POLSKA AKADEMIA NAUK, INSTYTUT MATEMATYCZNY

DISSERTATIONES MATHEMATICAE

(ROZPRAWY MATEMATYCZNE)

CCLXXXIII

DAVID W. BOYD, JANICE COOK and PATRICK MORTON

On sequences of ± 1's defined by binary patterns

WARSZAWA 1989
PAŃSTWOWE WYDAWNICTWO NAUKOWE

PRINTED IN POLAND

ISBN 83-01-09037-5 ISSN 0012-3862

WROCŁAWSKA DRUKARNIA NAUKOWA

CONTENTS

1. Introduction

The Rudin–Shapiro sequence $\{a(n)\}$, defined by the recursive rules

$$a(2n) = a(n), \quad a(2n+1) = (-1)^n a(n), \quad n \geqslant 0; \quad a(0) = 1,$$

has been studied extensively since its introduction by Shapiro [33] in 1951. Recently Mendes France, and others [9], [26], [27], have shown that the Rudin–Shapiro sequence can be generated by an infinite sequence of alternating folds of an ordinary piece of paper. More generally they have shown that an arbitrary sequence of folds gives rise to so-called "direction sequences" $u_n = \pm 1$, which have the property that

$$\max_{0 \leqslant \theta < 1} \Big| \sum_{n=0}^{N} u_n e^{2\pi i n\theta} \Big| \leqslant c\sqrt{N}, \quad N \geqslant 1.$$

This is a generalization of the same property first discovered by Shapiro [33] and later Rudin [32] for the sequence $\{a(n)\}$. (See also Littlewood [22].)

Rider [31] and Brillhart and Carlitz [2] were the first to recognize $a(n)$ as coming from the binary representation of n. Brillhart and Carlitz derived the above definition of $a(n)$, and showed that $a(n)$ is plus or minus one according as the number of occurrences of the pattern 11 in the binary representation of n is even or odd. (In [7] this property is interpreted by means of automata.) On the basis of this definition special arithmetic properties of the partial sums

$$s(x) = \sum_{k=0}^{[x]} a(k)$$

were considered in [5] and [6].

In this paper we follow a suggestion of the late Ernst Straus given to us in 1980, and study the sequence $a_p(n)$ for a general pattern P of 0's and 1's (of length $d \geqslant 2$). The term $a_p(n)$ equals $(-1)^{e_p(n)}$, where $e_p(n)$ is the number of occurrences of the pattern P in the binary representation of n. In this definition we allow leading zeros, so that for example $a_{0011}(6) = -1$, $a_{0011}(51) = +1$. If P is a string of 0's, this causes some trouble; in this case we use the convention that an occurence of P is to be counted only if it preceded by a 1. This

---

1980 *Mathematics Subject Classification*: Primary *10A30, 10H25.*

Key words: binary patterns, partial sums, period polynomials, autocorrelation tree.

definition can easily be expressed using a recursion similar to the definition of $a_{11}(n)$ given above. (See Section 2, (2).) These sequences have also been considered in [7], [28], and [1], where they are associated with some remarkable infinite products.

We focus for the most part on the partial sums

$$s(x) = s_p(x) = \sum_{k=0}^{[x]} a_p(k),$$

and we prove that $s_p(x) = O(x^\tau)$, where $\frac{1}{2} \leqslant \tau < 1$ and $\tau \to 1$ as the number of digits d of P tends to ∞. More specifically, $\tau = \log\xi/\log 2$, where ξ is an algebraic number: $\tau = \frac{1}{2}$ when $d = 2$, and $\tau > \frac{1}{2}$ when $d \geqslant 3$. The number τ is easily seen to be transcendental if $d \geqslant 3$. (For a proof see the appendix.)

The algebraic number ξ arises in the following manner. We show that the subsums

$$s_p(x, i) = \sum_{\substack{k=0 \\ k \equiv i (\mathrm{mod} M)}}^{[x]} a(k) \qquad (M = 2^{d-1})$$

form a vector $\bar{s}(x) = (s_p(x, 0), \ldots, s_p(x, M-1))^t$ which satisfies the recursion

$$\bar{s}(2x) = A\bar{s}(x) + \bar{j}(x),$$

where $\bar{j}(x)$ is a bounded vector and A is an $M \times M$ matrix of a very simple type (see Section 2). The number ξ is the maximum modulus of the eigenvalues of A. In all but 14 early cases, ξ is the unique dominant eigenvalue of A, a fact which is proved in Section 5 using Rouche's theorem.

The upper bound for $s_p(x)$ follows from the explicit formula

$$s_p(x) = \sum_{k=1}^{d} \lambda_k(x) x^{\log\xi_k/\log 2} + s_0(x),$$

where $s_0(x)$ is bounded, the ξ_k are the nonzero eigenvalues of A, and the $\lambda_k(x)$: $\boldsymbol{R}^+ \to \boldsymbol{C}$ are bounded continuous functions satisfying $\lambda_k(2x) = \lambda_k(x)$. (See Section 7.)

In the case that $\xi_1 = \xi$ is positive and dominates the ξ_k, the above formula implies that

$$\lim_{n\to\infty} \frac{s(2^n x)}{(2^n x)^\tau} = \lambda_1(x) = \lambda(x)$$

for all $x > 0$. This is the analogue of the limit considered in [6]. See also [8] and [10].

The curve $y = \lambda(x)$, for $1 \leqslant x \leqslant 2$, represents the limiting behavior of the function $s(x)/x^\tau$ on the intervals $2^n \leqslant x \leqslant 2^{n+1}$, as $n \to \infty$, and has several striking properties. In the first place, $\lambda(x)$ is non-differentiable almost everywhere. Secondly the function $x^\tau\lambda(x)$ maps $\boldsymbol{Q}^+$ into $\boldsymbol{Q}(\xi)$. Both properties

follow from an explicit formula for $\lambda(x)$. (See Section 8, (69).) As a consequence of this formula, the sum $s_p(x, i)$ over the residue class $i(\mathrm{mod}\, 2^{d-1})$ has the same order of magnitude as $s_p(x)$ whenever the i-th component of an eigenvector of A corresponding to the eigenvalue ξ is nonzero.

We conjecture that the curve $y = \lambda(x)$ is a fractal (see [24]), with Hausdorff dimension equal to $2-\tau$, but so far we have been unable to prove this.

In particular, $\lambda(x)$ is not always 0, and this implies that $s(x) = \Omega(x^\tau)$, i.e. $s(x) \neq O(x^{\tau-\varepsilon})$ for any $\varepsilon > 0$. For the 14 exceptional cases in which ξ is not the unique dominant eigenvalue of A, we prove $s(x) = \Omega(x^\tau)$ in Section 10 by computing $s(2^n)$. For $d \geqslant 3$, this fact shows, since $\tau > \frac{1}{2}$, that the sequence $\{a_p(n)\}$ does not arise as the direction sequence of a paper-folding sequence. On the other hand all four sequences $a_p(n)$ for which $d = 2$ do arise from paper folding. This fact was pointed out to us by Wanda Mourant (private communication). (See [28].)

Thus for all but 4 patterns, $s(x) \neq O(\sqrt{x})$. This is interesting in view of the fact that almost all sequences of ± 1's are known to have partial sums which are $cx^{1/2+\varepsilon}$ in size.

In Section 9 we consider the special case $P = 111$. Here $\lambda(x)$ is always positive, from which it follows that

$$c_1 x^\tau \leqslant s(x) \leqslant c_2 x^\tau, \qquad \tau = .823172455\ldots,$$

for positive constants c_1 and c_2; in this case ξ is the largest root of $x^3-2x-2 = 0$.

Finally, the nonzero eigenvalues of the matrix A satisfy the irreducible polynomial

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

where $\pi_k = 0$ or 1 and $\pi_k = 1$ iff k is a period of the pattern P (See Section 3). This result implies that $P(x)$ – and therefore the order of magnitude of $s_p(x)$ – depends only on the periods of P, or in other words, on the "autocorrelation" of the pattern P, introduced by Guibas and Odlyzko in [13]. This is the digit string

$$\alpha = \pi_{d-1}\pi_{d-2}\ldots\pi_1\pi_0, \qquad \pi_0 = 1,$$

written backwards from the convention adopted in [13].

We call the polynomial $P(x)$ the *period polynomial* of the pattern P. From the characterization of autocorrelations proved in [13] it follows that the polynomials $P(x)$ satisfy a set of recursions (see Section 4): if P' is a pattern of length $d+1$, with associated polynomial $P_{d+1}(x)$, then

$$P_{d+1}(x) = \begin{cases} xP_d(x)-2x+2 & \text{(B)}, \\ xP_d(x)-2 & \text{(R)}, \end{cases}$$

for a unique period polynomial $P_d(x)$ corresponding to some pattern P of length d. Therefore the period polynomials $P_d(x)$ form a tree, with blue (B) or red (R) edges according to the rule which leads from $P_d(x)$ to $P_{d+1}(x)$.

Equivalently, the autocorrelations α of all binary patterns form an infinite tree T, whose structure we investigate in Section 11. It turns out that T is the union of an infinite number of disjoint periodic subtrees T_k, each of which has an initial subtree isomorphic to the first $k-1$ "levels" of T. (The level of a correlation is d, the number of its digits. See Section 11, Theorem 14.) Thus the tree T reproduces itself to a greater and greater extent in its subtrees T_k. We also determine which branches of T terminate and which are infinite. (See Figures 4–9 in Section 11.)

In conclusion we refer the reader to the papers [12] and [14], in which the properties of binary patterns are considered from a different point of view. In these papers the autocorrelations of digit patterns play a central role.

We would like to thank Wanda Mourant for her help in proving Theorem 7 of Section 6, and Andrew Odlyzko for the proof of Theorem 4 in Section 4. We are also grateful to Lou Groner for his help with APL and APL-2 programming, to Chris Moller for providing interfaces to several graphics devices, to Helen Mason for producing Figures 2, 3 and 7–9 to Ron and Wanda Mourant for producing Figures 1 and 4–6; and last but not least, to the Muses who inspired this work.

2. The matrix recursion

As in the introduction we let d be an integer $\geqslant 2$, and we fix our attention on a binary bit pattern P with d digits and possible leadings zeros. We let p be the non-negative integer corresponding to P, and set

$$p = 2b+e, \quad e = 0 \text{ or } 1, \quad 0 \leqslant b \leqslant 2^{d-1}-1. \tag{1}$$

We define the numbers $a_p(n)$ by

DEFINITION. $a_p(0) = 1$, and for $r = 0$ or 1,

$$a_p(2n+r) = \begin{cases} -a_p(n), & \text{if } n \equiv b (\text{mod } 2^{d-1}) \text{ and } r = e, \\ a_p(n), & \text{otherwise,} \end{cases} \tag{2}$$

for $n \geqslant 0$, or for $n \geqslant 1$ if $p = 0$ and $r = 0$.

This definition implies that $a_p(2n+r)$ agrees with $a_p(n)$ unless the last d binary digits of $2n+r$ coincide with those of P, and so $a_p(n)$ counts the parity of the number of occurences of P in n (allowing for possible leading zeros in P). A modification needs to be made in this rule when P is a string of zeros, in which case we consider that P occurs in n only if a 1 occurs as a digit

somewhere in front of P. By (2) and the rule $a_p(0) = 1$, this amounts to the convention that P does not occur in the integer 0. For example, if $P = 00$, then $a(2) = a(10_2) = 1$, $a(4) = a(100_2) = -1$, $a(8) = +1$.

We will usually write $a(n)$ for $a_p(n)$ where no confusion is possible and we let

$$s_p(x) = s(x) = \sum_{k \leqslant x} a_p(k) = \sum_{k=0}^{[k]} a_p(k) \tag{3}$$

be the partial sums of the sequence $\{a(n)\}$. We also introduce the subsums

$$s_p(x, i) = s(x, i) = \sum_{\substack{k \leqslant x \\ k \equiv i (\mathrm{mod}\, M)}} a(k), \tag{4}$$

where $M = 2^{d-1}$ and the sum is over all integers $k \equiv i \pmod{M}$, for a fixed i with $0 \leqslant i \leqslant M-1$. With these sums we define the M-dimensional column vector

$$\bar{s}_p(x) = \bar{s}(x) = \big(s(x, 0), s(x, 1), \ldots, s(x, 2^{d-1}-1)\big)^t. \tag{5}$$

It turns out that $s(x)$ satisfies a recursion formula which is most conveniently stated in matrix terms, using $\bar{s}(x)$. This is the content of

THEOREM 1. *For any real number* $x \geqslant 0$,

$$\bar{s}(2x) = A\bar{s}(x) + \bar{j}(x), \tag{6}$$

where A is a $2^{d-1} \times 2^{d-1}$ matrix of 0's, 1's *and a single* -1, *and $\bar{j}(x)$ is a column vector which takes on a finite set of integer-vector values.* (See (13) *and* (16).)

To determine the matrix A and the "junk term" $\bar{j}(x)$, we first assume $p \neq 0$, and we consider the $2i$-th entry of $\bar{s}(2x)$:

$$\begin{aligned} s(2x, 2i) &= \sum_{\substack{k \leqslant 2x \\ k \equiv 2i (\mathrm{mod}\, M)}} a(k) = \sum_{\substack{2k \leqslant 2x \\ 2k \equiv 2i (\mathrm{mod}\, M)}} a(2k) \\ &= \sum_{\substack{k \leqslant x \\ k \equiv i (\mathrm{mod}\, M)}} a(2k) + \sum_{\substack{k \leqslant x \\ k \equiv i + M/2 (\mathrm{mod}\, M)}} a(2k) \qquad (M = 2^{d-1}). \end{aligned}$$

In order to apply (2) we introduce the symbol

$$\delta_k^{r,e} = \begin{cases} -1, & \text{if } r = e \text{ and } k \equiv b \pmod{2^{d-1}}, \\ 1, & \text{otherwise.} \end{cases} \tag{7}$$

With this we have

$$s(2x, 2i) = \sum_{\substack{k \leqslant x \\ k \equiv i (\mathrm{mod}\, M)}} \delta_k^{0,e} a(k) + \sum_{\substack{k \leqslant x \\ k \equiv i + M/2 (\mathrm{mod}\, M)}} \delta_k^{0,e} a(k)$$

or

$$s(2x, 2i) = \delta_i^{0,e} s(x, i) + \delta_{i+M/2}^{0,e} s(x, i + M/2). \tag{8}$$

Similarly, the $(2i+1)$-st entry of $\bar{s}(2x)$ is

$$s(2x, 2i+1) = \sum_{\substack{2k+1 \leqslant 2x \\ 2k+1 \equiv 2i+1 (\mathrm{mod} M)}} a(2k+1)$$

$$= \sum_{\substack{k \leqslant x \\ k \equiv i, i+M/2 (\mathrm{mod} M)}} a(2k+1) - \sum_{\substack{2x < 2k+1 \leqslant 2x+1 \\ k \equiv i, i+M/2 (\mathrm{mod} M)}} a(2k+1),$$

so that

$$s(2x, 2i+1) = \delta_i^{1,e} s(x, i) + \delta_{i+M/2}^{1,e} s(x, i+\tfrac{1}{2}M) + j(x, 2i+1), \tag{9}$$

where

$$j(x, 2i+1) = \begin{cases} -a([2x+1]), & \text{if } [2x+1] \equiv 2i+1 (\mathrm{mod} M), \\ 0, & \text{otherwise,} \end{cases} \tag{10}$$

Note that $j(x, 2i+1) = 0, \pm 1$, and is only nonzero when $[2x+1]$ is odd, i.e. when the binary digit of x in the "halves" place is 0. We set $j(x, 2i) = 0$.

If P is the zero string the only change occurs in (8), and then only when $i \equiv 0$ or $M/2$ $(\mathrm{mod}\, M)$. In this case we have

$$s(2x, 0) = -s(x, 0) + s(x, \tfrac{1}{2}M) + 2, \quad p = 0. \tag{11}$$

Hence we put $j(x, 0) = 2$, for $p = 0$, and $j(x, 2i) = 0$, otherwise.

Together, (8)–(11) give

$$s(2x, 2i+r) = \delta_i^{r,e} s(x, i) + \delta_{i+M/2}^{r,e} s(x, i+\tfrac{1}{2}M) + j(x, 2i+r) \tag{12}$$

from which Theorem 1 follows.

From (12) we can read off the entries of the matrix A. The typical row consists of 1's in two columns whose indices differ by $M/2$, and zeros elsewhere. The only exception to this occurs when $\delta_i^{r,e} = -1$ or $\delta_{i+M/2}^{r,e} = -1$, i.e. when the row index is congruent to $2b+e = p$ and the column index is b. (Row and column indices are to be read $(\mathrm{mod}\, 2^{d-1})$). Moreover, the nonzero entries occur in positions (i, j) in which $2j$ or $2j+1 \equiv i$ $(\mathrm{mod}\, M)$. Leaving zero entries blank, we have (e.g. in case p is even)

$$A = \begin{array}{c} \\ 0 \\ 1 \\ \\ \\ p \\ \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{cccccccc} 0 & 1 & & M/2 & & & & b \end{array} \\ \begin{bmatrix} 1 & & & 1 & & & & \\ 1 & & & 1 & & & & \\ & 1 & & & & 1 & & \\ & 1 & & & & 1 & & \\ & & 1 & & & & -1 & \\ & & 1 & & & & 1 & \\ & & & 1 & & & & 1 \\ & & & 1 & & & & 1 \end{bmatrix} \end{array} \tag{13}$$

It is convenient to write

$$A = B - 2E, \tag{14}$$

where

$$B = \begin{array}{c} \begin{array}{cccc} 0 & 1 & M/2 & \end{array} \\ \begin{bmatrix} 1 & & 1 & \\ 1 & & 1 & \\ & 1 & & 1 \\ & 1 & & 1 \\ & & & \\ & 1 & & 1 \\ & 1 & & 1 \end{bmatrix} \end{array}$$

and E has the single nonzero entry $+1$ in the (pb)-th position.

As for the term $\bar{j}(x)$, it has at most one nonzero entry, unless $p = 0$, in which case it can have 1 or 2 nonzero entries. This term is important for the limit theory which we discuss in Section 8.

If we iterate (6), we obtain the following result.

COROLLARY.

$$\bar{s}(2^k x) = A^k \bar{s}(x) + \sum_{r=0}^{k-1} A^{k-1-r} \bar{j}(2^r x), \tag{15}$$

where $k \geqslant 1$, A is given by (13) and the components of $\bar{j}(x)$ are

$$j(x, i) = \begin{cases} -a([2x+1]), & \text{if } i \text{ is odd and } [2x+1] \equiv i (\bmod 2^{d-1}), \\ 2, & \text{if } p = 0 \text{ and } i = 0, \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

EXAMPLES. For the patterns $P = 00, 01, 10, 11$ of length 2 we have respectively

$$A = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

In the original Rudin–Shapiro case, $P = 11$, equation (6) is the matrix from of the recursions given in Satz 2 of [5].

As examples with $d = 3$, we take the patterns $P = 010$ and $P = 111$, which correspond to the respective matrices

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

From (6) we shall develop a linear recurrence for the partial sums $s_p(x) = (1, 1, \ldots, 1)\, \bar{s}(x)$ (see (3)–(5)), but first we compute the characteristic polynomial of the matrix A.

3. The characteristic polynomial of A

We take the equation $A = B-2E$ as the starting point for our computations, and first consider the matrix B. (See [18], [23] for a discussion of a class of related matrices.)

LEMMA 1. *The minimal polynomial of the matrix B in* (14) *is* x^d-2x^{d-1}

PROOF. Consider the action of B on the standard basis vector $\bar{e}_i$, which has a 1 in the i-th row, $0 \leqslant i \leqslant 2^{d-1}-1$. From (14) it is clear that

$$B\bar{e}_i = \bar{e}_{2i}+\bar{e}_{2i+1}.$$

Iterating this gives

$$B^2\bar{e}_i = \bar{e}_{4i}+\bar{e}_{4i+1}+\bar{e}_{4i+2}+\bar{e}_{4i+3}$$

and in general

$$B^r\bar{e}_i = \bar{e}_{2^ri}+\ldots+\bar{e}_{2^ri+2^r-1}, \qquad r \leqslant d-1, \tag{17}$$

reading subscripts modulo 2^{d-1}. For $r = d-1$ we have

$$B^{d-1}\bar{e}_i = \bar{e}_0+\ldots+\bar{e}_{M-1} = \bar{e}, \tag{18}$$

and

$$B^d\bar{e}_i = \sum_{j=0}^{M-1} B\bar{e}_j = \sum_{j=0}^{M-1} (\bar{e}_{2j}+\bar{e}_{2j+1}) = 2\bar{e},$$

which implies $B^d-2B^{d-1} = 0$. From these computation it is also clear that $B^{d-1} \neq 0$ and $B^{d-1}-2B^{d-2} \neq 0$, and the lemma follows. ∎

For future reference note from (18) that

$$B^{d-1} \text{ is the all ones matrix.} \tag{19}$$

LEMMA 2. *If $\varphi(x) = x^d-2x^{d-1}$ and I is the $2^{d-1}\times 2^{d-1}$ identity matrix, then*

$$(xI-B)^{-1} = \Psi(x)/\varphi(x), \tag{20}$$

where $\Psi(x)$ is the matrix

$$\Psi(x) = Ix^{d-1}+(B-2I)x^{d-2}+(B^2-2B)x^{d-3}+ \quad +B^{d-1}-2B^{d-2} \tag{21}$$

Proof. Using Lemma 1, it is straight-forward to verify that

$$(xI-B)\Big(Ix^{d-1}+\sum_{r=1}^{d-1}(B^r-2B^{r-1})x^{d-1-r}\Big) = \varphi(x)I,$$

from which the lemma follows. ∎

Now consider the perturbed matrix $A = B-2E$, and write

$$\begin{aligned}(xI-A) &= (xI-B)(xI-B)^{-1}(xI-A)\\ &= (xI-B)(xI-B)^{-1}(xI-B+2E)\\ &= (xI-B)\big(I+2(xI-B)^{-1}E\big)\\ &= (xI-B)\frac{1}{\varphi(x)}\big(\varphi(x)I+2\Psi(x)E\big),\end{aligned}$$

using (20). Since E has a single nonzero entry in the (pb)-th position, the only nonzero column of $2\Psi(x)E$ is the b-th, and the diagonal entry in this column in $2\Psi(x)_{bp}$. It is easy to see that

$$\det\big(\varphi(x)I+2\Psi(x)E\big) = \varphi(x)^{M-1}\big(\varphi(x)+2\Psi(x)_{bp}\big),$$

and taking determinants above yields

$$\begin{aligned}\det(xI-A) &= \frac{\det(xI-B)}{\varphi(x)}\big(\varphi(x)+2\Psi(x)_{bp}\big)\\ &= \frac{\Phi(x)}{\varphi(x)}\big(\varphi(x)+2\Psi(x)_{bp}\big),\end{aligned} \tag{22}$$

where $\Phi(x) = \det(xI-B)$ is the characteristic polynomial of B.

Since the largest root of $\varphi(x) = x^{d-1}(x-2)$ is 2, the Perron–Frobenius theorem ([11]) shows that $\Phi(x)$ has 2 at a simple root, i.e. that $\Phi(x)/\varphi(x)$ is a power of x. Let $P(x)$ be the d-th degree polynomial

$$P(x) = \varphi(x)+2\Psi(x)_{bp}. \tag{23}$$

Then we have

THEOREM 2. *The characteristic polynomial of A is $x^{M-d}P(x)$, where*

$$P(x) = (x-2)(x^{d-1}+2\pi_1 x^{d-2}+ \quad +2\pi_k x^{d-1-k}+\ldots+2\pi_{d-1})+2, \tag{24}$$

$$\pi_k = (B^{k-1})_{bp}, \quad 1 \leqslant k \leqslant d-1. \tag{25}$$

Proof. It only remains to derive (24) and (25) from (23). From (23) and (21) we have

$$\begin{aligned}P(x) = x^d+(2I_{bp}-2)x^{d-1}+2(B_{bp}-2I_{bp})x^{d-2}+\\ +2[(B^{d-2})_{bp}-2(B^{d-3})_{bp}]x+2(B^{d-1})_{bp}-4(B^{d-2})_{bp}.\end{aligned} \tag{26}$$

But $(B^{d-1})_{bp} = 1$ by (19), and (24) follows easily by comparing that expression with the right side of (26). ∎

Formula (24) shows that $P(x)$ is an Eisenstein polynomial with respect to the prime 2 and is therefore irreducible over Q. Also, the computations in Lemma 1 make it clear that $\pi_k = 0$ or 1. In the following theorem we give

a precise interpretation of π_k. In order to state it easily let the pattern P be written as

$$P = p_{d-1}p_{d-2}\cdots p_1 p_0, \qquad p_i = 0 \text{ or } 1,$$

so that $p = \sum_{i=0}^{d-1} 2^i p_i$. (Notice that we are writing the digits of P from right to left.)

THEOREM 3. *For* $1 \leqslant k \leqslant d-1$, $\pi_k = 1$ *iff the pattern* P *has period* k, *i.e. iff*

$$p_{i+k} = p_i \quad \textit{for } 0 \leqslant i \leqslant d-1-k. \tag{27}$$

Proof. From (25), $\pi_k = 1$ iff the (bp)-th entry of B^{k-1} is 1. From (17) the p-th column of B^{k-1} is

$$B^{k-1}\bar{e}_p = \bar{e}_{2^{k-1}p} + \quad + \bar{e}_{2^{k-1}p+2^{k-1}-1}.$$

It follows that $\pi_k = 1$ iff

$$b \equiv 2^{k-1}p + r (\bmod 2^{d-1}),$$

for some r, $0 \leqslant r \leqslant 2^{k-1}-1$, which holds iff

$$b = 2^{k-1}p + r + l2^{d-1},$$

for some $l \in \boldsymbol{Z}$. This is equivalent in turn to the equation

$$p = 2b + e = 2r + e + 2^k p + 2^d l,$$

or

$$p = \frac{2r+e+2^d l}{1-2^k} = \frac{r'+2^d l}{1-2^k}, \qquad \text{with } 0 \leqslant r' \leqslant 2^k-1,\ r' \equiv e \ (\bmod 2).$$

Therefore $\pi_k = 1$ iff

$$(1-2^k)p \equiv r' \ (\bmod 2^d).$$

Expanding $1-2^k$ 2-adically, we write this as

$$p \equiv \frac{r'}{1-2^k} \equiv r'(1+2^k+2^{2k}+\ldots)$$

or

$$p \equiv r' + 2^k r' + 2^{2k} r' + \ldots (\bmod 2^d).$$

But this congruence is equivalent to the condition (27), i.e. that k is a period of P This proves the theorem. ∎

By (24) and Theorem 3 the polynomial $P(x)$ is determined by the symmetry properties of the pattern P. In particular, all patterns with the same period set have the same characteristic polynomial. We call $P(x)$ the *period polynomial* of the pattern P, and denote the set of periods of P by Π.

We list the period polynomials, patterns and period sets for $d = 2, 3, 4$, in Table 1.

Table 1
Period Polynomials

P	Π	$P(x)$
00, 11	$\{1\}$	x^2-2
01, 10	$\emptyset$	x^2-2x+2
000, 111	$\{1, 2\}$	x^3-2x-2
010, 101	$\{2\}$	x^3-2x^2+2x-2
001,011 100,110	$\emptyset$	x^3-2x^2+2
0000,1111	$\{1, 2, 3\}$	x^4-2x^2-2x-2
0101,1010	$\{2\}$	$x^4-2x^3+2x^2-4x-2$
0100,0010 0110,1001 1101,1011	$\{3\}$	x^4-2x^3+2x-2
0001,1000 1110,0111 0011,1100	$\emptyset$	x^4-2x^3+2

The most commonly occurring polynomials are $x^d-2x^{d-1}+2$, corresponding to the period set $\Pi = \emptyset$, and $x^d-2x^{d-1}+2x-2$, corresponding to $\Pi = \{d-1\}$. Note that $d-1$ is a period of P iff the first and last digits of P coincide. Also from (24) it is easy to see that all coefficients of $P(x)$ are $0, \pm 2$ or -4.

For $d = 1$, we let the period polynomial for both patterns $P = 0$ and $P = 1$ be $P(x) = x$, corresponding to the period set $\Pi = \emptyset$. This convention agrees with (24). Note that $s(x)$ is bounded for both of these patterns. (See [19] for a discussion of this case. Also, see [14] for similar results in connection with Theorems 2 and 3.)

4. The autocorrelation tree

The following relations between period polynomials are particularly striking. Let Π_d and Π_{d+1} be the period sets corresponding to patterns P_d and P_{d+1} of length d and $d+1$, respectively, and let $P_d(x)$ and $P_{d+1}(x)$ be the corresponding polynomials. Then it follows easily from (24) that

$$\begin{aligned} P_{d+1}(x) &= xP_d(x)-2x+2 \quad \text{if } \Pi_{d+1} = \Pi_d, \quad \text{(B)}, \\ P_{d+1}(x) &= xP_d(x)-2 \quad \text{if } \Pi_{d+1} = \Pi_d \cup \{d\}, \quad \text{(R)}. \end{aligned} \tag{28}$$

We refer to these formulas as the "Blue Rule" and the "Red Rule", respectively. Consideration of the above examples shows that for $d \leqslant 4$, each

period set arises from a period set for patterns of one less digit by either the Blue Rule or the Red Rule, i.e. each Π_{d+1} coincides with some Π_d, or differs from some Π_d by the addition of the single period d. Thus for $d \leqslant 3$, each $P_{d+1}(x)$ arises from a unique $P_d(x)$ by one of the formulae in (28).

This phenomenon continues to hold in general. In order to prove this remarkable result we make the following definition. See [13].

DEFINITION. Let P be binary digit pattern. Define the *autocorrelation* (or correlation) of P, denoted $\alpha = \alpha(P)$, by putting

$$\alpha = \pi_{d-1}\pi_{d-2}\ldots\pi_1\pi_0,$$

where $\pi_0 = 1$ and π_k is defined by (25) or by the property: $\pi_k = 1$ iff k is a period of P, $1 \leqslant k \leqslant d-1$. (Again we write from right to left; this is the opposite of the convention adopted in [13].)

In the same way, if Π is a period set of patterns of length d, $\alpha(\Pi)$ is defined as the autocorrelation of any pattern with period set Π, i.e.

$$\alpha = \alpha_{d-1}\alpha_{d-2}\ldots\alpha_1\alpha_0 \quad \text{with } \alpha_0 = 1 \text{ and } \alpha_k = 1 \text{ iff } k \in \Pi.$$

The string α is itself a digit pattern of length d, and encodes the set of periods of a pattern P. Furthermore, these autocorrelations are clearly in one-to-one correspondence with all period polynomials, by (24) and Theorem 3.

If (28) holds for the period sets Π_{d+1} and Π_d then their autocorrelations satisfy the following relations:

$$\alpha(\Pi_d) = \alpha_{d-1}\alpha_{d-2}\ldots\alpha_1\alpha_0, \tag{29}$$

$$\alpha(\Pi_{d+1}) = \alpha_d\alpha_{d-1}\ldots\alpha_1\alpha_0, \tag{29'}$$

where

$$\alpha_d = \begin{cases} 0 & \text{for rule (B)}, \\ 1 & \text{for rule (R)}. \end{cases} \tag{29''}$$

Thus $\alpha(\Pi_{d+1})$ arises from $\alpha(\Pi_d)$ by the addition of a single digit.

We now prove the following theorem.

THEOREM 4. *If $\alpha' = \alpha(\Pi_{d+1})$ is the autocorrelation of a period set Π_{d+1}, then α' arises from some $\alpha(\Pi_d)$ by the rules* (29)–(29″).

Proof. We use the characterization of correlations given in [13]. A binary pattern $v = v_{n-1}v_{n-2}\ldots v_1v_0$ of length n is the autocorrelation of some binary pattern of length n iff $v_0 = 1$ and v satisfies the following two rules:

1. *Forward propagation rule.* If $v_k = v_l$, with $k < l$, then $v_m = 1$ for all m of the form $m = k+i(l-k)$, $i \geqslant 1$, which lie in the range $k \leqslant m < n$.

2. *Backward propagation rule.* Let k, l be any indices for which $k < l \leqslant 2k$ and $v_k = v_l = 1$, but $v_{2k-l} = 0$. (Note $2k-l = k-(l-k)$.) If $s = [(n-k)/(l-k)]$,

then $v_m = 0$ for all m of the form $m = k-i(l-k)$, $1 \leqslant i \leqslant s$, which lie in the range $0 \leqslant m \leqslant 2k-l$.

By this result α', being an autocorrelation, must satisfy both propagation rules 1 and 2. But then the pattern v which results by dropping the leading digit α_d has length d, and also satisfies rules 1 and 2. Hence $v = \alpha(\Pi_d)$ must be the autocorrelation of some pattern of length d, and $\alpha(\Pi_d)$ and $\alpha(\Pi_{d+1})$ are related by (29)–(29″). ■

COROLLARY. *Every period set Π_{d+1} arises from a period set Π_d by one of the rules*

$$\Pi_{d+1} = \Pi_d \quad \text{or} \quad \Pi_{d+1} = \Pi_d \cup \{d\}.$$

Every period polynomial $P_{d+1}(x)$ arises from some $P_d(x)$ by either the Blue Rule or the Red Rule (28).

In other words, the autocorrelations form a connected tree (see Figs. 1 and 2 and Figs. 7–9 in Section 11). In this tree the nodes are the autocorrelations, and the edges are colored blue or red corresponding to the rule which leads from one correlation to the next. Note that each $\alpha(\Pi_{d+1})$ has exactly one "predecessor" $\alpha(\Pi_d)$, but $\alpha(\Pi_d)$ can give rise to 0, 1, or 2 autocorrelations at the next level.

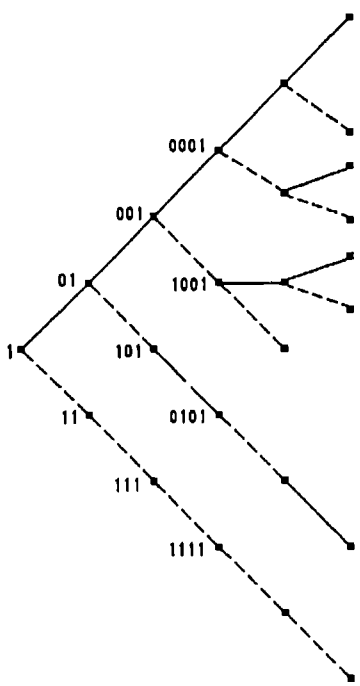


Fig. 1. The autocorrelation tree ($d \leqslant 6$). Key: blue —, red ---.

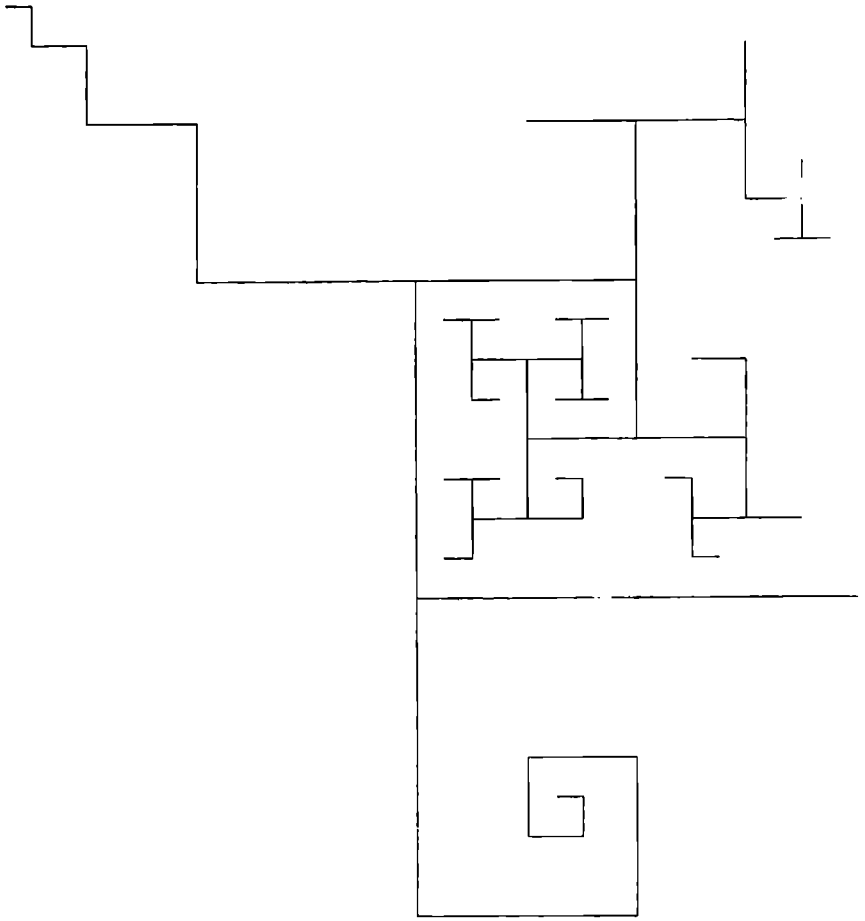


Fig. 2. The autocorrelation tree ($d \leqslant 9$). (Nodes are represented by segments and edges by rotations: $-90°$ for blue edges, $+90°$ for red edges.)

We shall discuss the structure of this tree in more detail in Section 11. For now we make the following remarks. The forward and backward propagation rules (see Theorem 4) make it clear that the patterns

$$\varrho_d = \overbrace{1\,1\ldots 1}^{d} \quad \text{and} \quad \beta_d = \overbrace{0\ldots 0\,1}^{d}$$

are correlations and correspond to the period sets

$$\Pi_\varrho = \{1, 2, \ldots d-1\} \quad \text{and} \quad \Pi_\beta = \emptyset.$$

The ϱ's are all connected by red branches, and the β's, by blue branches. The forward propagation rule shows further that ϱ_{d+1} is the only "descendant" of ϱ_d. In fact, if a correlation contains two consecutive 1's, it can only generate red branches. From this it is easy to see that every correlation in the tree is either a ϱ_d or is descended from some β_d. For convenience we shall refer to the subtree $\{\varrho_1, \varrho_2, \ldots\}$ as the red branch and the subtree $\{\varrho_1, \beta_2, \beta_3, \ldots\}$ as the blue branch.

5. The roots of the period polynomials

In order to investigate bounds for the sum function $s_p(x)$, it is necessary to give an analysis of the roots of $P(x)$ for a given pattern P. We begin with the period polynomials on the red and blue branches.

The polynomial corresponding to $\varrho = \varrho_d$ is

$$P_\varrho(x) = (x-2)(x^d+2x^{d-1}+\quad+2)+2 = x^d-2\sum_{k=0}^{d-2} x^k = \frac{x^{d+1}-x^d-2x^{d-1}+2}{x-1}. \tag{30}$$

By Descartes' rule of signs, $P_\varrho(x)$ has a unique positive real root ξ_d. Concerning this root we prove

LEMMA 3 (Red branch). *If ξ_d is the positive real root of $P_\varrho(x)$, where $\varrho = \varrho_d = 1\ldots1$ (d digits) then the sequence $\{\xi_d\}$ is strictly increasing, $\xi_2 = \sqrt{2}$, $\sqrt{2} < \xi_d < 2$ for $d \geqslant 3$, and $\lim_{d\to\infty} \xi_d = 2$.*

For $d \geqslant 3$ all roots of $P_\varrho(x)$ distinct from ξ_d have modulus $< \xi_d$.

Proof. Writing $P_d(x)$ for $P_\varrho(x)$ and using the Red Rule gives

$$P_{d+1}(\xi_d) = -2 \quad \text{for } d \geqslant 2.$$

But $P_{d+1}(2) = 2$ (see (24)), and it follows immediately that $\xi_d < \xi_{d+1} < 2$. This proves the first three assertions, since $P_2(x) = x^2-2$.

To prove $\lim_{d\to\infty} \xi_d = 2$, put $x = \xi_d$ in (30); this gives

$$\lim_{d\to\infty} (\xi_d^2-\xi_d-2) = \lim_{d\to\infty} \frac{-2}{\xi_d^{d-1}} = 0,$$

and since $(\xi_d^2-\xi_d-2) = (\xi_d-2)(\xi_d+1)$ we find $\lim_{d\to\infty} \xi_d = 2$.

To prove the final assertion consider (30):

$$|P_d(x)| \geqslant |x|^d-2\Big|\sum_{k=0}^{d-2} x^k\Big| \geqslant |x|^d-2\Big|\sum_{k=0}^{d-2} x^k\Big| = P_d|(x)|. \tag{31}$$

This inequality implies $P_d(x) \neq 0$ if $|x| > \xi_d$. In addition, if $|x| = \xi_d$ and $P_d(x) = 0$, then we must have equality at each step in (31), so

$$\Big|\sum_{k=0}^{d-2} x^k\Big| = \sum_{k=0}^{d-2} |x|^k.$$

But for $d \geqslant 3$ this can happen only if all x^k lie on the same ray through the origin, i.e. x must be positive real, and $x = \xi_d$. ∎

On the blue branch the polynomial corresponding to $\beta = \beta_d$ is

$$P_\beta(x) = x^{d-1}(x-2)+2 = x^d-2x^{d-1}+2. \tag{32}$$

Clearly $P_\beta(2) = 2$ and $P_\beta(x) > 0$ for $x > 2$. Descartes' rule of signs shows that $P_\beta(x)$ has either 0 or 2 roots in $(0, 2)$ and one or no negative real roots according as d is odd or even. In the former case this root lies in $(-1, 0)$.

To determine whether $P_\beta(x)$ has 0 or 2 real roots in $(0, 2)$ we evaluate $P_\beta(x)$ at $x = 2-2/d$, which is the unique minimum of $P_\beta(x)$ in $(0, \infty)$. P_β has 2 real roots iff its minimum value is negative. On the other hand

$$P_\beta\left(2-\frac{2}{d}\right) = \left(2-\frac{2}{d}\right)^{d-1}\left(\frac{-2}{d}\right)+2$$

is negative iff

$$2-\frac{2}{d} > d^{1/(d-1)}. \tag{33}$$

Since the right-hand side is a decreasing function of d (for $d \geqslant 3$) and the left-hand side is increasing, we need only find the smallest value of d for which it holds. This value is $d = 5$, and so (33) is true for all $d \geqslant 5$.

LEMMA 4. (Blue branch). *For $d \geqslant 5$, the polynomial $P_\beta(x)$ has a positive root ξ_d in the internal $2-2/d < \xi_d < 2$. All other roots of $P_\beta(x)$ lie in the disc $|x| \leqslant d^{1/(d-1)} < \xi_d$, and $\lim_{d\to\infty} \xi_d = 2$.* For $d = 2$ *or* 4, $P_\beta(x)$ *has no real roots; for $d = 3$, P_β has a unique real root lying in the interval* $(-1, 0)$.

Proof. The only assertion left to prove is that all roots of P_β other than ξ_d lie in the disc $|x| = d^{1/(d-1)}$, for $d \geqslant 5$. But this is immediate from a theorem of van Vleck ([25], p. 153), according to which a polynomial of the form (32) has at least $d-1$ of its roots in the disc $|x| \leqslant d^{1/(d-1)}$. ∎

For future reference we list the exceptional polynomials of Lemmas 3 and 4 in Table 2. These are the period polynomials which do not have a unique dominant real root.

Table 2

Exceptional Polynomials

d	Π	$\alpha(\Pi)$	$P(x)$
2	$\{1\}$	11	x^2-2
2	$\emptyset$	01	x^2-2x+2
3	$\emptyset$	001	x^3-2x^2+2
4	$\emptyset$	0001	x^4-2x^3+2

We will now show that all other period polynomials do have a positive dominant real root.

First consider polynomials which are the direct descendants of polynomials on the blue branch. These are polynomials corresponding to the autocorrelations $\delta_d = 10\ldots01$ of length d, i.e.

$$(34)\qquad P_\delta(x) = (x-2)(x^{d-1}+2)+2 = x^d-2x^{d-1}+2x-2, \quad d \geqslant 3,$$
$$\Pi_\delta = \{d-1\}.$$

Since

$$P_\delta(2) = 2 \text{ and } P_\delta\left(\frac{3}{2}\right) = \left(\frac{3}{2}\right)^{d-1}\left(\frac{-1}{2}\right)+1 \leqslant \frac{-1}{8},$$

P_δ has a root ξ_d between $\frac{3}{2}$ and 2. The next lemma shows that all subsequent polynomials inherit this property.

LEMMA 5. *If $P_d(x)$ (on any branch) has a root $\xi > 1$, then any immediate descendant $P_{d+1}(x)$ has a root in the interval $(\xi, 2)$.*

Proof. If $P_{d+1}(x)$ is a blue descendant of $P_d(x)$, then by (28), $P_{d+1}(\xi) = -2\xi+2 < 0$. If $P_{d+1}(x)$ is a red descendant, then $P_{d+1}(\xi) = -2$. In either case the assertion follows from $P_{d+1}(2) = 2$ (see (24)). ∎

COROLLARY. *If $P_d(x)$ is any polynomial not on the red or blue branches, $P_d(x)$ has a real root between $\frac{3}{2}$ and 2.*

Proof. This is because any such polynomial is a descendant of one of the polynomials (34). (Recall the penultimate remark in Section 4.) ∎

Note. For $d \geqslant 5$, (34) shows that

$$P_\delta\left(\frac{9}{5}\right) = \left(\frac{9}{5}\right)^{d-1}\left(\frac{-1}{5}\right)+\frac{8}{5} < 0.$$

It follows that all polynomials, except possibly those descended from $P_3(x) = x^3-2x^2+2x-2$ ($\delta_3 = 101$) or from $P_4(x) = x^4-2x^3+2x-2$ ($\delta_4 = 1001$), have real roots in $(\frac{9}{5}, 2)$. In addition, all degree 5 descendants of these two polynomials have roots in $(\frac{9}{5}, 2)$ as well, as can easily be checked by hand. These descendants are the period polynomials of the correlations 10101, 11001, and 01001. Consequently all period polynomials of degree 5 which are not on the red or blue branches have roots in $(\frac{9}{5}, 2)$.

We require one more lemma in order to prove that the real root of $P_d(x)$ in $(\frac{9}{5}, 2)$ dominates the other roots.

LEMMA 6. *Let $\Pi \neq \emptyset$ be the period set of a pattern P of length d, and let $m = \min \Pi$. If $l \in \Pi$, then either $m|l$ or $l > d-m$.*

Proof. This is trivially true if $m = 1$. Assume $m \geqslant 2$, let l be an element of Π which is not a multiple of m, and set

$$l \equiv l' (\mathrm{mod}\, m), \quad \text{where } 0 \leqslant l' \leqslant m-1.$$

Since m is a period of P, we can write P as

$$P = p_{r-1} \cdots p_0 p_{m-1} \cdots p_0 \cdots p_{m-1} \cdots p_0$$

where $r < m$; in other words, we can read subscripts in P modulo m. Consider any subscript i, $0 \leqslant i \leqslant d-1$, and put $i \equiv i' \pmod{m}$, where $i' < m$.

Assuming that $l+m \leqslant d$, we have $l+i' < d$, for all $i' < m$, and because l is a period of P,

$$p_i = p_{i'} = p_{i'+l} = p_{i'+l'} = p_{i+l'}$$

for all i which $i+l' < d$. But this implies l' is a period of P, i.e. $l' \in \Pi$. This contradicts the definition of m, so that $l+m > d$, or $l > d-m$. ∎

This lemma is related to the GCD rule in Guibas and Odlyzko [13].

Now consider (24) in the form

$$P_d(x) = (x-2)Q_d(x)+2, \qquad Q_d(x) = x^{d-1}+2\pi_1 x^{d-2}+ \quad +2\pi_{d-1}, \tag{35}$$

and let the reciprocal polynomials of P and Q be

$$\hat{P}_d(x) = x^d P_d\left(\frac{1}{x}\right), \qquad \hat{Q}_d(x) = x^{d-1} Q_d\left(\frac{1}{x}\right).$$

From (35) we see that

$$\begin{aligned} \hat{P}_d(x) &= (1-2x)\hat{Q}_d(x)+2x^d, \\ \hat{Q}_d(x) &= 1+2\pi_1 x+2\pi_2 x^2+ \quad +2\pi_{d-1}x^{d-1}. \end{aligned} \tag{36}$$

Since we are considering polynomials which is do not lie on the red or blue branches, we take $m \geqslant 2$ in Lemma 6, which implies that

$$\hat{Q}_d(x) = 1+2x^m+2x^{2m}+ \quad + \sum_{l>d-m} 2x^l, \tag{37}$$

where the last sum is over certain exponents $l > d-m$. If we let $d \to \infty$ in such a way that the period set Π_d has least period m, we find

$$\hat{Q}_d(x) \to 1+\sum_{i=1}^{\infty} 2x^{mi} = \frac{1+x^m}{1-x^m} \quad \text{for } |x| < 1, \text{ as } d \to \infty,$$

$$\lim_{d\to\infty} \hat{P}_d(x) = (1-2x)\frac{1+x^m}{1-x^m} \quad \text{for } |x| < 1,\ m = \min \Pi_d. \tag{38}$$

(By the propagation rules in Section 4, the patterns β_m^j, consisting of the pattern β_m repeated j times, are correlations having least period m. Thus the limit in (38) makes sense.)

By a theorem of Hurwitz ([34], p. 119), $\frac{1}{2}$ is the only limit point of the zeros of the $\hat{P}_d(x)$ in $|x| < 1$, which shows that 2 is the only limit point of the zeros of the $P_d(x)$ in $|x| > 1$. Since $\frac{1}{2}$ is a simple zero in (38), this also shows that for large d, all but one of the roots of $P_d(x)$ cluster onto the unit disc $|x| \leqslant 1$. The

roots guaranteed by Lemma 5 (and its corollary) must therefore tend to 2, and must dominate all other roots of $P_d(x)$ for large enough d. Note that a similar phenomenon occurs also on the red and blue branches, since (30) and (32) show that the corresponding polynomials $\hat{P}_d(x)$ on these branches tend respectively to

$$\frac{1-x-2x^2}{1-x} = (1-2x)\frac{1+x}{1-x} \quad \text{and} \quad 1-2x.$$

In order to prove a more precise result we estimate $\hat{Q}_d(x)$ in the disc $|x| \leqslant r$ for a fixed r in $(\frac{1}{2}, 1)$. If

$$\hat{Q}_d(x) = 1+2\sum_{i=1}^{k} x^{im} + \sum_{l>d-m} 2x^l,$$

then $k = [(d-1)/m]$, and

$$\hat{Q}_d(x) = \frac{2x^{m(k+1)}-x^m-1}{x^m-1} + 2\sum_{l>d-m} x^l; \tag{39}$$

for $|x| \leqslant r$ it follows that

$$|\hat{Q}_d(x)| \geqslant \frac{1-|x|^m-2|x|^{m(k+1)}}{|x|^m+1} - 2\sum_{i=l}^{\infty} |x|^i \geqslant \frac{1-r^m-2r^{m(k+1)}}{1+r^m} - \frac{2r^l}{1-r},$$

where l is the least element of Π_d which is not a multiple of m. We use the facts $(k+1)m \geqslant d$ and $l > \max(d-m, m) \geqslant \frac{1}{2}(d-m+m) = \frac{1}{2}d$, to deduce

$$|\hat{Q}_d(x)| \geqslant \frac{1-r^m-2r^d}{1+r^m} - \frac{2r^{d/2}}{1-r}, \quad \text{for } |x| \leqslant r. \tag{40}$$

Now apply Rouche's Theorem ([34]) to (36): if

$$|(1-2x)\hat{Q}_d(x)| > |2x^d| \quad \text{on } |x| = r, \tag{41}$$

then $\hat{P}_d(x)$ has the same number of zeros as $(1-2x)\hat{Q}_d(x)$ in $|x| \leqslant r$. If the inequality

$$(2r-1)\left(\frac{1-r^m-2r^d}{1+r^m} - \frac{2r^{d/2}}{1-r}\right) > 2r^d \tag{42}$$

holds, then by (40), $\hat{Q}_d(x) \neq 0$ in $|x| \leqslant r$, and (41) holds, so that $\hat{P}_d(x)$ has *one* zero in $|x| \leqslant r$; but then $P_d(x)$ has one zero in $|x| > 1/r$, and this zero dominates the roots in $|x| < 1/r$.

To finish the argument we only need to pick r and find the values of d for which (42) holds. First, it is clear that the left-hand side of (42) is an increasing function of d, for $\frac{1}{2} < r < 1$, and the right-hand side is a decreasing function of d. If the inequality holds for some d, it will hold for all larger d.

We handle the cases $m = 2$ and $m \geqslant 3$ separately. For $m \geqslant 3$ it is enough to show that

$$(2r-1)\left(\frac{1-r^3-2r^d}{1+r^3}-\frac{2r^{d/2}}{1-r}\right) > 2r^d \quad (m \geqslant 3).$$

With $r = 5/9$ we find that the inequality holds for all $d \geqslant 8$.

When $m = 2$, the last term in (39) is absent because all periods in Π are even. (By Lemma 6, the only odd period which could occur is $l = d-1$, in case d is even. The autocorrelation of Π would then have the form $110101\ldots01$. But the Backward Propagation Rule does not hold for this pattern (take $k = d-2$, $l = d-1$), and so $d-1 \notin \Pi$.) We can replace (42) in this case by the inequality

$$(2r-1)\frac{1-r^2-2r^d}{1+r^2} > 2r^d \quad (m = 2),$$

and find that this holds with $r = \frac{5}{8}$ for $d \geqslant 7$.

This analysis implies that all $P_d(x)$ with $d \geqslant 8$ have a unique positive real root $> \frac{9}{5}$ which is also the root of largest modulus. It only remains to check the period polynomials not on the red or blue branches with $d \leqslant 7$. There are 21 such polynomials, and it can be checked that for each of these polynomials, the positive real root guaranteed by Lemma 5 is the unique root of largest modulus. This completes the proof of

THEOREM 5. *Except for the exceptional polynomials listed in Table* 2, *each polynomial $P_d(x)$ has a positive dominant real root which is less than* 2 *and which tends to* 2 *as $d \to \infty$. As $d \to \infty$ the other roots of $P_d(x)$ cluster onto the unit disc $|x| \leqslant 1$.*

We finish the section with a proof of

LEMMA 7. *If P is any pattern, then its period polynomial $P(x)$ has no roots on the unit circle.*

Proof. Let $\gamma = e^{i\theta}$ be a root of $P(x) = 0$. Then γ is an algebraic integer, and a root of $x^d P(1/x) = 0$, the reciprocal equation. The polynomial $P(x)$ is Eisenstein with respect to the prime 2, and therefore

$$0 = \gamma^d P\left(\frac{1}{\gamma}\right) = 1+2\gamma u(\gamma),$$

where $u(x) \in \boldsymbol{Z}[x]$. But $u(\gamma)$ is also an algebraic integer, and $-1 = 2\gamma u(\gamma)$, an equation which is clearly impossible since it implies 2 divides -1. This proves the lemma. ∎

6. A recurrence relation for $s_p(x)$

In the next section we shall derive an explicit formula for $s(x)$, which will easily give a nontrivial upper bound. This formula will follow from a recurrence relation satisfied by $s(x)$, which depends in turn on the curious fact that the matrix $P(A)$ has column sums equal to zero.

In order to derive this relation let E be the operator defined by

$$(Ef)(x) = f(2x) \quad \text{for } f\colon \boldsymbol{R}^+ \to \boldsymbol{C}.$$

With this operator we write (15) in the form

$$E^k \bar{s}(x) = A^k \bar{s}(x) + \bar{b}_k(x), \tag{43}$$

where

$$\bar{b}_k(x) = \sum_{r=0}^{k-1} A^{k-1-r} \bar{j}(2^r x), \quad k \geqslant 1. \tag{44}$$

Note that $\bar{b}_k(x)$ is a bounded function of x with rational integral entries.

Let $P(x)$ be the period polynomial (24) associated to the matrix A and the pattern P. If $P(x) = \sum_{k=0}^{d} c_k x^k$, then by (43) and (44)

$$\begin{aligned} P(E)\bar{s}(x) &= P(A)\bar{s}(x) + \sum_{k=1}^{d} c_k \bar{b}_k(x) \\ &= P(A)\bar{s}(x) + \sum_{r=0}^{d-1} \Big(\sum_{k=r+1}^{d} c_k A^{k-1-r} \Big) \bar{j}(2^r x). \end{aligned}$$

Multiplying on the left by $\bar{e}^t = (1, 1, \ldots, 1)$ gives

$$P(E)s(x) = \bar{e}^t P(A)\bar{s}(x) + \sum_{r=0}^{d-1} \bar{u}_r \bar{j}(2^r x),$$

with the row vectors

$$\bar{u}_r = \bar{e}^t \sum_{k=r+1}^{d} c_k A^{k-1-r} \tag{45}$$

We shall prove shortly that $\bar{e}^t P(A) = 0$. Assuming this and writing

$$b(x) = \sum_{r=0}^{d-1} \bar{u}_r \bar{j}(2^r x), \tag{46}$$

we first have

THEOREM 6. *The partial sums $s(x) = s_p(x)$ satisfy the recurrence relation*

$$P(E)s(x) = b(x), \quad x \geqslant 0, \tag{47}$$

where $b(x)$ is given by (45) *and* (46), *and* $P(x) = \sum_{k=0}^{d} c_k x^k$ *is the period polynomial corresponding to the pattern P The function $b(x)$ takes on bounded integral values.*

EXAMPLE. If P is the pattern 111, $P(x) = x^3 - 2x - 2$, and $\bar{j}(x)$ is a 4-dimensional column vector whose entries are

$$j(x, i) = \begin{cases} -a([2x+1]), & \text{if } [2x+1] \equiv i (\bmod 4),\ i \text{ odd}, \\ 0 & \text{otherwise}. \end{cases}$$

Thus (47) becomes

$$s(8x) - 2s(2x) - 2s(x) = b(x). \tag{48}$$

with

$$b(x) = \bar{u}_0 \bar{j}(x) + \bar{u}_1 \bar{j}(2x) + \bar{u}_2 \bar{j}(4x).$$

Computing the vectors $\bar{u}_r$ from (45) we find

$$\bar{u}_0 = \bar{e}^t (A^2 - 2I) = (2, 0, 2, 0),$$
$$\bar{u}_1 = \bar{e}^t A = (2, 2, 2, 0),$$
$$\bar{u}_2 = \bar{e}^t I = (1, 1, 1, 1).$$

Further, since $\bar{j}(x)$ has nonzero entries only in odd positions, $(2, 0, 2, 0)\bar{j}(x) = 0$, and (48) simplifies to:

$$s(8x) - 2s(2x) - 2s(x) = 2j(2x, 1) + j(4x, 1) + j(4x, 3).$$

Note that $|b(x)| \leqslant 3$ since at most one of $j(4x, 1)$ and $j(4x, 3)$ is nonzero.

Using the defining recurrence for $a_p(n) = a(n)$, namely,

$$a(2n+1) = \begin{cases} -a(n), & n \equiv 3 \ (\bmod 4), \\ a(n), & \text{otherwise}, \end{cases}$$
$$a(2n) = a(n),$$

it is not difficult to check that

$$b(n) = 2j(2n, 1) + j(4n, 1) + j(4n, 3) = -3a(n)$$

if n is an integer $\geqslant 0$. Therefore, putting $x = n$ in the above formula gives

$$s(8n) - 2s(2n) - 2s(n) = -3a(n), \qquad n \geqslant 0.$$

Recurrences for the other residue classes (mod 8) follow in a similar manner: e.g., if $x = n + \frac{5}{8}$, then

$$s(8n+5) - 2s(2n+1) - 2s(n) = b(n + \tfrac{5}{8}) = 0.$$

To show $\bar{e}^t P(A) = 0$ we need a lemma.

LEMMA 8. *If P has length $d \geqslant 2$, then*

$$a(i)A^r\bar{e}_i = \sum_{k=2^ri}^{2^ri+2^r-1} a(k)\bar{e}_k, \quad r \geqslant 0,$$

where $\bar{e}_i$ $(0 \leqslant i \leqslant 2^{d-1}-1)$ is the i-th standard basis vector, and subscripts are read modulo $M = 2^{d-1}$. Here $i \geqslant 1$ if P is the zero pattern, and $i \geqslant 0$ otherwise.

Proof. (By induction.) This is obvious for $r = 0$. For $r = 1$ the claim is that

$$a(i)A\bar{e}_i = a(2i)\bar{e}_{2i} + a(2i+1)\bar{e}_{2i+1}. \tag{49}$$

However, the definition of A implies

$$A\bar{e}_i = \delta_i^{0,e}\bar{e}_{2i} + \delta_i^{1,e}\bar{e}_{2i+1},$$

and (49) follows on multiplying through by $a(i)$ and using the recursive formula (2).

Assume the assertion holds for r. Then by (49)

$$a(i)A^{r+1}\bar{e}_i = \sum_{k=2^ri}^{2^ri+2^r-1} a(k)A\bar{e}_k = \sum_{k=2^ri}^{2^ri+2^r-1} \big(a(2k)\bar{e}_{2k} + a(2k+1)\bar{e}_{2k+1}\big)$$

$$= \sum_{k=2^{r+1}}^{2^{r+1}i+2^{r+1}-1} a(k)\bar{e}_k. \quad \blacksquare$$

Now consider the row vector

$$\bar{e}_t P(A) = \bar{e}^t \sum_{r=0}^{d} c_r A^r = \sum_{r=0}^{d} c_r \bar{e}^t A^r$$

This vector is zero iff its components $\bar{e}^t P(A)\bar{e}_i = 0$, for all $0 \leqslant i \leqslant M-1$. Thus $\bar{e}^t P(A) = 0$ is equivalent to

$$\sum_{r=0}^{d} c_r \bar{e}^t A^r \bar{e}_i = 0 \quad \text{for } 0 \leqslant i \leqslant M-1.$$

Substituting for $A^r\bar{e}_i$ from Lemma 8 gives

$$\sum_{r=0}^{d} c_r \sum_{k=2^ri}^{2^ri+2^r-1} a(k) = 0, \tag{50}$$

on multiplying through by $a(i)$ and noting $\bar{e}^t\bar{e}_k = 1$. We set

$$N_r = \sum_{j=0}^{2^r-1} a(2^ri+j). \tag{51}$$

With this notation, together with the values of the coefficients c_r from (24), equation (50) becomes

$$\sum_{r=0}^{d} c_r N_r = 0, \tag{52}$$

or

$$(2-4\pi_{d-1})N_0+(2\pi_{d-1}-4\pi_{d-2})N_1+ \quad +(2\pi_2-4\pi_1)N_{d-2} \\ +(2\pi_1-2)N_{d-1}+N_d = 0.$$

Rearranging, we have to prove that

$$2N_0+\sum_{k=1}^{d-1} 2\pi_{d-k}(N_k-2N_{k-1})+(N_d-2N_{d-1}) = 0, \tag{53}$$

for the appropriate values of i.

Note that $N_0 = a(i)$, and

$$N_r-2N_{r-1} = \sum_{j=0}^{2^r-1} a(2^r i+j)-2\sum_{j'=0}^{2^{r-1}-1} a(2^{r-1}i+j').$$

Setting

$$\delta_a = \begin{cases} 1, & a \equiv b \pmod{2^{d-1}} \ (b \text{ from } (i)) \\ 0, & \text{otherwise,} \end{cases}$$

and using (2) applied to $a(2^r i+j) = a(2(2^{r-1}i+j')+\varepsilon)$, $\varepsilon = 0, 1$, we find

$$N_r-2N_{r-1} = -2\sum_{j'=0}^{2^{r-1}-1} a(2^{r-1}i+j')\delta_{2^{r-1}i+j'} = -2Q_{r-1},$$

for $1 \leqslant r \leqslant d$. This is because two consecutive terms in N_r give $2a(2^{r-1}i+j')$ if $2^{r-1}i+j' \not\equiv b \pmod{2^{d-1}}$ and 0 if $2^{r-1}i+j' \equiv b \pmod{2^{d-1}}$. However, the last congruence determines j' uniquely $\bmod 2^{r-1}$, since $r \leqslant d$. Therefore

$$Q_r = \begin{cases} a(2^r i+j) & \text{if } 2^r i+j \equiv b \pmod{2^{d-1}}, \text{ for some } j \in [0, 2^r-1], \\ 0, & \text{otherwise,} \end{cases}$$

for $0 \leqslant r \leqslant d-1$, and (53) is equivalent to

$$a(i) = 2\pi_{d-1}Q_0+2\pi_{d-2}Q_1+ \quad +2\pi_1 Q_{d-2}+Q_{d-1}, \tag{54}$$

where $Q_{d-1} = a(2^{d-1}i+b)$.

We shall now prove (54) for all patterns P and all $i \geqslant 0$ ($i \geqslant 1$ if $P = 0\ldots0$). We focus on the term Q_{d-1}. By considering the binary representation of $2^{d-1}i+b$ it is clear that

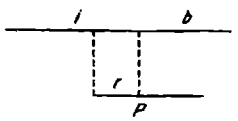
$$a(2^{d-1}i+b) = (-1)^n a(i)a(b) = (-1)^n a(i),$$

where n is the number of overlap occurrences of P, those which use digits from

i and b (note $a(b) = 1$). Let the overlap length of such an occurrence be r, where $1 \leqslant r \leqslant d-1$. Then it is easy to see that $i \equiv b_r$ (the first r digits of b) $(\mathrm{mod}\, 2^r)$ and r is a period of P, i.e. $\pi_r = 1$. (See the figure below.)

On the other hand, $Q_{d-1-r} \neq 0$ iff there is a solution $j(\mathrm{mod}\, 2^{d-1-r})$ to the congruence

$$2^{d-1-r}i+j \equiv b(\mathrm{mod}\, 2^{d-1}), \tag{55}$$



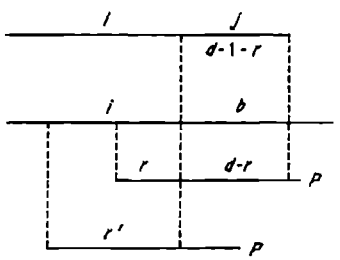
which is equivalent to

$$b \equiv j(\mathrm{mod}\, 2^{d-1-r}) \quad \text{and} \quad i \equiv \frac{b-j}{2^{d-1-r}}(\mathrm{mod}\, 2^r).$$

Clearly such a j exists iff $i \equiv b_r(\mathrm{mod}\, 2^r)$. Hence overlap occurrences of P in $2^{d-1}i+b$ correspond one-to-one to nonzero terms $\pi_r Q_{d-1-r}$ in (54).

In particular, if P has no periods, all $\pi_l = 0$ and no overlap occurrences are possible, so $Q_{d-1} = a(i)$ and (54) holds.

Now order the overlap occurrences in order of increasing r and consider a term $\pi_r Q_{d-1-r}$ in this ordering. By similar reasoning as above, $Q_{d-1-r} = \pm a(i)$ according as the number of overlap occurrences of P in $2^{d-1-r}i+j$ is even or odd, where j satisfies (55). From the diagram every overlap occurrence in $2^{d-1}i+b$ with $r' > r$ is an overlap occurrence in $2^{d-1-r}i+j$, and the *original* overlap occurrence (corresponding to $\pi_r Q_{d-1-r}$) is not an overlap occurrence in $2^{d-1-r}i+j$. Hence each term has one fewer overlap occurrences than the preceding term, implying that the nonzero terms in (54) alternate in sign.



If n is even, $a(2^{d-1}i+b) = a(i)$ and the n nonzero terms $2\pi_r Q_{d-1-r}$ in (54) cancel. Thus (54) is a true equation.

If n is odd, $a(2^{d-1}i+b) = -a(i)$; all terms but the first term $2\pi_r Q_{d-1-r}$ (with smallest r) cancel, giving

$$2\pi_r Q_{d-1-r} + Q_{d-1} = 2a(i) - a(i) = a(i),$$

and (54) holds in every case.

The result we have proved seems fundamental enough to state as

THEOREM 7. *For any pattern P, $\bar{e}^t P(A) = 0$, where $P(x)$ is the period polynomial corresponding to P. This is equivalent to the formula* (52), *for* $0 \leqslant i \leqslant 2^{d-1}-1$ (*or* $1 \leqslant i \leqslant 2^{d-1}$, *if* $P = 0\ldots0$), *where N_r is defined by* (51) *and the c_r's are the coefficients of $P(x)$. In fact, the equivalent formulas* (52) *and* (54) *hold for all $i \geqslant 0$ ($i \geqslant 1$ if $P = 0\ldots0$).*

7. An explicit formula

Returning to equation (47) we set

$$x = 2^y \quad \text{and} \quad \hat{s}(y) = s(x).$$

Under this substitution we have $E\hat{s}(y) = \hat{s}(y+1)$, and $P(E)\hat{s}(y) = 0$ is a standard linear difference equation with constant coefficients. All solutions of this equation are of the form

$$\hat{s}(y) = \sum_{k=1}^{d} \hat{\lambda}_k(y)\xi_k^y, \qquad y \in \boldsymbol{R},$$

where $\hat{\lambda}_k$: $\boldsymbol{R} \to \boldsymbol{C}$, $\hat{\lambda}_k(y+1) = \hat{\lambda}_k(y)$, the ξ_k are the roots of $P(x) = 0$, and ξ_i^y is any fixed branch of the exponential function. (See [29], pp. 356, 386.) Thus for some complex-valued function $\lambda_k(x) = \hat{\lambda}_k(\log x/\log 2)$, defined for $x > 0$, we have

$$s(x) = \sum_{k=1}^{d} \lambda_k(x)\xi_k^{\log x/\log 2} + s_0(x) = \sum_{k=1}^{d} \lambda_k(x)x^{\log \xi_k/\log 2} + s_0(x) \tag{56}$$

where $s_0(x)$ is a particular solution of (47). Note that $\lambda_k(2x) = \lambda_k(x)$ for $x > 0$ and $k = 1, \ldots, d$.

To find a suitable s_0, we divide by $P(E)$ in (47) and write

$$s_0(x) = \frac{1}{P(E)}b(x) = \sum_{k=1}^{d} \frac{a_k}{E - \xi_k I}b(x), \qquad a_k = \frac{1}{P'(\xi_k)} \tag{57}$$

using partial fractions. To make sense of the terms in this sum we consider the roots with $|\xi_k| > 1$ and $|\xi_k| < 1$ separately. (Remember that $|\xi_k| \neq 1$ by Lemma 7.)

If $|\xi_k| > 1$, then we set

$$\frac{1}{E-\xi_k I}b(x) = \frac{-1}{\xi_k}\sum_{r=0}^{\infty}\left(\frac{E}{\xi_k}\right)^r b(x) = -\sum_{r=0}^{\infty}\frac{b(2^r x)}{(\xi_k)^{r+1}}, \qquad x \geqslant 0,$$

where the sum converges absolutely and uniformly.

If $|\xi_k| < 1$, we take

$$\frac{1}{E-\xi_k I}b(x) = E^{-1}\frac{1}{I-\xi_k E^{-1}}b(x) = E^{-1}\sum_{r=0}^{\infty}\xi_k^r E^{-r}b(x)$$
$$= \sum_{r=0}^{\infty}\xi_k^r b\left(\frac{x}{2^{r+1}}\right),$$

and the sum is again absolutely and uniformly convergent. It now follows easily from (57) that

$$s_0(x) = -\sum_{\substack{k\\|\xi_k|>1}}\frac{1}{P'(\xi_k)}\sum_{r=0}^{\infty}\frac{b(2^r x)}{\xi_k^{r+1}} + \sum_{\substack{k\\|\xi_k|<1}}\frac{1}{P'(\xi_k)}\sum_{r=0}^{\infty}\xi_k^r b\left(\frac{x}{2^{r+1}}\right) \tag{58}$$

is a particular solution to (47). (See [29], p. 399.) We remark that $s_0(x)$ is bounded for $x \geqslant 0$, because $b(x)$ is bounded. Further, since $\bar{j}(x)$ is continuous from the right for all $x \geqslant 0$, this property is shared by $b(x)$ and $s_0(x)$.

Now set

$$t(x) = s(x) - s_0(x), \tag{59}$$

and replace x by $x, 2x, \ldots, 2^{d-1}x$ in (56). Writing τ_k for $\log\xi_k/\log 2$, we have the following system of linear equations in the $\lambda_k(x)$:

$$t(x) = \sum_{k=1}^{d}\lambda_k(x)x^{\tau_k},$$

$$t(2x) = \sum_{k=1}^{d}\lambda_k(x)\xi_k x^{\tau_k},$$

$$t(2^{d-1}x) = \sum_{k=1}^{d}\lambda_k(x)\xi_k^{d-1}x^{\tau_k}.$$

Since the determinant of the system is

$$\begin{vmatrix} x^{\tau_1} & x^{\tau_d} \\ (\xi_1)x^{\tau_1} & (\xi_d)x^{\tau_d} \\ \vdots & \\ (\xi_1^{d-1})x^{\tau_1} & (\xi_d^{d-1})x^{\tau_d} \end{vmatrix} = x^{\tau_1}\ldots x^{\tau_d}\Delta,$$

where $\Delta = \pm\sqrt{D} \neq 0$ and D is the discriminant of $P(x)$, it follows by Cramer's rule and elementary operations that for $x > 0$,

$$(60) \qquad \lambda_k(x) = \frac{1}{\Delta x^{\tau_k}} \begin{vmatrix} 1 & t(x) & 1 \\ \xi_1 & t(2x) & \xi_d \\ & \cdot & \\ \xi_1^{d-1} & t(2^{d-1}x) & \xi_d^{d-1} \end{vmatrix} \qquad \tau_k = \frac{\log \xi_k}{\log 2}.$$

It is not difficult to see directly from this formula that $\lambda_k(2x) = \lambda_k(x)$. Moreover the elements of the determinant are bounded, for $1 \leqslant x \leqslant 2$, so the functions $\lambda_k(x)$ are also bounded in this interval. By the rule $\lambda_k(2x) = \lambda_k(x)$ it follows that $\lambda_k(x)$ is bounded for all $x > 0$. Clearly $\lambda_k(x)$ is also right-continuous for all $x > 0$ by the remarks made above about $s_0(x)$. By equation (56), the boundedness of the $\lambda_k(x)$ immediately implies

THEOREM 8. *For $x \geqslant 0$, $s_p(x) = O(x^\tau)$, where $\tau = \log\xi/\log 2$ and ξ is the maximum modulus of the roots of the period polynomial $P(x)$ corresponding to the pattern P.*

By the results of Section 5, $\tau < 1$ and $\tau \to 1$ as $d \to \infty$. (Also see Table 2 and Section 10 for the roots of the four exceptional polynomials.)

COROLLARY. *Let $N_e(x)$ and $N_0(x)$ represent the number of integers $\leqslant x$ in whose binary expansions the pattern P occurs respectively an even, and an odd, number of times.*

Then $N_e(x) \simeq N_0(x)$ as $x \to \infty$ and

$$N_e(x) = \tfrac{1}{2}x + O(x^\tau), \qquad \text{as } x \to \infty.$$

Proof. This is immediate from Theorem 8, since

$$N_e(x) + N_0(x) = x + O(1), \qquad N_e(x) - N_0(x) = s(x) = O(x^\tau). \qquad \blacksquare$$

We finish this section by completing the proof of

THEOREM 9. *The partial sum function $s_p(x) = s(x)$ can be expressed as the sum of the continuous function*

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

(with $\lambda_k(x)$ defined by (59) and (60)) and the bounded function

$$s_0(x) = \sum_{r=0}^{\infty} b\left(\frac{x}{2^{r+1}}\right) \sum_{\substack{k \\ |\xi_k| < 1}} \frac{\xi_k^r}{P'(\xi_k)} - \sum_{r=0}^{\infty} b(2^r x) \sum_{\substack{k \\ |\xi_k| > 1}} \frac{1}{P(\xi_k)\xi_k^{r+1}}.$$

The coefficients $\lambda_k(x)$ are continuous functions of x, for $x > 0$, and satisfy $\lambda_k(2x) = \lambda_k(x)$. Here the ξ_k are all the conjugate roots of the period polynomial $P(x)$ corresponding to P.

Proof. It only remains to prove that $\lambda_k(x) = 0$ if $|\xi_k| < 1$, and that $\lambda_k(x)$ is continuous in x. Set

$$(61)\qquad s_k(x) = \begin{cases} \lambda_k(x)x^{\tau_k} - \sum_{r=0}^{\infty} \dfrac{b(2^r x)}{P'(\xi_k)\xi_k^{r+1}}, & \text{if } |\xi_k| > 1, \\ \lambda_k(x)x^{\tau_k} + \sum_{r=0}^{\infty} \dfrac{\xi_k^r}{P'(\xi_k)} b\left(\dfrac{x}{2^{r+1}}\right), & \text{if } |\xi_k| < 1, \end{cases}$$

so that

$$s(x) = \sum_{k=1}^{d} s_k(x)$$

by (56) and (58). Applying $E - \xi_k I$ to $s_k(x)$ gives in either case that

$$(E - \xi_k I)s_k(x) = \frac{1}{P'(\xi_k)} b(x) = a_k b(x),$$

and iterating this equation gives

$$(62)\qquad s_k(2^m x) = \xi_k^m s_k(x) + a_k \sum_{r=0}^{m-1} \xi_k^{m-1-r} b(2^r x), \qquad x > 0,\ m \geqslant 1.$$

Note the resemblance of this formula to (15).

We shall use (62) to express $\lambda_k(x)$ as a uniform limit involving $s_k(2^m x)$. If $|\xi_k| > 1$ in (62), we have

$$(63)\qquad \lim_{m\to\infty} \frac{s_k(2^m x)}{\xi_k^m} = s_k(x) + a_k \sum_{r=0}^{\infty} \frac{b(2^r x)}{\xi_k^{r+1}} = \lambda_k(x)x^{\tau_k}.$$

If $|\xi_k| < 1$, put $x/2^m$ for x in (62) and solve for $\xi_k^m s_k(x/2^m)$. This gives

$$\lim_{m\to\infty} \xi_k^m s_k\left(\frac{x}{2^m}\right) = s_k(x) - a_k \sum_{r=0}^{\infty} \xi_k^r b\left(\frac{x}{2^{r+1}}\right) = \lambda_k(x)x^{\tau_k}.$$

But by (60) and (61) the limit $\lim_{h\to 0+} s_k(h) = s_k(0)$ exists and is finite. From the last equation it follows that

$$(64)\qquad \lambda_k(x) = x^{-\tau_k} \lim_{m\to\infty} \xi_k^m s_k\left(\frac{x}{2^m}\right) = 0, \qquad \text{if } |\xi_k| < 1.$$

This proves the first assertion.

We now claim that the limit in (63) is uniform. For example, (61) implies that

$$\left|\frac{s_k(x)}{x^{\tau_k}} - \lambda_k(x)\right| \leqslant \frac{c}{x^{\tau_k}}, \qquad x > 0, \text{ if } |\xi_k| > 1.$$

Putting $2^m x$ for x gives

$$\left|\frac{s_k(2^m x)}{\xi_k^m x^{\tau_k}} - \lambda_k(x)\right| \leqslant \frac{c}{\xi_k^m x^{\tau_k}}, \quad \text{for } |\xi_k| > 1,$$

and this proves our claim on any interval $[a, b]$ with $a > 0$.

Now from (46), (58)–(61), it follows that $s_k(x)$ is right-continuous, with bounded jumps. Therefore, $s_k(2^m x)/\xi_k^m$ has jumps which tend to 0. A standard argument can now be used to show that $\lambda_k(x)$ is continuous. (See the theorem in §3 of [6].) ∎

Remark. By this result the roots of $P(x)$ inside the unit circle only contribute to the bounded term $s_0(x)$. This contribution, which by (61) and (64) is

$$\sum_{|\xi_k|<1} s_k(x) = \sum_{r=1}^{\infty} b\left(\frac{x}{2^{r+1}}\right) \sum_{|\xi_k|<1} \frac{\xi_k^r}{P'(\xi_k)},$$

is a rational function of the roots ξ_k, since $b(x/2^{r+1}) = b(0)$ for sufficiently large r (depending on x).

8. The limit function

We assume in this section that P is a pattern for which the corresponding period polynomial $P(x)$ has a single dominant real root. By the results of Section 5 this is the case for all but four nodes in the autocorrelation tree, and thus for all but four classes of binary patterns.

We denote the dominant real root of $P(x)$ by $\xi_1 = \xi$, and the corresponding coefficient in (56) by $\lambda_1(x) = \lambda(x)$. Writing $\tau = \log\xi/\log 2$, we have

$$\lim_{n\to\infty} \frac{s(2^n x)}{(2^n x)^\tau} = \lim_{n\to\infty} \frac{s(2^n x)}{\xi^n x^\tau} = \lambda(x). \tag{65}$$

It follows that the curve $(x, \lambda(x))$, for $1 \leqslant x \leqslant 2$, represents the limiting behavior of the function $s(x)/x^\tau$ on the intervals $2^n \leqslant x \leqslant 2^{n+1}$, as $n \to \infty$. We have already seen that $\lambda(x)$ is continuous. We now derive an explicit formula for $\lambda(x)$ and use this to show that $\lambda(x)$ is non-differentiable almost everywhere. In particular, this will show that τ is the exponent of the correct order of magnitude for $s(x)$. We also prove a result which generalizes a theorem of [6] concerning the case $P = 11$.

We first prove

LEMMA 9. *The limit* $\lim_{n\to\infty}(\xi^{-1}A)^n$ *exists and equals the matrix* $L = \bar{v}\bar{u}$, *where* $\bar{u}$ *and* $\bar{v}$ *and left and right eigenvectors of* A *corresponding to the eigenvalue* ξ, *and* $\bar{u}\cdot\bar{v} = 1$.

Proof. (See also [11], pp. 89–92, applied to the regular, but non-stochastic, matrix $\xi^{-1}A$.) Let K be the normal closure of $Q(\xi)$, and let T be a matrix with coefficients in K for which $T^{-1}AT = \tilde{A}$ is in Jordan normal form:

$$T^{-1}AT = \tilde{A} = \begin{bmatrix} \xi_1 & & & & \\ & \xi_2 & & & \\ & & & & \\ & & & \xi_d & \\ & & & & N \end{bmatrix},$$

N being composed of blocks of the form

$$\begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & 1 & & \\ & & & & \\ & & & 1 & 0 \end{bmatrix}$$

In the corresponding basis of K^M we have

$$A\bar{v}_i = \xi_i\bar{v}_i, \quad 1 \leqslant i \leqslant d,$$
$$A\bar{v}_i = \bar{v}_{i+1} \text{ or } 0, \; i > d.$$

It follows that

$$A^n\bar{v}_i = \begin{cases} \xi_i^n\bar{v}_i, & i \leqslant d, \\ 0, & i > d, \end{cases}$$

for $n \geqslant M-d$, and hence $\lim_{n\to\infty}(\xi^{-1}A)^n = L$, with $L\bar{v}_1 = \bar{v}_1$, $L\bar{v}_i = 0$ for $i \geqslant 2$. In the standard basis, therefore

$$L = T\begin{bmatrix} 1 & & & \\ & 0 & & \\ & & & \\ & & & 0 \end{bmatrix}T^{-1}$$

But from elementary linear algebra $T = [\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_M]$. If $\bar{u} = (u_1, \ldots, u_M)$ is the first row of T^{-1}, then

$$L = [\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_M]\begin{bmatrix} 1 & & & \\ & 0 & & \\ & & & \\ & & & 0 \end{bmatrix}\begin{bmatrix} \bar{u} \\ * \end{bmatrix} = [\bar{v}_1, 0, \ldots, 0]\begin{bmatrix} \bar{u} \\ * \end{bmatrix} = \bar{v}\bar{u},$$

with $\bar{v} = \bar{v}_1$. The vector $\bar{v}$ is an eigenvector of A corresponding to the eigenvalue ξ. Since $\tilde{A}^t = (T^t)A^t(T^t)^{-1}$, the first column of $(T^{-1})^t$, or $\bar{u}^t$, is an eigenvector of A^t corresponding to ξ, and therefore $\bar{u}$ is a left eigenvector of A. Finally, the fact $\bar{u}\cdot\bar{v} = 1$ follows from $T^{-1}T = I$. ∎

We remark that the property $\bar{u}\cdot\bar{v} = 1$ determines the matrix product $L = \bar{v}\bar{u}$ completely. In particular, L has entries in $Q(\xi)$.

COROLLARY.

$$\lim_{n\to\infty} \frac{\bar{e}^t A^n}{\xi^n} = (\bar{e}\cdot\bar{v})\bar{u}.$$

LEMMA 10. *$\bar{e}\cdot\bar{v} = 2v_b(2-\xi)^{-1} \neq 0$, where v_b denotes the b-th component of $\bar{v}$.*

Proof. Let $\bar{v} = (v_0, \ldots, v_{M-1})^t$, $M = 2^{d-1}$, and read subscripts (mod M). From $A\bar{v} = \xi\bar{v}$ and (13) we have

$$\xi v_i = \begin{cases} v_{[i/2]} + v_{[i/2]+M/2}, & i \neq p, \\ -v_b + v_{b+M/2}, & i = p. \end{cases} \tag{66}$$

Consequently $\bar{e}\cdot\bar{v} = v_0 + \ldots + v_{M-1}$ and

$$\xi\bar{e}\cdot\bar{v} = 2v_0 + \ldots + 2v_{m-1} - 2v_b = 2\bar{e}\cdot\bar{v} - 2\bar{v}_b,$$

so that

$$\bar{e}\cdot\bar{v} = \frac{2}{2-\xi} v_b.$$

If this expression is 0, $v_b = 0$, and then (66) holds with v_b in place of $-v_b$. But (see (14)) this implies $B\bar{v} = \xi\bar{v}$, a contradiction, since B only has the eigenvalues 0 and 2. ∎

THEOREM 10.

$$\lim_{n\to\infty} \frac{\bar{s}(2^n x)}{\xi^n} = L\bar{s}(x) + L\sum_{r=0}^{\infty} \frac{\bar{j}(2^r x)}{\xi^{r+1}}, \quad \textit{for } x \geqslant 0. \tag{67}$$

Proof. The proof is modelled after the proof of the Toeplitz theorem on summability ([17], p. 35). Let $\|N\|$ denote the largest absolute value of an entry of the matrix N. From (15) we have

$$\begin{aligned} \xi^{-n}\bar{s}(2^n x) &= (\xi^{-1}A)^n\bar{s}(x) + \sum_{r=0}^{n-1} (\xi^{-1}A)^{n-1-r} \frac{\bar{j}(2^r x)}{\xi^{r+1}} \\ &= (\xi^{-1}A)^n\bar{s}(x) + \sum_{r=0}^{n-1} (\xi^{-1}A)^r \frac{\bar{j}(2^{n-1-r}x)}{\xi^{n-r}}. \end{aligned} \tag{68}$$

Given $\varepsilon > 0$, choose n_0 so that

$$\|(\xi^{-1}A)^n - L\| < \varepsilon \quad \text{for } n \geqslant n_0.$$

Then for $n \geqslant n_0$

$$\sum_{r=0}^{n-1} (\xi^{-1}A)^n \frac{\bar{j}(2^{n-1-r}x)}{\xi^{n-r}}$$

$$= \sum_{r=0}^{n_0-1} \left(\left(\frac{A}{\xi}\right)^r - L\right) \frac{\bar{j}(2^{n-1-r}x)}{\xi^{n-r}} + \sum_{r=n_0}^{n-1} \left(\left(\frac{A}{\xi}\right)^r - L\right) \frac{\bar{j}(2^{n-1-r}x)}{\xi^{n-r}} + L \sum_{r=0}^{n-1} \frac{\bar{j}(2^r x)}{\xi^{r+1}}$$

$$= \Sigma_1 + \Sigma_2 + \Sigma_3.$$

The components of Σ_1 tend to 0 as $n \to \infty$, while the components of Σ_2 are bounded by

$$\sum_{r=n_0}^{n-1} \varepsilon M \frac{2}{\xi^{n-r}} < 2M\varepsilon \sum_{r=1}^{\infty} \frac{1}{\xi^r} = c\varepsilon.$$

Since Σ_3 is the partial sum of a convergent series we have

$$\lim_{n\to\infty} \sum_{r=0}^{n-1} (\xi^{-1}A)^n \frac{\bar{j}(2^{n-1-r}x)}{\xi^{n-r}} = L \sum_{r=0}^{\infty} \frac{\bar{j}(2^r x)}{\xi^{r+1}},$$

and (67) follows from (68). ∎

We deduce

COROLLARY 1. *For* $x > 0$,

$$\lambda(x) = \lim_{n\to\infty} \frac{s(2^n x)}{(2^n x)^\tau} = \frac{\bar{e}\cdot\bar{v}}{x^\tau} \bar{u}\bar{s}(x) + \frac{\bar{e}\cdot\bar{v}}{x^\tau} \sum_{r=0}^{\infty} \frac{\bar{u}\bar{j}(2^r x)}{\xi^{r+1}}. \tag{69}$$

Proof. This follows from (67) by multiplying on the left by $\bar{e}^t$ and using Lemma 9.

Remark 1. The proof of Theorem 10 shows that the limit in (69) is uniform in x on any finite interval $[a, b]$ with $a > 0$.

Remark 2. If $\bar{w}$ is any M-dimensional column-vector, it follows easily from (67) and (69) that

$$\lim_{n\to\infty} \frac{\bar{w}\cdot\bar{s}(2^n x)}{(2^n x)^\tau} = \left(\frac{\bar{w}\cdot\bar{v}}{\bar{e}\cdot\bar{v}}\right)\lambda(x). \tag{70}$$

By Remark 1, every linear combination of the subsums $s_p(x, i)$ (see (4)) is $O(x^\tau)$. In particular, if the i-th component of $\bar{v}$ is nonzero, then $s_p(x, i)$ has the same limiting behavior as $s(x)$ (take $\bar{w} = \bar{e}_i$). This is certainly true of the sum $s_p(x, b)$, by Lemma 10. From these remarks we may deduce the following corollary, the proof of which is analogous to the proof of the corollary to Theorem 8.

COROLLARY 2. *Let* $N_e(x, i)$ *and* $N_0(x, i)$ *represent the number of integers* $\leqslant x$ *and* $\equiv i \pmod{2^{d-1}}$ *in whose binary expansions the pattern* P *occurs*

respectively an even, and an odd, number of times. Then $N_e(x, i) \simeq N_0(x, i)$ *as* $x \to \infty$, *and*

$$N_e(x, i) = \frac{x}{2^d} + O(x^\tau), \qquad \tau = \frac{\log \xi}{\log 2}.$$

We also note the relation

$$N_e(x, i) - N_0(x, i) = s_p(x, i),$$

which shows that the sign of the limit in (70) (with $\bar{w} = \bar{e}_i$) determines the relative sizes of $N_e(x, i)$ and $N_0(x, i)$ as $x \to \infty$.

Remark 3. Formula (67) and Lemma 9 imply that the limit

$$\bar{\lambda}(x) = \lim_{n \to \infty} \frac{\bar{s}(2^n x)}{(2^n x)^\tau}$$

is a multiple of $\bar{v}$, for every $x > 0$, and so

$$A\bar{\lambda}(x) = \xi \bar{\lambda}(x), \qquad x > 0.$$

Returning to formula (69), we set

$$\mu_p(x) = \mu(x) = -\sum_{r=0}^{\infty} \frac{\bar{u}\bar{j}(2^r x)}{\xi^{r+1}}, \tag{71}$$

in terms of which we have

$$\lambda(x) = \frac{\bar{e} \cdot \bar{v}}{x^\tau}(\bar{u}\bar{s}(x) - \mu(x)). \tag{71'}$$

Let $x = \sum_{r=0}^{\infty} d_r 2^{-r}$ be the binary expansion of x, where $d_0 \in \boldsymbol{Z}$ and $d_i = 0$ or 1 for $i \geqslant 1$ (infinitely many $d_i \neq 1$), and set

$$x_k = [2^k x] = \sum_{r=0}^{k} d_r 2^{k-r} \tag{72}$$

We use the x_k to rewrite the formula for $\mu(x)$, as follows. From (16) we have

$$j(2^r x, 2i+1) = \begin{cases} -a(x_{r+1}+1), & \text{if } x_{r+1}+1 \equiv 2i+1 \pmod{2^{d-1}}, \\ 0, & \text{otherwise.} \end{cases}$$

Also, $\bar{j}(x)$ has at most one nonzero coordinate, unless P is the all zero string, in which case $j(x, 0) = 2$. If we let $\bar{u} = (u_0, u_1, \ldots, u_{M-1})$ and define

$$u(n) = \begin{cases} u_{2i+1}, & \text{if } n \equiv 2i \pmod{2^{d-1}}, \\ 0, & \text{otherwise,} \end{cases} \tag{73}$$

then (71) and (73) imply the formula

$$\mu(x) = \begin{cases} \sum_{r=1}^{\infty} \frac{a(x_r+1)u(x_r)}{\xi^r}, & \text{if } p \neq 0, \\ \frac{-2u_0}{\xi-1} + \sum_{r=1}^{\infty} \frac{a(x_r+1)u(x_r)}{\xi^r}, & \text{if } p = 0. \end{cases} \tag{74}$$

To show that $\mu(x)$ is not constant, we prove

LEMMA 11. *For some* i, $u_{2i+1} \neq 0$.

Proof. The vector $\bar{u}$ satisfies $\bar{u}A = \xi\bar{u}$, so from (13) we have

$$\xi u_i = \begin{cases} u_{2i}+u_{2i+1}, & i \neq b, \\ -u_p+u_{p\pm 1}, & i = b. \end{cases} \tag{75}$$

Assume $u_{2i+1} = 0$ for all i. Then replacing i by $2i+1$ in (75) shows that $u_{4i+2} = 0$ for all i. Similarly, replacing i by $4i+2$ gives $u_{8i+4} = 0$ for all i. Continuing, we find $u_j = 0$ for all $j = 2^k i+2^{k-1}$ and $1 \leqslant k \leqslant d-1$. The only residue class $(\bmod 2^{d-1})$ left out in this process is 0. But then $\xi\bar{u}$ has one nonzero coordinate u_0 (since $\bar{u} \neq 0$), while $\bar{u}A$ has two nonzero coordinates. This contradiction proves the lemma. ∎

We now apply (74) and this lemma to prove

THEOREM 11. *For almost all real numbers* $x_0 > 0$, $\lambda(x)$ *is nondifferentiable at* x_0.

Proof. By (71′) it is enough to prove the theorem with $\mu(x)$ in place of $\lambda(x)$, since $\bar{u}\bar{s}(x)$ is a step function with jumps at the integers. We show that

$$\frac{1}{h}(\mu(x_0+h)-\mu(x_0)) \quad \text{is unbounded as } h \to 0+. \tag{76}$$

Let x_0 be normal for the base 2. Then all digit patterns occur infinitely often in the binary expansion of x_0. Since almost all real numbers are normal (base 2), it suffices to prove (76) for such an x_0. (See [30].)

Pick an i for which $u_{2i+1} \neq 0$. Let $p_0 \ldots p_n p_{n+1} \ldots p_{n+m}$ be a sequence of 0's and 1's chosen that

$$2i \equiv 2^{d-2}p_0+2^{d-3}p_1+ \quad +p_n \ (\bmod 2^{d-1}), \quad \text{so } p_n = 0 \text{ and } p_{n+j} = 1 \text{ for } 1 \leqslant j \leqslant m.$$

Also, let m be large enough that

$$\frac{2U}{\xi^m(\xi-1)} < |u_{2i+1}|, \quad \text{with } U = \max_{0 \leqslant j \leqslant M-1} |u_j|.$$

Since x_0 is normal the pattern $p_0 \ldots p_{n+m}$ occurs infinitely often as a digit string in the representation $x_0 = \sum_{r=0}^{\infty} d_r 2^{-r}$; i.e., there is an infinite sequence of

k's for which $x_k \equiv 2i \pmod{2^{d-1}}$, so that x_k ends in the digit string $p_0 p_1 \ldots p_n$, and d_k is followed by m ones:

$$d_{k+j} = 1 \quad \text{for } 1 \leqslant j \leqslant m.$$

Put $h = 2^{-k}$ and $x_0 + h = \sum_{r=0}^{\infty} d'_r 2^{-r}$, where

$$d'_r = \begin{cases} d_r, & r < k, \\ 1, & r = k, \\ d_r, & r > k. \end{cases}$$

Now if $x_r = [2^r x_0]$ and $x'_r = [2^r(x_0+h)]$, then by (72)

$$x_r \equiv x'_r \equiv 1 \pmod 2 \quad \text{for } n+1 \leqslant r \leqslant n+m.$$

From (73) and (74) it follows that

$$\begin{aligned} \mu(x_0+h) - \mu(x_0) &= \sum_{r=1}^{\infty} \frac{a(x'_r+1)u(x'_r)}{\xi^r} - \sum_{r=1}^{\infty} \frac{a(x_r+1)u(x_r)}{\xi^r} \\ &= -\frac{a(x_k+1)u(x_k)}{\xi^k} + \sum_{r=k+m+1}^{\infty} \frac{a(x'_r+1)u(x'_r) - a(x_r+1)u(x_r)}{\xi^r} \\ &= -\frac{a(x_k+1)u_{2i+1}}{\xi^k} + \varepsilon_k, \end{aligned}$$

with

$$|\varepsilon_k| \leqslant 2U \sum_{r=k+m+1}^{\infty} \frac{1}{\xi^r} = \frac{2U}{\xi^{k+m}(\xi-1)}.$$

Therefore

$$\left|\frac{\mu(x_0+h)-\mu(x_0)}{h}\right| \geqslant \left(\frac{2}{\xi}\right)^k \left(|u_{2i+1}| - \frac{2U}{\xi^m(\xi-1)}\right) \to \infty, \quad \text{as } k \to \infty,$$

since $\xi < 2$. This completes the proof of the theorem. ∎

For the pattern $P = 11$ it can be shown that the function $\lambda(x)$ considered in [6] is differentiable *nowhere*. This is probably also true of $\lambda_p(x)$ for any pattern P, but we have not checked this. The proof would require a finer analysis of the properties of $a_p(n)$.

COROLLARY. *The "correct" order of magnitude of $s(x)$ is x^τ, i.e. $s(x) = \Omega(x^\tau)$.*

Proof. From (56)

$$s(x) = \lambda(x)x^\tau + O(x^{\tau-\delta}), \quad \text{as } x \to \infty,$$

where $\delta > 0$. By the theorem, $\lambda(x_0) \neq 0$ for some x_0. If we set $y_k = 2^k x_0$, then

$$s(y_k) \simeq \lambda(x_0) y_k^\tau \quad \text{as } y_k = 2^k x_0 \to \infty. \quad ∎$$

Using (74) we can also prove the following result.

THEOREM 12. *If* $x \in Q$, *then* $x^{\tau}\lambda(x) \in Q(\xi)$.

To put this in perspective, note that $x^{\tau}\lambda(x)$ is certainly an element of $K = Q(\xi_1, \xi_2, \ldots, \xi_d)$ whenever $s_0(2^n x) \in K$ for $0 \leqslant n \leqslant d-1$, by (59) and (60). However the more precise result in Theorem 12 follows from the *analytic* formula (69) for $\lambda(x)$.

Proof. It is clear from (71′) and (74) that we need only show

$$\sum_{r=1}^{\infty} \frac{a(x_{r+1})u(x_r)}{\xi^r} \in Q(\xi) \tag{77}$$

when $x = \sum_{r=0}^{\infty} d_r 2^{-r} \in Q$. With suitable integers $q \geqslant 2$ and $r_0 \geqslant 1$ we may assume that $d_{r+q} = d_r$, for $r \geqslant r_0$. The object of the proof is to show that

$$a(x_{k+2q(d-1)}+1)u(x_{k+2q(d-1)}) = a(x_{k+1})u(x_k), \qquad k \geqslant r_0+q(d-1). \tag{78}$$

The assertion (77) will then follow.

We first consider

$$x_{k+q(d-1)} = \sum_{s=0}^{k+q(d-1)} d_s 2^{k+q(d-1)-s} = 2^{q(d-1)} \sum_{s=0}^{k} d_s 2^{k-s} + \sum_{s=k+1}^{k+q(d-1)} d_s 2^{k+q(d-1)-s}$$
$$= 2^{q(d-1)} x_k + \sum_{s=1}^{q(d-1)} d_{s+k} 2^{q(d-1)-s} = 2^{q(d-1)} x_k + x'_k.$$

Clearly $x_{k+q(d-1)} \equiv x'_k \pmod{2^{d-1}}$ and $x'_{k+q} = x'_k$, for $k \geqslant r_0$. Also $x_k \equiv d_k \pmod 2$.

Now if $d_k = 1$ and $k \geqslant r_0$, then $d_{k+2q(d-1)} = 1$, $u(x_{k+2q(d-1)}) = u(x_k) = 0$, and (78) holds trivially. We may therefore assume that $d_k = 0 = d_{k+q(d-1)}$.

We compute the expression $a(x_{k+q(d-1)}+1)$. For this consider $x_{k+q(d-1)}+1 = 2^{q(d-1)}x_k + x'_k + 1$ as being pieced together from the digits of x_k and x'_k+1. Since $a(n)$ counts occurrences of the pattern P in strings of length d, the contributions to $a(x_{k+q(d-1)}+1)$ come from $a(x_k)$, $a(x'_k+1)$, and the overlap

$$o_k = \underbrace{d_{k-d+2} \ldots d_{k-1}}_{d-1} d_k \underbrace{d_{r+1} \ldots d_{r+d-1}}_{d-1}.$$

Thus

$$a(x_{k+q(d-1)}+1) = a(x_k)a(x'_k+1)a(o_k). \tag{79}$$

(In this formula compute $a(x'_k+1)$ and $a(o_k)$ by treating x'_k+1 and o_k as strings of binary digits rather than rational integers, i.e. ignore our usual conventions about leading zeros.) Since x_k and x_k+1 differ by a 1 in the last place we have the relation

$$a(x_k) = a(x_k+1)a(o'_k)a(o''_k), \tag{80}$$

where o'_k and o''_k are the digit strings

$$o'_k = d_{k-d+1} \ldots d_{k-1} 0, \qquad o''_k = d_{k-d+1} \ldots d_{k-1} 1.$$

Putting (80) into (79) gives

$$a(x_{k+q(d-1)}+1) = a(x_k+1)\varepsilon_k, \quad \text{with } \varepsilon_k = a(o_k)a(o_k')a(o_k'')a(x_k'+1).$$

Now $o_{k+q} = o_k$, $o_{k+q}' = o_k'$ and $o_{k+q}'' = o_k''$ as long as $k \geqslant r_0+d-1$, so for these k we have $\varepsilon_{k+q} = \varepsilon_k = \pm 1$ and

(81)

$$a(x_{k+2q(d-1)}+1) = a(x_{k+q(d-1)}+1)\varepsilon_{k+q(d-1)} = a(x_k+1)\varepsilon_k\varepsilon_{k+q(d-1)} = a(x_k+1).$$

Finally, we note that

$$x_{k+2q(d-1)} \equiv x_{k+q(d-1)}' = x_{k-q(d-1)}' \equiv x_k \pmod{2^{d-1}},$$

for $k \geqslant r_0+q(d-1)$. Hence $u(x_{k+2q(d-1)}) = u(x_k)$, and formula (78) is a consequence of (81). ∎

As a consequence of Theorem 12 it follows that $\lambda(x)$ is zero or transcendental if $x \in Q$ and x^τ is transcendental. The latter is most likely the case for any rational $x \neq 2^a$.

9. The case $P = 111$

We compute $\lambda(x)$ explicitly in the case $P = 111$. The matrix A has eigenvectors

$$\bar{u} = (\xi+1, \xi^2-1, \xi+1, 1), \quad \bar{v}_1 = (1,1, \xi-1, \xi^2-\xi-1)^t,$$

where ξ is the real root of $P(x) = x^3-2x-2 = 0$:

$$\xi = \frac{1}{3}\sqrt[3]{27+3\sqrt{57}}+\frac{1}{3}\sqrt[3]{27-3\sqrt{57}} = 1.76929235\ldots$$

Since $\bar{u}\bar{v}_1 = 3\xi^2-2 = P'(\xi)$, we put $\bar{v} = \bar{v}_1/P'(\xi)$. Then $\bar{e}\cdot\bar{v} = \xi^2/(3\xi^2-2)$, and from (71′) and (74) we have

$$x^\tau\lambda(x) = \frac{\xi^2}{3\xi^2-2}\left(\bar{u}\bar{s}(x)-\sum_{r=1}^{\infty}\frac{a(x_r+1)u(x_r)}{\xi^r}\right), \tag{82}$$

where $x = \sum_{r=0}^{\infty} d_r 2^{-r}$ and $x_r = [2^r x]$. In this case (73) gives

$$u(n) = \begin{cases} \xi^2-1, & n \equiv 0 \pmod 4, \\ 1, & n \equiv 2 \pmod 4, \\ 0, & n \equiv 1, 3 \pmod 4. \end{cases}$$

If $1 \leqslant x < 2$, $\bar{s}(x) = \bar{s}(1) = (1,1,0,0)^t$ and (82) becomes

$$x^\tau\lambda(x) = \frac{\xi^2}{3\xi^2-2}\left(\xi^2+\xi-\sum_{r=1}^{\infty}\frac{a(x_r+1)u(x_r)}{\xi^r}\right), \quad 1 \leqslant x < 2. \tag{83}$$

The infinite sum in this formula is bounded by

$$\left|\sum_{r=1}^{\infty} \frac{a(x_r+1)u(x_r)}{\xi^r}\right| \leqslant \frac{1}{\xi}+\sum_{r=2}^{\infty} \frac{\xi^2-1}{\xi^r} = \frac{\xi+2}{\xi}.$$

(Note that $x_1 = 2$ or 3 since $1 \leqslant x < 2$.) This leads to the bounds

$$\frac{\xi^2(\xi+1)}{3\xi^2-2} \leqslant x^\tau\lambda(x) \leqslant \frac{\xi(\xi^2+3\xi+4)}{3\xi^2-2}, \qquad 1 \leqslant x < 2,$$

or

$$\frac{1}{19}(\xi^2+8\xi+5) \leqslant x^\tau\lambda(x) \leqslant \frac{1}{19}(15\xi^2+6\xi-1), \qquad \text{for } 1 \leqslant x < 2, \tag{84}$$

by virtue of the relations

$$\frac{1}{3\xi^2-2} = \frac{1}{38}(-6\xi^2+9\xi+8), \qquad \xi^3 = 2\xi+2. \tag{85}$$

Dividing by x^τ and using $1 \leqslant x^\tau < \xi$ gives that

$$\frac{1}{38}(5\xi^2+2\xi+6) < \lambda(x) \leqslant \frac{1}{19}(15\xi^2+6\xi-1), \qquad x > 0, \tag{86}$$

or

$$.66290952\ldots < \lambda(x) \leqslant 2.97745713\ldots$$

Since $\lambda(x) > 0$ and the components of $\bar{v}$ are positive, the remarks following Theorem 10 show that

$$\lim_{n\to\infty} \frac{s(2^n x, i)}{(2^n x)^\tau} = v_i \frac{3\xi^2-2}{\xi^2}\lambda(x) > 0$$

for all $x > 0$ and $0 \leqslant i \leqslant 3$. In every residue class (mod 4), the integers with an even number of occurrences of $P = 111$ predominate, but the excess is smallest for the residue class 3, and largest for the classes 0 and 1.

Further, (86) implies that

$$c_1 < \frac{s(x)}{x^\tau} < c_2, \qquad \tau = .82317245\ldots$$

for positive constants c_1 and c_2 and large enough x. Actually this inequality holds for all $x \geqslant 1$ and appropriate c_1, c_2, as can be seen from the fact that $s(x) \geqslant 1$ for $x \geqslant 0$. This follows in turn by an easy induction argument from the recurrence

$$s(8x) = 2s(2x)+2s(x)+b(x), \qquad |b(x)| \leqslant 3.$$

(See Section 6, (48).)

We now calculate $\lambda(x)$ using (83) for the two special values $x = 1$, $x = 12/7$. When $x = 1$, $x_r = 2^r$ for all $r \geqslant 1$, and certainly $a(2^r+1) = 1$. Equation (83) gives

$$\lambda(1) = \frac{\xi^2}{3\xi^2-2}\left(\xi^2+\xi-\frac{1}{\xi}-\sum_{r=2}^{\infty}\frac{\xi^2-1}{\xi^r}\right) = \frac{\xi}{3\xi^2-2}(\xi^2+\xi) \tag{87}$$

$$= \frac{1}{19}(\xi^2+8\xi+5) = 1.17288075\ldots$$

This shows that the lower bound in (84) is sharp. In particular

$$s(2^n) \simeq \frac{\xi^n}{19}(\xi^2+8\xi+5), \quad \text{as } n \to \infty.$$

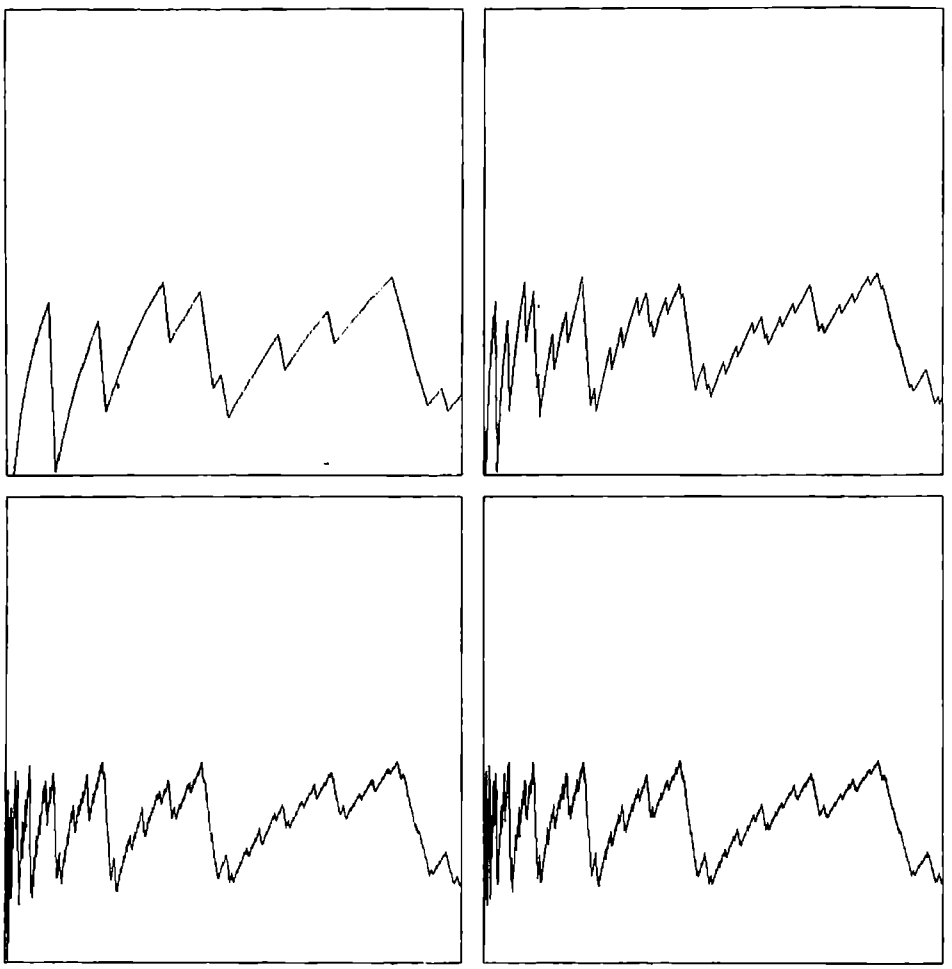


Fig 3. Approximations to $\lambda_{111}(x)$. Pictured are graphs of $s_{111}(n)/n^t$, where n lies in the intervals $[1, 4^{a+3}]$ and $a = 0, 1, 2, 3$. Note the gradual emergence of the property $\lambda(2x) = \lambda(x)$.

Note that the norm of $\lambda(1)$ from the field $Q(\xi)$ is

$$\text{Norm}(\lambda(1)) = \text{Norm}\left(\tfrac{1}{19}(\xi^2+8\xi+5)\right) = \tfrac{1}{19},$$

which is interesting in view of the fact that the discriminant of x^3-2x-2 is $-4\cdot 19$.

If $x = 12/7 = 1.101101101\ldots = 1.\overline{101}_2$, we have

$$x_1 = 11_2, \quad x_2 = 110_2, \quad x_3 = 1101_2, \quad x_4 = 11011_2, \ldots$$

and it is easy to see that $a(x_r+1) = -1$ for $r \geqslant 2$. Since x_r is only even when $r \equiv 2 \pmod 3$, in which case $x_r \equiv 2 \pmod 4$, we find

$$\left(\frac{12}{7}\right)^\tau \lambda\left(\frac{12}{7}\right) = \frac{\xi^2}{3\xi^2-2}\left(\xi^2+\xi+\sum_{r\equiv 0}^{\infty}\frac{1}{\xi^{3r+2}}\right) = \frac{\xi^3}{3\xi^2-2}\left(\xi+1+\frac{1}{2\xi+1}\right)$$

$$= \frac{1}{171}(49\xi^2+88\xi+74) = 2.24027545\ldots,$$

(using $1/(2\xi+1) = (4\xi^2-2\xi-7)/9$), and

$$\lambda(\tfrac{12}{7}) = 1.43750900\ldots \tag{88}$$

From tables of $s(x)/x^\tau$ it appears that (87) and (88) give the actual minimum and maximum values of $\lambda(x)$ for $x > 0$.

10. The exceptional cases

For those patterns with autocorrelations 11, 01, 001, or 0001, we have yet to show that the bound

$$s(x) = O(x^\tau), \qquad \tau = \log\xi/\log 2,$$

cannot be improved, where ξ is the maximum modulus of the roots of the period polynomial. We do this by computing the values $s(2^n)$, $n \geqslant 0$. This in turn requires us to compute $b(2^n)$ and $s_0(2^n)$ for each pattern P.

Table 3

$s(2^n)$ for patterns of length 2

P	$s(2^n)$
00	$2,\ n=0;\quad 3\cdot 2^m-3,\ n=2m\geqslant 2;\quad 4\cdot 2^m-1,\ n=2m+1\geqslant 1$
11	$2^{[(n+1)/2]}+1$
01	$\frac{1+i}{2}(1+i)^n+\frac{1-i}{2}(1-i)^n-1$
10	$2,\ n=0;\quad \frac{1-i}{2}(1+i)^n+\frac{1+i}{2}(1-i)^n-1,\ n\geqslant 1$

For the correlations $\alpha = 11$, 01 these computations are straightforward and are given in Table 3. In each case $b(x)$ is computed using (45) and (46), $s_0(2^n)$ is obtained from (58), and $s(2^n) = t(2^n)+s_0(2^n)$ is evaluated using the fact that $t(x)$ satisfies the respective recursions

$$t(4x)-2t(x) = 0, \qquad \alpha = 11,$$

$$t(4x)-2t(2x)+2t(x) = 0, \qquad \alpha = 01.$$

In all cases with $d = 2$, $\xi = \sqrt{2}$, the roots of the respective polynomials being $\pm\sqrt{2}$ or $1\pm i$. Hence $\tau = \frac{1}{2}$ and

$$s_p(2^n) \geqslant c2^{n/2} = c2^{n\tau}$$

is clear for $P = 00$ or 11. For $P = 01$ or 10 the same estimate follows for an infinite sequence of n's by virtue of

$$s_p(2^{8n}) = 16^n-1, \quad n \geqslant 1, \qquad \text{for } P = 01 \text{ or } 10.$$

Hence $s(x) = \Omega(x^{1/2})$ in all cases with $d = 2$.

It is important to note that modified limit functions $\lambda_p(x)$ exist in the four cases just considered. For example, it can be shown that the limits

$$\lim_{n\to\infty} \frac{s(4^n x)}{\sqrt{4^n x}} = \lambda_p(x) \qquad \text{for } P = 00, 11,$$

$$\lim_{n\to\infty} \frac{s(256^n x)}{\sqrt{256^n x}} = \lambda_p(x) \qquad \text{for } P = 01, 10$$

exist for all $x > 0$, using (15) and the facts

$$A^2 = 2I \qquad \text{for } P = 00, 11,$$

$$A^8 = 16I \qquad \text{for } P = 01, 10.$$

The proof amounts to replacing (15) by a recursion for $\bar{s}(x)$ in which A is replaced by A^2 or A^8, and then using the arguments in Theorem 10. We suppress further details, but refer the reader to [6] for a discussion of the Rudin–Shapiro case $P = 11$.

In the cases with $\alpha = 001$ or 0001, the period polynomials are

$$x^3-2x^2+2 \quad (\alpha = 001) \qquad \text{and} \qquad x^4-2x^3+2 \quad (\alpha = 0001)$$

with respective discriminants -44 and $5\cdot 2^6$ In both cases there are two roots of largest modulos, ξ_1 and ξ_2, which are complex conjugates, the remaining roots having absolute value less than 1. (See below for their explicit values.) The results of Section 7 imply that

$$s(x) = \lambda_1(x)x^{\tau_1}+\lambda_2(x)^{\tau_2}+s_0(x), \qquad x > 0.$$

Since we may take $\tau_2 = \log\xi_2/\log 2 = \log\bar{\xi}_1/\log 2 = \bar{\tau}_1$, equation (60) implies easily that

$$\lambda_2(x) = \bar{\lambda}_1(x), \tag{89}$$

and therefore

$$s(x) = 2\operatorname{Re}(\lambda_1(x)x^{\tau_1}) + s_0(x). \tag{90}$$

Note that (89) holds in general for complex conjugate ξ_1, $\xi_2 = \bar{\xi}_1$, using the well-known fact that Δ (see (60)) is invariant under complex conjugation iff the number of pairs of complex roots of $P(x)$ is even.

Putting $x = 2^n$ in (90) and using $\lambda_1(2^n) = \lambda_1(1)$ gives

$$\frac{s(2^n)}{|\xi_1|^n} = 2\operatorname{Re}(\lambda_1(1)e^{in\operatorname{Arg}\xi_1}) + O(|\xi_1|^{-n}). \tag{91}$$

To show $s(x) = \Omega(x^\tau)$, $\tau = \log|\xi_1|/\log 2$, we prove

LEMMA 12. *If* $\operatorname{Re}\lambda_1(1) \neq 0$, *then* $|s(2^n)| \geqslant c|\xi_1|^n$ *for infinitely many* n, *for some positive constant* c.

Proof. Let $\lambda_1(1) = a+bi$, $a \neq 0$, and $\theta = \operatorname{Arg}\xi_1$; then

$$\operatorname{Re}(\lambda_1(1)e^{in\theta}) = a\cos n\theta - b\sin n\theta.$$

If $\theta/(2\pi) \in \boldsymbol{Q}$, choose n so that $n\theta/(2\pi) \in \boldsymbol{Z}$. Then $\operatorname{Re}(\lambda_1(1)e^{in\theta}) = a$ and the lemma follows easily.

If $\theta/(2\pi) \notin \boldsymbol{Q}$, then the theory of continued fractions implies that the inequality

$$\left|\frac{n\theta}{2\pi} - k\right| < \frac{1}{\sqrt{5}}, \qquad k \in \boldsymbol{Z},$$

is solvable for infinitely many n, i.e.

$$n\theta = 2\pi k_n + \varepsilon_n, \qquad |\varepsilon_n| < \frac{2\pi}{\sqrt{5}n}, \qquad k_n \in \boldsymbol{Z},$$

for infinitely many n (see [16], p. 34). For all such $n > 2\pi/\sqrt{5}$

$$\cos n\theta = \cos\varepsilon_n \geqslant 1 - \frac{\varepsilon_n^2}{2} > 1 - \frac{2\pi^2}{5n^2}, \qquad |\sin n\theta| \leqslant |\varepsilon_n| < \frac{2\pi}{5n}.$$

Hence,

$$|\operatorname{Re}\lambda_1(1)e^{in\theta}| \geqslant |a|\left(1 - \frac{2\pi^2}{5n^2}\right) - |b|\left(\frac{2\pi}{\sqrt{5}n}\right) = |a| + O\left(\frac{1}{n}\right),$$

as $n \to \infty$, for an infinite sequence of n's. From (91) it is then clear that $|s(2^n)| \geqslant |a|\,|\xi_1|^n$ for this same sequence and large enough n. ∎

For the patterns with $\alpha = 001$, we have computed $\lambda_1(1)$ from equation (60), which leads to the formula

$$\lambda_1(1) = \frac{\xi_3 - \xi_2}{-2\sqrt{11}\,i}(\xi_2\xi_3 t(1) - (\xi_3 + \xi_2)t(2) + t(4)).$$

Taking conjugates and adding respectively substracting gives

$$\begin{aligned} \operatorname{Re}\lambda_1(1) &= \frac{1}{2(3\xi^2 - 4\xi)}[(2\xi^2 - 2\xi)t(1) - (\xi - 2)t(2) - t(4)], \\ \operatorname{Im}\lambda_1(1) &= \frac{1}{4\sqrt{11}}[(2\xi^2 - 4\xi)t(1) - (3\xi^2 - 4)t(2) + (3\xi - 2)t(4)], \end{aligned} \tag{92}$$

where $\xi = \xi_3$ is now the real root of $x^3 - 2x^2 + 2 = 0$:

$$\xi = \xi_3 = \tfrac{2}{3} - \tfrac{1}{3}\sqrt[3]{19 - 3\sqrt{33}} - \tfrac{1}{3}\sqrt[3]{19 + 3\sqrt{33}} = -.839286755\ldots$$

To compute the required values of $t(x)$ we have computed $b(x)$ from (45) and (46), and $s_0(1)$, $s_0(2)$, and $s_0(4)$ from (58).

Finally, the values $t(1)$, $t(2)$, and $t(4)$ were computed from (59) and then substituted into (92); this gave the results in Table 4 for $\lambda_1(1)$.

Table 4

$\lambda_1(1)$ for $d = 3$, $\alpha = 001$

P	$\operatorname{Re}\lambda_1(1)$	$\operatorname{Im}\lambda_1(1)$
001	$\frac{1}{22}(-7\xi^2 + \xi + 16)$	$\frac{1}{2\sqrt{11}}(\xi^2 - 5\xi + 2)$
100	$\frac{1}{22}(6\xi^2 - 4\xi + 2)$	$\frac{1}{\sqrt{11}}(-\xi^2 + 2\xi)$
110	$\frac{1}{22}(6\xi^2 - 4\xi + 2)$	$\frac{1}{\sqrt{11}}(-\xi^2 + 2\xi)$
011	$\frac{1}{22}(2\xi^2 - 5\xi + 8)$	$\frac{1}{2\sqrt{11}}(-2\xi^2 + \xi + 2)$

$\xi = -.83928675\ldots,\ \xi^3 - 2\xi^2 + 2 = 0$

We note the curious fact that $\lambda_1(1)$ is the same for the patterns 100 and 110. Thus the leading term in (90) is identical for these two patterns when $x = 2^n$. For the sake of completeness we give the explicit values of ξ_1, $|\xi_1|$ and

Arg ξ_1 using Cardan's formulas:

$$\xi_1 = \frac{4+m+n+i(m-n)\sqrt{3}}{6} = 1.41964337\ldots+i.60629072\ldots,$$

$$|\xi_1| = \tfrac{1}{3}\sqrt{(m-n)^2+(m+2)(n+2)} = 1.54368901\ldots,$$

$$\operatorname{Arg}\xi_1 = \tan^{-1}\sqrt{3}\left(\frac{m-n}{4+m+n}\right) = .40362481\ldots^{R},$$

where

$$m = \sqrt[3]{19+3\sqrt{33}}, \qquad n = \sqrt[3]{19-3\sqrt{33}}.$$

Our computations show that $s(x) = \Omega(x^{.62638\ldots})$.

There are 6 patterns having length $d = 4$ and autocorrelation $\alpha = 0001$. For the sake of simplicity we content ourselves here with the computation of $\operatorname{Re}\lambda_1(1) = \frac{1}{2}t(1)$ (see (90)). Again, the values of $s_0(1)$ were computed from (45), (46) and (58); by these computations $s_0(1)$ is naturally expressed in terms of the power sums

$$\eta_k = \xi_1^k+\xi_2^k, \qquad k = 1,\ 2,\ 3,$$

where ξ_1 and ξ_2 are the roots of largest modulus of $x^4-2x^3+2 = 0$.

In order to see that $\operatorname{Re}\lambda_1(1) \neq 0$ we compute the η_k explicitly, starting with the cubic resolvent $y^3-8y-8 = (y+2)(y^2-2y-4)$ of $x^4-2x^3+2 = 0$, whose roots are $\xi_1\xi_2+\xi_3\xi_4$ and its "conjugates" under the symmetric group S_4. (See [15], p.52.) We find

$$\xi_1\xi_2+\xi_3\xi_4 = 1+\sqrt{5}, \qquad \xi_1\xi_3+\xi_2\xi_4 = -2, \qquad \xi_1\xi_4+\xi_2\xi_3 = 1-\sqrt{5}.$$

From these equations and $\xi_1\xi_2\xi_3\xi_4 = 2$ we find easily that

$$\xi_1\xi_2 = \tfrac{1}{2}(1+\sqrt{5})+\sqrt{\tfrac{1}{2}(-1+\sqrt{5})} = \gamma-\gamma'\sqrt{\gamma},$$

where $\gamma = (1+\sqrt{5})/2$ is the fundamental unit in $Q(\sqrt{5})$ and $\gamma' = (1-\sqrt{5})/2$. Further, the relations

$$(\xi_1+\xi_2)(\xi_3+\xi_4) = -1-\sqrt{5}, \qquad \xi_1+\xi_2+\xi_3+\xi_4 = 2,$$

imply that

$$\eta_1 = \xi_1+\xi_2 = 1+\sqrt{2+\sqrt{5}} = 1+\gamma\sqrt{\gamma},$$

$$\eta_2 = \eta_1^2-2\xi_1\xi_2 = 2+2\sqrt{\gamma},$$

$$\eta_3 = \eta_1\eta_2-\xi_1\xi_2\eta_1 = 4+2\sqrt{\gamma}.$$

Table 5

$\mathrm{Re}\lambda_1(1)$ for $d = 4$, $\alpha = 0001$

P	$\mathrm{Re}\lambda_1(1)$
0001	$\frac{1}{20}(5+24\sqrt{\gamma}-13\gamma\sqrt{\gamma})$
1110	$\frac{1}{20}(5-12\sqrt{\gamma}+9\gamma\sqrt{\gamma})$
0011	$\frac{1}{20}(5+8\sqrt{\gamma}-\gamma\sqrt{\gamma})$
1000	$\frac{1}{20}(15+12\sqrt{\gamma}+\gamma\sqrt{\gamma})$
0111	$\frac{1}{20}(5-6\sqrt{\gamma}+7\gamma\sqrt{\gamma})$
1100	$\frac{1}{20}(5-12\sqrt{\gamma}+9\gamma\sqrt{\gamma})$

$\gamma = \frac{1}{2}(1+\sqrt{5})$

Using these expressions gives the values of $\mathrm{Re}\lambda_1(1)$ listed in Table 5. Since $\sqrt{\gamma}$ has degree 4, it is clear that $\mathrm{Re}\lambda_1(1) \neq 0$ in all cases. Again, for the sake of explicitness we note that

$$\xi_1 = \tfrac{1}{2}(1+\sqrt{2+\sqrt{5}})+\tfrac{1}{2}i(\sqrt{-1+\sqrt{5}+(-3+\sqrt{5})\sqrt{\gamma}}), \qquad \xi_2 = \bar{\xi}_1,$$

$$|\xi_1| = \sqrt{\xi_1\xi_2} = \sqrt{\tfrac{1}{2}(1+\sqrt{5}+\sqrt{2\sqrt{5}-2})} = 1.55054357$$

$$\mathrm{Arg}\,\xi_1 = \tan^{-1}\left(\frac{(\sqrt{-1+\sqrt{5}+(-3+\sqrt{5})\sqrt{\gamma}}}{1+\sqrt{2+\sqrt{5}}}\right) = .16655985\ldots^R,$$

$$\xi_3 = \tfrac{1}{2}(1-\sqrt{2+\sqrt{5}})+\tfrac{1}{2}i(\sqrt{-1+\sqrt{5}+(3-\sqrt{5})\sqrt{\gamma}}), \qquad \xi_4 = \bar{\xi}_3.$$

The Galois group of $P(x) = x^4-2x^3+2 = 0$ is dihedral of order 8, and $\sqrt{\gamma} = \sqrt{(1+\sqrt{5})/2}$ generates the real subfield of the splitting field of $P(x) = 0$.

Note again that $\lambda_1(1)$ is the same for the patterns 1110 and 1100. These computations show that $s(x) \asymp \Omega(x^{\cdot 632774\ldots})$, when $d = 4$ and $\alpha = 0001$.

Putting together the results of this section and the corollary to Theorem 11, we obtain

THEOREM 13. *If P is any pattern, and ξ is the maximum modulus of the roots of the period polynomial $P(x)$,* then $\tau = \log\xi/\log 2$ *is the "correct" order of magnitude of the sum $s(x)$, i.e. $s(x) = \Omega(x^\tau)$.*

COROLLARY. *For any pattern P having length $d \geqslant 3$, the sequence $\{a_p(n)\}$ does not arise from a paper-folding sequence, in the sense of Mendes France* ([9], [27]).

Proof. From the theorem and from the results of Section 5, according to which $\tau > \frac{1}{2}$ if $d \geqslant 3$, $s(x) = O(x^{1/2})$ is false for $d \geqslant 3$. Hence $\{a_p(n)\}$ cannot be a direction sequence for a paper-folding sequence, since the partial sums of all such sequences are $cx^{1/2}$ in size. (See Section 1.) ∎

11. The structure of the autocorrelation tree

We finish by looking at the autocorrelation tree introduced in Section 4 in more detail. In this section we shall deduce several structural statements in a simple manner from Lemma 6 and the forward and backward propagation rules (See Theorem 4.) For convenience we refer to these rules as rules 1 and 2. Some of these results are implicit in [13]. (See especially the recursive predicate Ξ of that paper.)

In Sections 4 and 5 we saw that every correlation is on the red or blue branches or is descended from some correlation of the form $\delta_k = 100\ldots1$ of length $k+1$, $k \geqslant 2$. Let T_k denote the subtree beginning with the node δ_k, $k \geqslant 2$. Then T_k consists of the autocorrelations of all patterns whose least periods are k. Also, let T_0 and T_1 denote the blue and red branches, and let T denote the tree taken as a whole. T_0 and T_1 are obviously periodic subtrees of T, and T is the union of the T_k. (See Figs. 1 and 4.) We shall prove

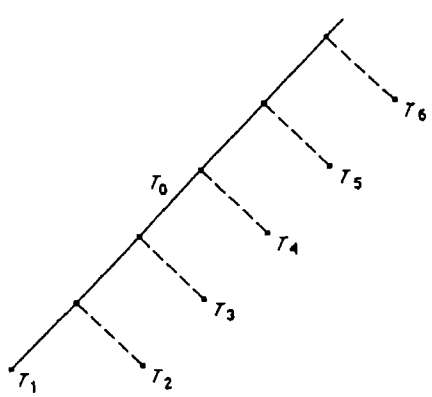


Fig. 4. T as the union of the subtrees T_k.

THEOREM 14. *T_k is a periodic tree, if $k \geqslant 2$: in fact, T_k is the union of an infinite number of isomorphic finite subtrees.*

Furthermore, if $T(k-1)$ denotes the finite subtree of T consisting of the first $k-1$ "levels" of the tree T (all autocorrelations of length $\leqslant k-1$), and $T_k(k-1)$ denotes the first $k-1$ "levels" of T_k (its autocorrelations of length $\leqslant 2k-1$), then

$$T_k(k-1) \simeq T(k-1), \quad k \geqslant 2. \tag{93}$$

We begin by proving the following lemma.

LEMMA 13. *Let $\beta_k = 00\ldots01$ be of length $k \geqslant 2$, and let α be any autocorrelation of length n. If $\gamma = \alpha\beta_k$ is a correlation, and p is any nonzero index for which $\alpha_p = 1$, then either $k|p$ or $p+k > n$.*

Proof. Note that $\alpha_p = \gamma_{p+k}$ for any p, $0 \leqslant p < n$. By assumption γ is the autocorrelation of some binary pattern P, with least period k, since $\gamma_k = \alpha_0 = 1$. Because $\gamma_{p+k} = \alpha_p = 1$, $p+k$ is also a period of P. By Lemma 6, $k|(p+k)$ or $p+k > (n+k)-k = n$. The lemma follows immediately. ∎

Note that Lemma 13 is concerned with exactly the autocorrelations γ in the tree T_k. Next, we prove

LEMMA 14. *If $k \geqslant 2$ and α is any correlation, then $\gamma = \alpha\beta_k$ is a correlation iff $\zeta = \alpha\beta_k\beta_k$ is a correlation, where β_k has the same meaning as in Lemma 13.*

Proof. Since any substring of a correlation which ends in a 1 is also a correlation, one direction is clear. It remains to show that ζ is a correlation if γ is. To do this we must verify that ζ satisfy the rules 1 and 2. Let n denote the length of α, and keep in mind that $\alpha_p = \gamma_{p+k} = \zeta_{p+2k}$, and $\gamma_p = \zeta_{p+k}$. ∎

Rule 1. Aside from trivial cases we must show that $\zeta_0 = \zeta_r = 1$ implies $\zeta_{ir} = 1$ for all ir in the range $0 \leqslant ir < n+2k$. If $r = k$ this is clear, since $\zeta_k = \gamma_0 = 1$, $\zeta_{ik} = \gamma_{(i-1)k}$, $i \geqslant 1$, and γ is a correlation with $\gamma_k = 1$. We can assume then that $r = p+2k$, $p \geqslant 0$, and that $2(p+2k) < n+2k$, or $2p+2k < n$. (Otherwise 0 and $p+2k$ are the only indices in the range $[0, n+2k)$.) But now $\alpha_0 = 1$ and $\alpha_p = \zeta_{p+2k} = 1$ imply $\alpha_{2p} = 1$ (since α satisfies rule 1); $\gamma_0 = 1$ and $\gamma_{p+k} = \alpha_p = 1$ imply $\gamma_{2p+2k} = \alpha_{2p+k} = 1$ (since γ satisfies rule 1); and $\alpha_{2p} = 1$ and $\alpha_{2p+k} = 1$ imply that $\alpha_{2p+2k} = 1$. Therefore $\alpha_p = 1$ and $\alpha_{2p+2k} = 1$ give together that $\alpha_{p+i(p+2k)} = \zeta_{(p+2k)(i+1)} = 1$, which was to be shown.

Rule 2. Now we must show that $\zeta_p = \zeta_q = 1$ $(p < q)$ and $\zeta_{p-(q-p)} = \zeta_{2p-q} = 0$ imply that $\zeta_r = 0$ for all r of the form $r = p-i(q-p)$ which lie in the range $0 \leqslant r \leqslant 2p-q$, with $1 \leqslant i \leqslant s$ and $s = [(n+2k-p)/(q-p)]$. This situation can arise only if $p, q \geqslant 2k$. Replacing p, q by $p+2k$, $q+2k$, we must show

$$\alpha_p = \alpha_q = 1,\ \zeta_{2p-q+2k} = 0 \ \rightarrow\ \zeta_{t+2k} = 0, \tag{94}$$

for all $t+2k = p+2k-i(q-p)$ in the range $[0, 2p-q+2k]$, with $1 \leqslant i \leqslant s$ and

$s = [(n-p)/(q-p)]$. Notice that

$$\alpha_p = \alpha_q = 1, \quad \gamma_{2p-q+k} = 0 \to \gamma_{t+k} = 0,$$

for all $t+k = p+k-i(q-p)$ in the range $[0, 2p-q+k]$, $1 \leqslant i \leqslant s$, since γ satisfies rule 2. Thus we need only check that $\zeta_{t+2k} = 0$ when $0 \leqslant t+2k < k$. From the form of ζ this amounts to showing that none of the subscripts $t+2k$ is 0.

We have

$$t+2k = p+2k-i(q-p) \geqslant p+2k-s(q-p)$$
$$\geqslant p+2k-(n-p) = 2p+2k-n.$$

By Lemma 13, either $k|p$ or $p+k > n$. In the second case $t+2k > p+k \geqslant 2$, as required. If $k|p$, then by Lemma 13 again, $k|q$ or $q+k > n$. If $k|q$, then $\zeta_{2p-q+2k}$ has a subscript divisible by k and must equal 1 (ζ satisfies rule 1). Hence we can assume $k|p$ and $q+k > n$. In this case

$$t+2k \geqslant 2p+2k-n > 2p+2k-(q+k) = 2p-q+k.$$

If $2p-q+k \geqslant 0$ we're done. If not, then $q-p > p+k \geqslant k$, and

$$s \leqslant \frac{n-p}{q-p} < \frac{q-p}{q-p} + \frac{k}{q-p} < 2,$$

giving $s = 1$, in which case (94) holds trivially. ∎

Let β_k^j represent the correlation obtained by repeating β_k j times, and $\delta_k^j = \delta_k \beta_k^{j-1} = 1\beta_k^j$, $j \geqslant 1$, where β_k^0 is empty. If T_k^j represents the subtree of T starting with the node δ_k^j (so $T_k^1 = T_k$), then Lemma 14 implies that

$$T_k^j \simeq T_k^{j+1}, \quad j \geqslant 1, k \geqslant 2,$$

by the isomorphism

$$\alpha\beta_k^j \leftrightarrow \alpha\beta_k^{j+1}, \quad \alpha\beta_k^j \in T_k^j, j \geqslant 1. \tag{95}$$

This map clearly preserves the tree structure, because two correlations $\gamma\beta_k$ and $\zeta\beta_k$ are connected by one of the red or blue rules iff γ and ζ are connected by the same rule. (See (29)–(29″).) In particular, $T_k \simeq T_k^j$ for $j \geqslant 1$. This is the sense in which T_k is periodic. (See Figs. 5 and 6, p. 54.)

Now we let U_k^j denote the subtree of all correlations which are descended from δ_k^j but not from δ_k^{j+1}, including both δ_k^j and δ_k^{j+1}. Except for δ_k^{j+1}, all such correlations have the form $\alpha\beta_k^j$, where α is a correlation which is either a substring of β_k, or has a nonzero subscript $m < k$ for which $\alpha_m = 1$. (Note that any correlation of the form $\alpha\beta_k^j$ with $\alpha_0 = 1$ satisfies $\alpha_r = 1$ for all r divisible by k, by the forward propagation rule.) The map (95) gives an isomorphism between U_k^j and U_k^{j+1}, and T_k is the union of the subtrees U_k^j. The

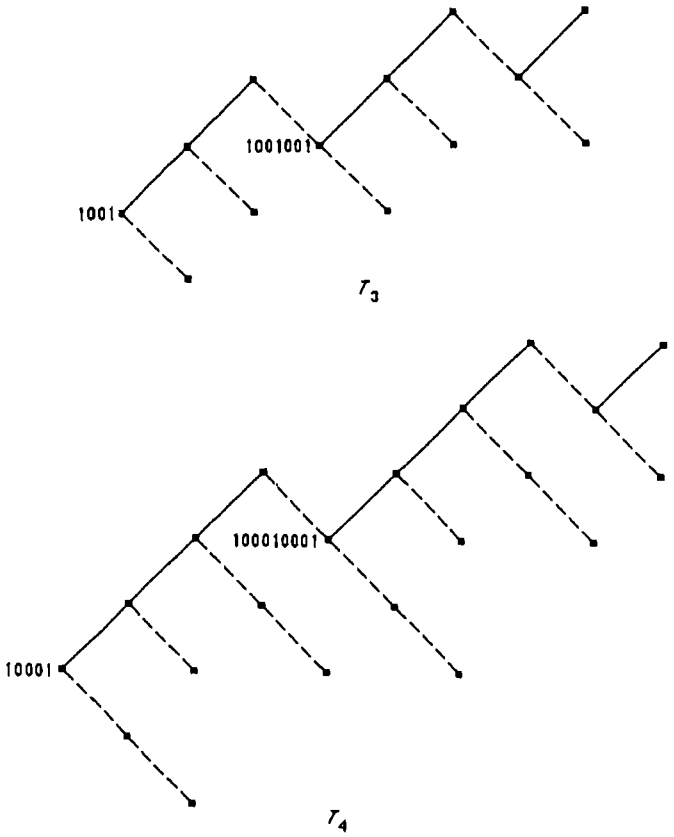


Fig. 5. The periodic subtrees T_3 and T_4. Key: blue —, red

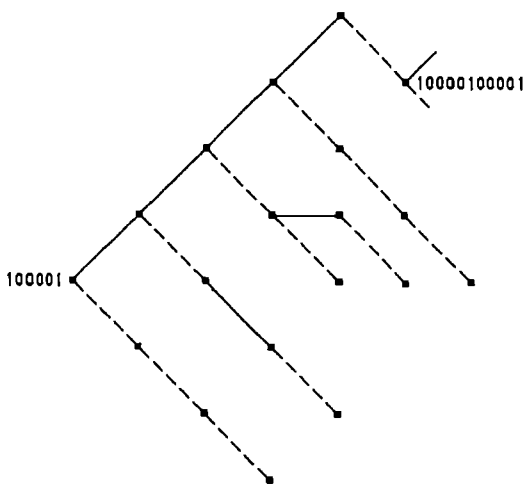


Fig. 6. The periodic subtree T_5. Key: blue —, red

next lemma shows that $U_k = U_k^1$ is *finite*. Hence T_k is the union of an infinite number of isomorphic finite subtrees, proving the first assertion of Theorem 14.

LEMMA 15. *Let $\alpha\beta_k$ be a correlation for which $\alpha_0 = 1$ and $k \geqslant 2$. If the smallest nonzero index m for which $\alpha_m = 1$ exists and is less than k, and if n is the length of α, then either*

a) *$m|k$ and $n \leqslant k-1$, or*

b) *$m \nmid k$ and $n \leqslant 2k-2$.*

Proof. a) If $m|k$, we use the fact that $\gamma = \alpha\beta_k$ satisfies the backward propagation rule. Because $\gamma_k = \gamma_{m+k} = 1$, $\gamma_{k-m} = 0$, we have $\gamma_{k-im} = 0$ for

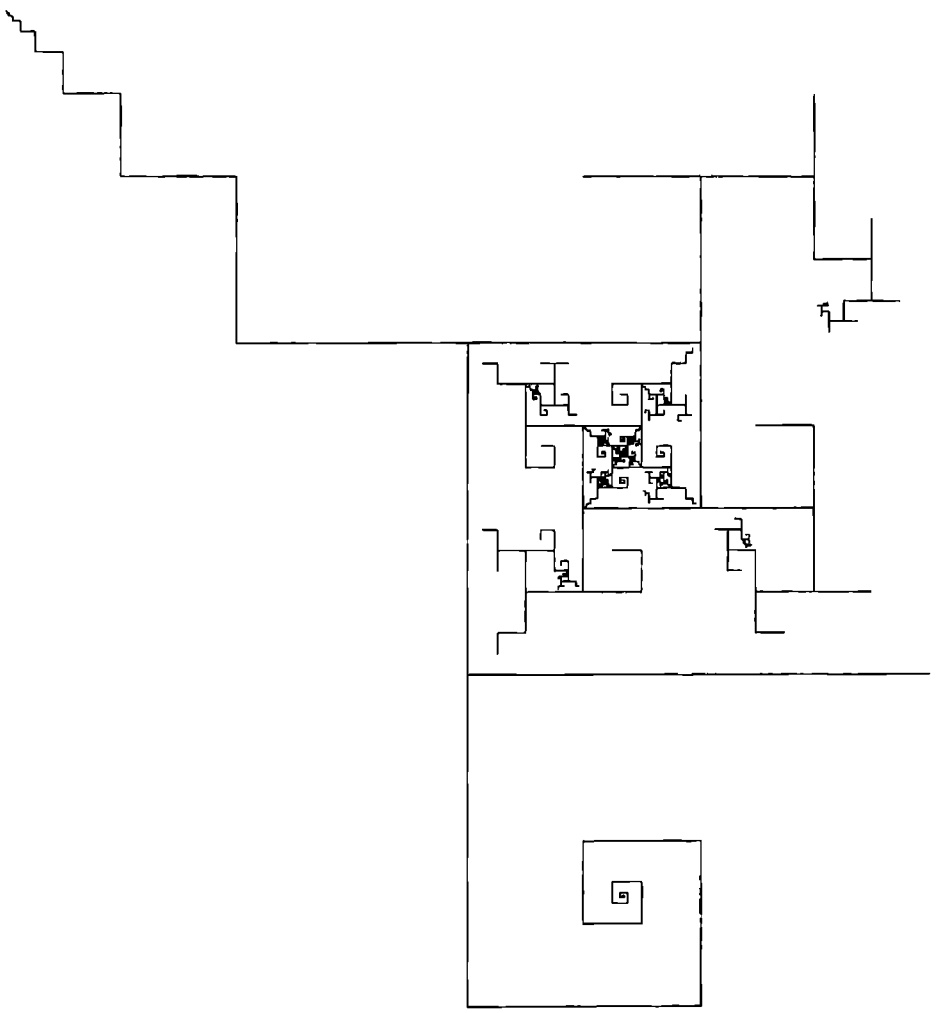


Fig. 7. The autocorrelation tree ($d \leqslant 23$). Key: blue – rotation by -90 , red – rotation by $+90°$

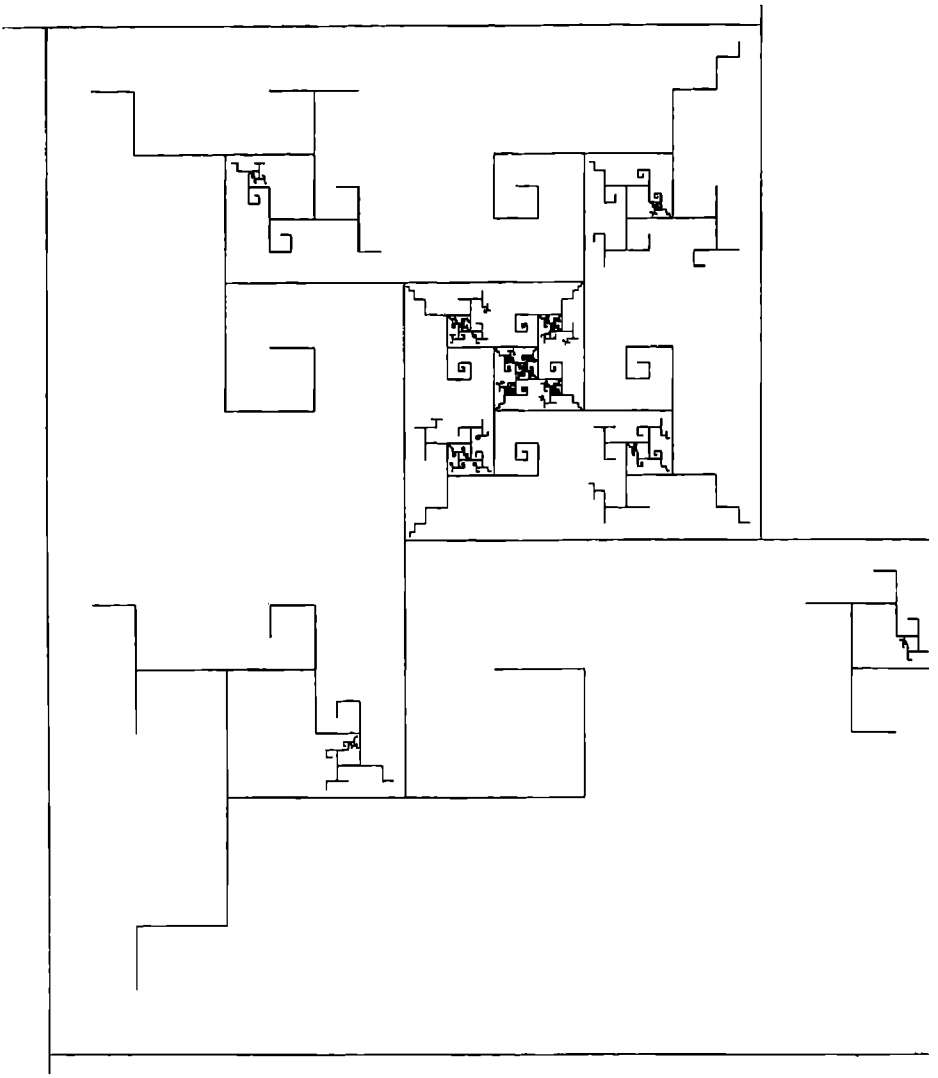


Fig. 8. Expanded view of the blue branch and descendants ($d \geqslant 23$).

$1 \leqslant i \leqslant s$, with $s = [(n+k-k)/(m+k-k)] = [n/m]$. Since $\gamma_0 = 1$ and $k-rm = 0$ for some r, we must have $s \leqslant r-1$, i.e.

$$\left[\frac{n}{m}\right] \leqslant \frac{k}{m}-1, \quad \text{or} \quad \frac{n}{m}-1 < \frac{k}{m}-1,$$

and this implies $n < k$.

b) If $m \nmid k$, then $\alpha_m = 1$ and Lemma 13 imply that $n \leqslant m+k-1 \leqslant 2k-2$. ∎

By this lemma and the fact that $U_k \simeq U_k^j$, $j \geqslant 1$, the only branch beginning at δ_k^j which does not terminate is the branch leading to δ_k^{j+1}. Hence

T_k has an infinite branch H_k, which contains all the correlations δ_k^j, $j \geqslant 1$, and all other branches terminate. It is easy to see that this infinite branch contains exactly those autocorrelations which are their own autocorrelations, i.e. which have the same periods as the set of patterns they encode. For any length d, there are exactly d correlations which this property, namely the correlations of length d on the infinite branches H_k, for $0 \leqslant k \leqslant d-1$.

We complete our discussion of the structure of T by proving (93). To do this it is sufficient to prove

LEMMA 16. *If α is any correlation of length $n \leqslant k-1$, then $\alpha\beta_k$ is also a correlation, for $k \geqslant 2$.*

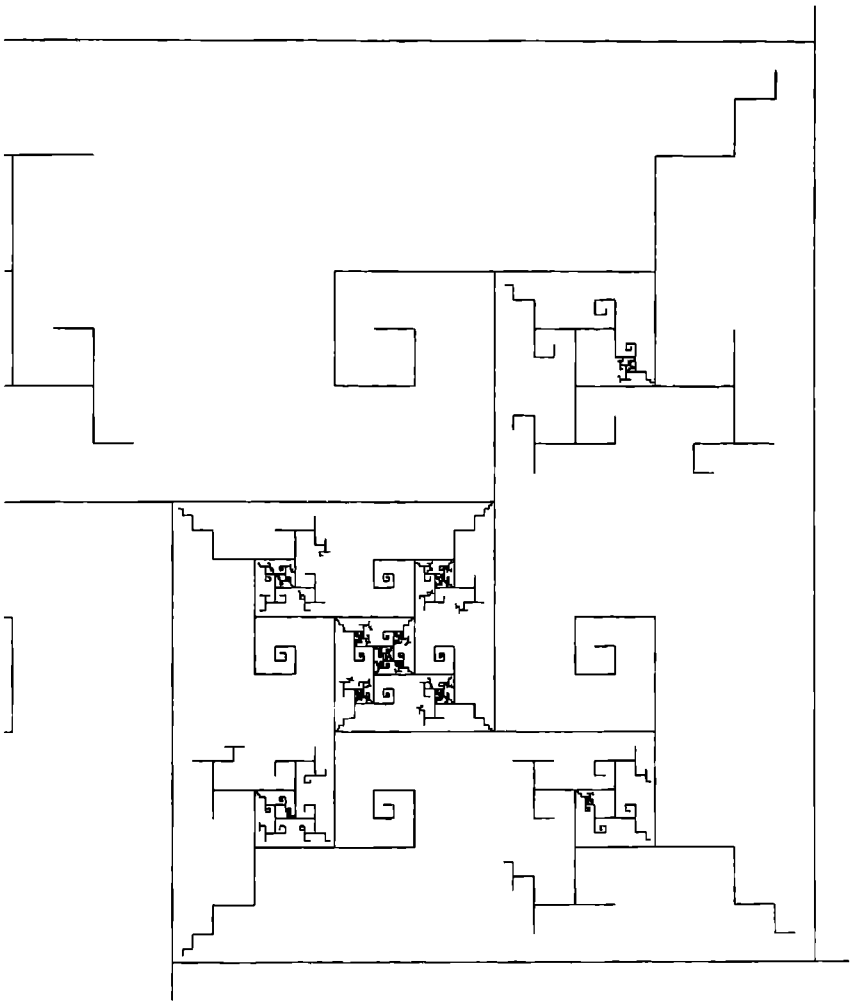


Fig. 9. Expanded view of branches descended from $\alpha = 00001$ ($d \leqslant 23$).

Remark. By this result $\alpha \leftrightarrow \alpha\beta_k$ provides an isomorphism between $T(k-1)$ and $T_k(k-1)$. (See equation (95).)

Proof. We must show that $\gamma = \alpha\beta_k$ satisfies rules 1 and 2. Rule 1 is easily verified, since α satisfies this rule itself.

As in the proof of Lemma 14, verifying rule 2 amounts to showing that

$$\alpha_p = \alpha_q = 0,\ \gamma_{2p-q-k} = 0 \ \rightarrow\ \gamma_{t+k} = 0,$$

for all $t+k = p+k-i(q-p)$ in the range $[0, 2p-q+k]$, where $1 \leqslant i \leqslant s$ and $s = [(n+k-(p+k))/(q+k-(p+k))] = [(n-p)/(q-p)]$. These indices satisfy the inequality

$$t+k \geqslant p+k-s(q-p) \geqslant k+p-(n-p) = k-n+2p \geqslant 1.$$

Therefore either $\gamma_{t+k} = \alpha_t = 0$ (because α satisfies rule 2) or $1 \leqslant t+k < k$, in which case $\gamma_{t+k} = 0$ by inspection. This completes the proof of the lemma. ∎

This lemma is sharp, because the string $\alpha\beta_k$ is not a correlation, if $\alpha = 111\ldots1$ has length k.

Appendix

THEOREM. *If $d \geqslant 3$, and ξ is the maximum modulus of the roots of the period polynomial $P_d(x)$, then $\tau = \log\xi/\log 2$ is transcendental.*

Proof. Since $2^\tau = \xi$ is algebraic, the Gelfond–Schneider theorem ([21]) implies that τ is either rational or transcendental. Assume

$$\tau = \frac{\log\xi}{\log 2} = \frac{p}{q}, \quad (p, q) = 1, \ \ p, q > 0. \tag{96}$$

First consider the cases in which ξ is the dominant real root of $P_d(x)$. Since the constant term of $P_d(x)$ is ± 2, the norm of ξ from the field $Q(\xi)$ is ± 2. By (96) $\xi^q = 2^p$, taking norms gives $\pm 2^q = 2^{dp}$. Therefore $q = dp$, giving that $p = 1$ and $\tau = 1/d \leqslant 1/3$. But $\tau > 1/2$ if $d \geqslant 3$ by the results of Section 5.

In the two exceptional cases $\xi^2 = \xi_1\xi_2$, where ξ_1 and ξ_2 are complex conjugates, and (96) gives $(\xi_1\xi_2)^q = 2^{2p}$ Now take norms from the normal closure of $Q(\xi_1)$. This normal closure has degree 6 if $d = 3$ and degree 8 if $d = 4$ (See Section 10). In both cases $N\xi_1 = N\xi_2 = 2^2$, and we find

$$2^{4q} = 2^{12p} \ \text{ for } d = 3, \qquad 2^{4q} = 2^{16p} \ \text{ for } d = 4.$$

Hence $\tau = p/q = 1/3$, respectively $1/4$, both of which are false by the computations of Section 10.

In either case (96) is impossible, so that τ must be transcendental. ∎

References

[1] J. P. Allouche and J. O. Shallit, *Infinite products associated with counting blocks in binary strings*, J. London Math. Soc. (to appear).

[2] J. Brillhart and L. Carlitz, *Note on the Shapiro polynomials*, Proc. Amer. Math. Soc. 25 (1970), 114–118.

[3] J. Brillhart, *On the Rudin–Shapiro polynomials*, Duke Math. J. 40 (1973), 335–353.

[4] J. Brillhart, J. S. Lomont and P. Morton, *Cyclotomic properties of the Rudin–Shapiro polynomials*, J. Reine Angew. Math. 288 (1976), 37–65.

[5] J. Brillhart and P. Morton, *Über Summen von Rudin–Shapiroschen Koeffizienten*, Illinois J. Math. 22 (1978), 126–148.

[6] J. Brillhart, P. Erdös and P. Morton, *On sums of Rudin–Shapiro coefficients, II*, Pacific J. Math. 107 (1983), 39–69.

[7] G. Christol, T. Kamae, M. Mendes France and G. Rauzy, *Suites algebriques, automata et substitutions*, Bull. Soc. Math. France 108 (1980), 401–419.

[8] J. Coquet, *A summation formula related to the binary digits*, Invent. Math. 73 (1983), 107–115.

[9] M. Dekking, M. Mendes France and A. van der Poorten, *Folds!, I, II, III*, Math. Intelligencer 4 (1982), 130–138, 173–181, 190–195.

[10] H. Delange, *Sur la function sommatoire de la fonction "summe des chiffres"*, Enseign. Math. (2) 21 (1975), 31–47.

[11] F. R. Gantmacher, *The Theory of Matrices*, Vol. II, Chelsea, 1964, p. 53.

[12] L. J. Guibas and A. M. Odlyzko, *Long repetitive patterns in random sequences*, Z. Wahrsch. Verw. Gebiete 53 (1980), 241–262.

[13] —, —, *Periods in strings*, J. Combin. Theory Ser. A 30 (1981), 19–42.

[14] —, —, *String overlaps, pattern matching, and nontransitive games*, J. Combin. Theory (A) 30 (1981), 183–208.

[15] I. Kaplansky, *Fields and Rings*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago 1972.

[16] A. Ya. Khinchine, *Continued Fractions*, University of Chicago Press, Chicago 1964.

[17] K. Knopp, *Infinite Sequences and Series*, Dover, 1956.

[18] D. H. Lehmer, *Mahler's matrices*, J. Austral. Math. Soc. 1 (1960), 385–395.

[19] D. H. and Emma Lehmer, *Picturesque exponential sums, I*, Amer. Math. Monthly 86 (1979), 725–733.

[20] —, —, *Picturesque exponential sums, II*, J. Reine Angew. Math. 318 (1980), 1–19.

[21] W. J. Leveque, *Topics in Number Theory*, vol. II, Addison-Wesley Publishing Co., 1956, p. 198.

[22] J. E. Littlewood, *Some Problems in Real and Complex Analysis*, Heath Math. Monographs, D. C. Heath & Co. 1968.

[23] K. Mahler, *A matrix representation of the primitive residue classes* (mod $2n$), Proc. Amer. Math. Soc. 8 (1957), 525–531.

[24] B. B. Mandelbrot, *The Fractal Geometry of Nature*. W. H. Freeman and Co., 1983.

[25] M. Marden, *Geometry of Polynomials*, Amer. Math. Soc. Mathematical Surveys no. 3, 1966.

[26] M. Mendes France and A. J. van der Porten, *Arithmetic and analytic properties of paper folding sequences*, Bull. Austral Math. Soc. 24 (1981), 123–131.
[27] M. Mendes France, *Paper folding, space-filling curves and Rudin–Shapiro sequences*, Contemp. Math. 9 (1982), 85–95.
[28] M. Mendes France and J. O. Shallit, *Some planar curves associated with sums of digits*, announcement (preprint).
[29] L. M. Milne-Thomson, *The Calculus of Finite Differences*, Chelsea.
[30] I. Niven, *Irrational Numbers*, Carus Monographs, No. 11, 1956, Ch. 8.
[31] D. Rider, *Closed subalgebras of $L^1(T)$*, Duke Math. J. 19 (1966), 347–355.
[32] W. Rudin, *Some theorems on Fourier coefficients*, Proc. Amer. Math. Soc. 10 (1959), 855–859.
[33] H. S. Shapiro, *Extremal problems for polynomials and power series*, Master's thesis, M.I.T., 1951.
[34] E. C. Titchmarsh, *Theory of Functions*, Oxford University Press, 1968.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, B.C., CANADA V6T 1Y4

IBM CORPORATION
DEPARTMENT 17W-BLDG. 630, ROUTE 52
HOPEWELL JUNCTION, NY 12533-0999, U.S.A.

DEPARTMENT OF MATHEMATICS
WELLESLEY COLLEGE
WELLESLEY, MA 02181, U.S.A.