T_{i_1}, ..., T_{i_r}$ are any $r$ of these maps (with possible repetitions), then obviously

$$(T_{i_1} \cdots T_{i_r} a)(n) = a(k^r n + k^{r-1} i_r + k^{r-2} i_{r-1} + \cdots + i_1).$$

Hence we can use these maps to characterize sequences in $\Gamma_{k^q}(G)$ : these are sequences $a$ for which $a^{-1}Ta$ is periodic for all monomials $T$ in $T_0, \cdots, T_{k-1}$ of degree $q$.

As is well-known, a sequence $a$ is $k$-automatic if and only if the set of sequences $\{a(k^q n + i)\}$ is a finite set. Using the maps $T$ we can state

*$a$ is $k$-automatic if and only if there are only finitely many images $T_a$, as $T$ runs over all monomials in $T_0, \cdots, T_{k-1}$.*

Using this we prove

THEOREM 3. *If $G$ is a finite abelian group, the sequences in $\Gamma_k(G)$ are all $k$-automatic.*

**Proof.** By definition, $a \in \Gamma_k(G)$ has the property that $a^{-1}T_ia$ is periodic for $0 \le i \le k-1$. Hence

$$T_ia = ap_i,$$

where $p_i$ is a periodic sequence, of period $M$, say. If $S$ is any monomial in $T_0, \cdots, T_{k-1}$, and $S = S'T_i$, then

$$Sa = S'T_i(a) = S'(ap_i) = S'(a)S'(p_i) = S'(a).p',$$

where $p' = S'(p_i)$ is also periodic with period M. Peeling off one term $T_i$ at a time gives finally that

$$Sa = ap'p'' \cdots = a\tilde{p},$$

where $\tilde{p}$ is periodic with period $M$. But there are only finitely many periodic sequences of a given period with terms that are taken from the given finite set $G$. Hence there are only finitely many distinct sequences $Sa$, implying that $a$ is $k$-automatic.

A different proof of this theorem appears in [5], where we show that the sequences in $\Gamma_k(G)$ are essentially fixed points of $k^r$-substitutions for suitable $r$. Shallit (private communication) has found yet a third proof.

If $A_k(G)$ is the set of all $k$-automatic sequences taken from the alphabet $G$, then it is a consequence of theorem 3 that

$$\Gamma_{k^q}(G) \subset A_k(G), \text{ for all } q \geq 0,$$

and so

$$\varinjlim \Gamma_{k^q} \subset A_k(G).$$

It is not hard to exhibit automatic sequences which are not in $\Gamma_{k^q}$ for any $q$. We do this for $q = 2$ in the following

**Example.** Let $G = Z_2$ and let $u$ be the Baum-Sweet sequence, defined by

$$\begin{aligned} u_0 &= 1, \\ u_{2n+1} &= u_n, \\ u_{4n} &= u_n, \\ u_{4n+2} &= 0. \end{aligned}$$

These formulae may be written as follows:

$$T_1 u = u, \; T_0^2 u = u, \; T_1 T_0 u = 0.$$

It is easy to check that

$$\{Su; S \text{ a monomial in } T_0, T_1\} = \{u, T_0 u, 0\}.$$

To show that $u \notin \Gamma_{2^q}(Z_2)$ for any $q$, we show that $-u + Su = u + Su$ is not periodic, for a suitable monomial $S$ of degree $q$. It suffices to take $q \geq 2$ and $S = T_1^{q-1} T_0$. Then

$$u + Su = u + T_1^{q-1} T_0 u = u + T_1^{q-2}(T_1 T_0 u) = u,$$

which is not periodic, since the series $\sum_{n=0}^{\infty} u_n x^n$ satisfies an irreducible cubic equation over $Z_2[x]$. Thus

$$u \in A_2(Z_2) - \varinjlim \Gamma_{2^q}(Z_2).$$

These considerations also raise the following question :

**Question.** Since $\Gamma_k(G)$ and $A_k(G)$ are abelian groups, they have associated rings of endomorphisms. What are

$$E_1 = \text{ End } (\Gamma_k(G)) \text{ and } E_2 = \text{ End } (A_k(G))?$$

It is obvious that $T_i \in E_2$ and not hard to check that $T_i \in E_1$ for all $i = 0, 1, \cdots, k-1$. Since the $T_i$ do not commute this shows that $E_1$ and $E_2$ are non-commutative rings.

**5. Paperfolding sequences.** I turn now to the connection between binary patterns and paperfolding sequences. First a little notation. I want to consider paperfolding sequences corresponding to an infinite sequence of folding instructions

$$w = \phi_1\phi_2 \cdots \phi_n \cdots,$$

where $\phi_n = o$ or $u$ and $o, u$ denote the operations of folding a rectangular piece of paper respectively over or under the left edge, (see [2],[3]). If $w_m$ is the finite initial word of $w$ of length $m$, then there are two sequences of $\pm 1$'s associated to $w_m$.

The first, $f_{w_m}$, encodes the sequence of up or down folds obtained by unfolding the paper after applying $w_m$, according to the rules

$$\vee \to +1$$

$$\wedge \to -1.$$

The second, denoted $d_{w_m}$, encodes horizontal and vertical directions in the plane curve obtained by making all the folds equal to right angles and looking at the paper edge-on. In $d_{w_m}$, horizontal and vertical directions are coded by

$$\begin{array}{ll} \rightarrow +1 & \leftarrow -1, \\ \uparrow +1 & \downarrow -1. \end{array}$$

For example,

$$f_{oou} = -1, 1, 1, 1, -1, -1, 1$$
$$d_{oou} = 1, -1, 1, 1, -1, 1, 1, 1.$$

The limit points of the sets $\{f_{w_m}, m \geq 1\}, \{d_{w_m}, m \geq 1\}$ in the set of all sequences of $\pm 1$'s are respectively called paperfolding sequences and direction sequences corresponding to the word $w$. The topology with respect to which these limits are to be taken is the topology in which two sequences are close if a large number of their initial terms agree. It is not hard to show that a subsequence of $\{f_{w_m}\}$ or $\{d_{w_m}\}$ converges in this topology if and only if the subsequence of corresponding words $w_m$ is *reverse*-convergent, i.e. if and only if an increasing number of the final instructions of the $w_m$ agree. Using this fact it is easy to see that paperfolding sequences and direction sequences of $w$ are in 1-1 correspondence with the reverse infinite words

$$\omega = \cdots \phi_n \cdots \phi_2 \phi_1, \; \phi_n = o \text{ or } u,$$

for which $\phi_n \cdots \phi_1$ are subwords of $w$. This sequence is called a sequence of *unfolding* instructions, (see [2] or [3]). Let me denote the paperfolding and direction sequences corresponding to $\omega$ by $f_\omega$ and $d_\omega$. If one defines

$$s_\omega(n) = \sum_{j=1}^{n} f_\omega(j), \; s_w(0) = 0,$$

then one has

$$d_\omega(n) = i^{s_\omega(n)+n^2+n}, n \geq 0,$$

and the following result holds:

THEOREM 4.. *(See [4], theorem 13.) Let the reverse infinite sequence $\omega$ be periodic, with period $\lambda$. Then*

$$s_\omega \in \Gamma_{2^\lambda}(Z) \text{ and } d_w \in \Gamma_{2^\lambda}(\pm 1).$$

In fact $s_\omega$ and $d_\omega$ have conductor 2, so they lie in the respective subgroups $\Lambda_{2^\lambda}(Z)$ and $\Lambda_{2^\lambda}(\pm 1)$. It follows by theorem 2 and analogous results that $s_\omega$ and $d_\omega$ are respectively a sum and product of pattern counting sequences. For example,

$$s_{(oou)^\infty} = -e_1 + e_3 + 2e_4 + e_5 + e_7 + \sum_{i,j=0}^{7} c_{ij} e_{ij},$$

where

$$c_{ij} = \begin{cases} 1, & i \text{ odd}, 0 \leq j \leq 3, \\ -1, & i \text{ odd}, 4 \leq j \leq 7, \\ 0, & \text{otherwise}, \end{cases}$$

and the digits $i, j$ in this formula are taken base 8. This leads to the representation

$$d_{(oou)^\infty} = a_1 a_4 \prod_{i \text{ odd}, \; 0 \leq j \leq 3} a_{ij},$$

where $a_P = (-1)^{e_P}$.

Theorem 4 raises the question: for which unfolding sequences $\omega$ is $d_\omega \in \Gamma_{2^\lambda(\pm 1)}$?

We phrase the answer to this question in terms of the map $\tau$ defined as follows. Let $D_w$ denote the set of direction sequences which arise from a given infinite word $w$, and set

$$\tau(d_{\omega\phi}) = d_\omega, \ \phi = o \text{ or } u.$$

Then $\tau : D_\omega \to D_w$ and in [4] we show that

$$d_{\omega\phi} = d_\omega(0)X, d_\omega(1)Y, d_\omega(2)X, d_\omega(3)Y, \ldots ,$$

for two fixed vectors $X, Y$ of length two; hence the sequence $\tau(d_{\omega\phi})$ is a "sign sequence" for $d_{\omega\phi}$ with respect to segments of length two. By iterating the map $\tau$ we prove that $\tau^k(d_\omega)$ is a sign sequence for $d_\omega$ with respect to segments of length $2^k$, and this fact leads to the following result, (see [4], theorems 17-18):

THEOREM 5.

*1) The sequence $d_\omega$ lies in $\Gamma_{2^\lambda}(\pm 1)$ if and only if $d_\omega \tau^\lambda(d_\omega)$ is a periodic sequence.*

*2) If $d_\omega \in \Gamma_{2^\lambda}(\pm 1)$, then $d_\omega \in \Lambda_{2^{\lambda'}}(\pm 1)$, for a suitable multiple $\lambda'$ of $\lambda$.*

*3) $d_\omega \in \Lambda_{2^{\lambda'}}(\pm 1)$ if and only if*

$$\tau^{\lambda'+n}(d_w) = \tau^n(d_w),$$

*for some $n \geq 0$.*

Hence direction sequences in $\Gamma_{2^\lambda}$ correspond to pre-periodic points of the map $\tau$, and the groups $\Gamma_{2^\lambda}(\pm 1)$ arise naturally in connection with the discrete dynamical system $(D_w, \tau)$.

Using the definition of $\tau$ and replacing $\lambda'$ by $\lambda$ in theorem 5, part 3) gives.

THÉORÈME 6. *(See [4], theorem 19.) The direction sequence $d_\omega$ is a finite product of pattern sequences base $2^\lambda$ (and so has a finite spectrum base $2^\lambda$) if and only if*

$$\omega = (w_1)^\infty w_0, \; w_1 \textit{ of length } \lambda,$$

*or*

$$\omega = (\overline{w_2} w_2)^\infty w_0, \; w_2 \textit{ of length } \lambda,$$

*where $\overline{w_2}$ is obtained from $w_2$ by switching $o$'s and $u$'s.*

As examples with $\lambda = 1$ let me note

$$d_{o^\infty} = a_{10}, \; d_{u^\infty} = a_1 a_{11}, \; d_{(uo)^\infty} = a_{11}, \; d_{(ou)^\infty} = a_1 a_{10},$$

where the patterns in these formulae are all binary patterns. From these relations it is possible to compute the spectra of all direction sequences which lie in $\Lambda_2(\pm 1)$. For example,

$$d_{(ou)^\infty o^k} = a_{10} \prod_P{}' a_P,$$

where the product is over all binary patterns beginning with 1 and having length $k+1$. This shows that every binary pattern (with a leading 1) occurs in the binary spectrum of some $d_\omega$.

I will finish by recalling a recent theorem of Mendès France and Shallit [3] (see their theorem 4.2 for the special case of paperfolding; note that their sequence of turns is here what we have called a paperfolding sequence):

THEOREM. *The paperfolding sequence $f_w$ is 2-automatic if and only if the sequence $\omega$ of unfolding instructions is ultimately periodic.*

Together with theorem 6, this gives

THEOREM 7. *A paperfolding sequence $f_\omega$ is 2-automatic if and only if its associated direction sequence $d_\omega$ has a finite spectrum base $2^\lambda$, for some $\lambda$.*

ACKNOWLEDGEMENTS : The author gratefully acknowledges the support of the Alexander von Humboldt Foundation, the C.N.R.S., the D.R.E.T. and the Mathematics Departements at the University of Bordeaux, the University of Paris-Sud and the University of Heidelberg during the period this article was written.

## REFERENCES

[1] D.W. BOYD, J. COOK and P. MORTON, *On sequences of* $\pm 1's$ *defined by binary patterns.* Dissertationes Math. **283**, to appear.

[2] M. DEKKING, M. MENDÈS FRANCE and A. van der POORTEN, *Folds I,II, III,* Math. Intelligencer 4 (1982), 130–138, 173–181, 190–195.

[3] M. MENDÈS FRANCE and J. O. ŜHALLIT, *Wire bending,* J. Comb. Theory **50** (1989), 1–23.

[4] P. MORTON and W.J. MOURANT, *Paper folding, digit patterns and groups of arithmetic fractals,* Proc. London Math. Soc. **59**, **2**, (1989), 253-293.

[5] P. MORTON and W.J. MOURANT, *Digit patterns and transcendental numbers.* submitted.

*Mots clefs*: Paperfolding, occurrences of words, fractals.

1980 *Mathematics subject classifications*: 10A30, 10L10, 68D22.

Mathematisches Institut, Universität Heidelberg
Heidelberg (GERMANY)
and
Dept. of Mathematics, Wellesley College,
Wellesley, MA 02181 (U.S.A.).