

## On a problem of Steinhaus concerning binary sequences

Shalom Eliahou      Delphine Hachez<sup>1</sup>

### Abstract

A finite  $\pm 1$  sequence  $X$  yields a binary triangle  $\Delta X$  whose first row is  $X$ , and whose  $(k + 1)$ st row is the sequence of pairwise products of consecutive entries of its  $k$ th row, for all  $k \geq 1$ . We say that  $X$  is *balanced* if its derived triangle  $\Delta X$  contains as many  $+1$ 's as  $-1$ 's. In 1963, Steinhaus asked whether there exist balanced binary sequences of every length  $n \equiv 0$  or  $3 \pmod{4}$ . While this problem has been solved in the affirmative by Harborth in 1972, we present here a different solution. We do so by constructing *strongly balanced* binary sequences, *i.e.* binary sequences of length  $n$  all of whose initial segments of length  $n - 4t$  are balanced, for  $0 \leq t \leq n/4$ . Our strongly balanced sequences do occur in every length  $n \equiv 0$  or  $3 \pmod{4}$ . Moreover, we provide a complete classification of sufficiently long strongly balanced binary sequences.

*2000 Mathematics Subject Classification:* 05A05, 05A15.

*Keywords and phrases:* Steinhaus, balanced binary sequence, derived sequence, derived triangle.

## 1 Introduction

Let  $X = (x_1, x_2, \dots, x_n)$  be a *binary sequence* of length  $n$ , *i.e.* a sequence with  $x_i = \pm 1$  for all  $i$ . We define the *derived sequence*  $\partial X$  of  $X$  by  $\partial X = (y_1, y_2, \dots, y_{n-1})$  where  $y_i = x_i x_{i+1}$  for all  $i$ . By convention, we agree that  $\partial X = \emptyset$  whenever  $n = 0$  or  $1$ , where  $\emptyset$  stands for the

---

<sup>1</sup>The present work is part of the PhD thesis work of the second author at the Université du Littoral Côte d'Opale under the direction of Professor S. Eliahou.

empty binary sequence of length 0. More generally, for  $k \geq 0$ , we shall denote by  $\partial^k X$  the  $k$ th derived sequence of  $X$ , defined recursively as usual by  $\partial^0 X = X$  and  $\partial^k X = \partial(\partial^{k-1} X)$  for  $k \geq 1$ .

We shall denote by  $\Delta X$  the collection of the derived sequences  $X, \partial X, \dots, \partial^{n-1} X$  of  $X$ . This collection may be pictured as a triangle, as in the following example: if  $X = (+1, +1, -1, +1, -1, +1, +1)$ , abbreviated as  $++-+-++$ , then  $\Delta X =$

$$\begin{array}{c} ++-+-++ \\ +- - - -+ \\ -+++ - \\ -++- \\ -+- \\ -- \\ + \end{array}$$

We shall henceforth refer to  $\Delta X$  as the *derived triangle* of  $X$ . If  $Y = (y_1, \dots, y_m)$  is any finite collection of numbers, we denote the sum of its entries by  $S(Y) = \sum_{i=1}^m y_i$ . For instance, if  $X = (x_1, x_2, \dots, x_n)$  is a binary sequence, then  $S(\Delta X)$  represents the sum of the entries in the derived triangle  $\Delta X$  of  $X$ , *i.e.*  $S(\Delta X) = \sum_{k=0}^{n-1} S(\partial^k X)$ .

**Definition 1.1** A binary sequence  $X = (x_1, x_2, \dots, x_n)$  is balanced if  $S(\Delta X) = 0$ . In other words,  $X$  is balanced if its derived triangle  $\Delta X$  contains as many  $+1$ 's as  $-1$ 's.

For example, the above binary sequence  $X = ++-+-++$  is balanced, as its derived triangle contains 14 positive signs and 14 negative signs in total. This sequence, as well as other balanced sequences of length 11, 12, 19 and 20, appear in [3], where the author proposed the following problem.

**Problem** Is there a balanced binary sequence of length  $n$  for every  $n \equiv 0$  or  $3 \pmod{4}$ ?

(The term “balanced” is not used by Steinhaus.) Note that the condition  $n \equiv 0$  or  $3 \pmod{4}$  is necessary for the existence of a balanced binary sequence  $X$  of length  $n$ . Indeed, the derived triangle of  $X$  contains  $n(n+1)/2$  entries; if  $n \equiv 1$  or  $2 \pmod{4}$ , this number of entries is odd, and therefore  $S(\Delta X)$  cannot vanish.

The above problem has been solved in the affirmative in [Harborth 1972]. In this paper, we shall present a new solution to the problem of

Steinhaus, by constructing binary sequences satisfying a much stronger condition.

**Definition 1.2** *A binary sequence  $X = (x_1, \dots, x_n)$  is strongly balanced if the initial segment  $(x_1, \dots, x_{n-4t})$  of  $X$  is balanced, for every  $0 \leq t \leq n/4$ .*

Alternatively, strongly balanced sequences may be defined recursively, as follows. As initial conditions, balanced sequences of length 0 or 3 are considered as strongly balanced. For  $n \geq 4$ , the sequence  $(x_1, \dots, x_n)$  is defined as strongly balanced if and only if it is balanced and  $(x_1, \dots, x_{n-4})$  is strongly balanced.

For instance, the above binary sequence  $X = ++-+-++$  is strongly balanced of length 7, as  $X$  and its initial segment of length 3, namely  $++-$ , are both balanced. Another example of a strongly balanced binary sequence is given by  $P = +-+++++---$ , of length 12. Indeed, the initial segments of length 4, 8 and 12 of  $P$ , namely  $-+++$ ,  $-+++++$  and  $P$  itself, are all balanced as easily seen. On the other hand, the sequences  $Y_7 = +++-+-$  and  $Y_8 = ++++--$  are both balanced, but not strongly so. Indeed, the initial segments of length 3 of  $Y_7$  and length 4 of  $Y_8$  are both constant  $+1$  sequences, and therefore cannot be balanced.

We shall denote by  $sb(n)$  the number of strongly balanced binary sequences of length  $n$ . There is no a priori reason to expect that strongly balanced sequences should exist at all for  $n$  large. But fortunately, the task of searching for all such sequences lends itself very well to computer experimentation (see below).

The outcome of our experiments is quite surprising. Initially, the number  $sb(n)$  for  $n \equiv 0 \pmod{4}$  strictly increases, from  $n = 4$  up to  $n = 36$ . Then it starts to decrease (non-strictly) up to length  $n = 92$ , where it finally stabilizes to the constant  $sb(n) = 4$  for all  $n = 4m \geq 92$ . For  $n \equiv 3 \pmod{4}$ , the situation is similar, though more complicated: provided  $n \geq 127$ , we find that  $sb(n) = 14$  if  $n \equiv 3, 7 \pmod{12}$ , and  $sb(n) = 12$  if  $n \equiv 11 \pmod{12}$ .

A convenient way to summarize the behavior of the numbers  $sb(n)$  is to exhibit properties of their generating function  $g(t) = \sum_{n=0}^{\infty} sb(n)t^n$ . For example, the eventual periodicity of  $sb(n)$  for  $n$  large is reflected by the property of the generating function  $g(t)$  of being a rational function. Our main result in this paper is the following

**Theorem 1.3** *The generating function  $g(t) = \sum_{n=0}^{\infty} sb(n)t^n$  of the number  $sb(n)$  of strongly balanced binary sequences of length  $n$  is given by the following rational function:*

$$g(t) = 4t^{92}/(1-t^4) + f_0(t) + (14 + 12t^4 + 14t^8)t^{127}/(1-t^{12}) + f_3(t),$$

where  $f_0(t)$  and  $f_3(t)$  are the following polynomials:

$$\begin{aligned} f_0(t) = & 1 + 6t^4 + 18t^8 + 30t^{12} + 52t^{16} + 80t^{20} + 88t^{24} + 106t^{28} \\ & + 116t^{32} + 124t^{36} + 106t^{40} + 92t^{44} + 92t^{48} + 90t^{52} + 64t^{56} \\ & + 44t^{60} + 38t^{64} + 32t^{68} + 20t^{72} + 20t^{76} + 8t^{80} + 8t^{84} + 6t^{88}, \end{aligned}$$

$$\begin{aligned} f_3(t) = & 4t^3 + 8t^7 + 16t^{11} + 26t^{15} + 36t^{19} + 48t^{23} + 48t^{27} + 66t^{31} \\ & + 88t^{35} + 108t^{39} + 114t^{43} + 90t^{47} + 88t^{51} + 104t^{55} + 92t^{59} \\ & + 60t^{63} + 48t^{67} + 28t^{71} + 26t^{75} + 26t^{79} + 20t^{83} + 16t^{87} \\ & + 18t^{91} + 14t^{95} + 14t^{99} + 14t^{103} + 14t^{107} + 16t^{111} + 14t^{115} \\ & + 14t^{119} + 16t^{123}. \end{aligned}$$

In the above formula for  $g(t)$ , the terms  $t^n$  are separated according as  $n \equiv 0$  or  $3 \pmod{4}$ , for better readability and because their behavior is different.

**Corollary 1.4** *For every natural number  $n \equiv 0$  or  $3 \pmod{4}$ , there exists a strongly balanced binary sequence of length  $n$ .*

*Proof:* Consider first the case  $n \equiv 0 \pmod{4}$ . By expanding the summand  $4t^{92}/(1-t^4)$  as  $4t^{92} + 4t^{96} + 4t^{100} + \dots$  in the formula for  $g(t)$ , we see that  $sb(n) = 4$  for every  $n = 4m \geq 92$ , as stated earlier. And the summand  $f_0(t)$  in  $g(t)$  gives the exact value of  $sb(n)$  for  $0 \leq n = 4m \leq 88$ , which is nowhere zero. Similarly, for the case  $n \equiv 3 \pmod{4}$ , we see that  $sb(n) = 14$  for every  $n \equiv 3$  or  $7 \pmod{12}$  with  $n \geq 127$ , and  $sb(n) = 12$  for every  $n \equiv 11 \pmod{12}$  with  $n \geq 131$ . This follows from expanding the summand  $(14 + 12t^4 + 14t^8)t^{127}/(1-t^{12})$  as an infinite series. Smaller values of  $n$  are taken care of by the polynomial  $f_3(t)$ . For example,  $sb(51) = 88$ ,  $sb(55) = 104$  and  $sb(59) = 92$ . Alternatively, one may note that, if there exists a strongly balanced binary sequence  $X$  of length  $n$ , then the initial segment of length  $n - 4$  of  $X$  is also a strongly balanced binary sequence. This follows directly from the definition.  $\square$

The set of all strongly balanced binary sequences of small length  $n$  ( $n \leq 127$ , say) may be constructed by the method described in Section 3.

The eventual periodicity of  $sb(n)$  is a consequence of Theorems 2.1 and 2.2 below.

## 2 A classification of long strongly balanced sequences

In this section, we shall describe the set of all strongly balanced binary sequences of length  $n \geq 92$  for  $n \equiv 0 \pmod{4}$ , and  $n \geq 127$  for  $n \equiv 3 \pmod{4}$ . These two sets admit periodic structures. In order to present the results, we introduce the following notation.

**Notation** If  $P, Q$  are finite binary sequences, we shall denote by  $PQ^\infty$  the infinite eventually periodic sequence which starts with  $P$  and continues with  $Q$  repeated periodically thereafter. If  $R$  is yet another finite binary sequence, and if  $k \in \mathbb{N}$ , we shall denote by  $PQ^kR$  the sequence starting with  $P$ , continuing with  $Q$  repeated  $k$  times, and ending with  $R$ . Finally, if  $T = (t_1, \dots, t_m, \dots)$  is any finite or infinite sequence of length  $\geq m$ , we shall denote by  $T[m] = (t_1, \dots, t_m)$  the initial segment of length  $m$  of  $T$ .

### 2.1 The case $n \equiv 0 \pmod{4}$

Let  $Q_1, \dots, Q_4$  denote the following infinite eventually periodic binary sequences. We will show that every initial segment  $Q_i[n]$  with  $n \equiv 0 \pmod{4}$  is strongly balanced, and that there are no other strongly balanced binary sequences of length  $n$ , provided  $n = 4m \geq 92$ . These statements are formalized in the next theorem.

$$\begin{aligned} Q_1 &= +-++(+ + - + + - + - - - ++)^{\infty}, \\ Q_2 &= (+-++-++++- - -)^{\infty}, \\ Q_3 &= +-+-(+ - - + + + + - - + +)^{\infty}, \\ Q_4 &= +-+-(- + - + + + - + - + +)^{\infty}. \end{aligned}$$

**Theorem 2.1** *For every  $n \equiv 0 \pmod{4}$ , the initial segment of length  $n$  of each of  $Q_1, Q_2, Q_3$  and  $Q_4$  is a strongly balanced binary sequence. Conversely, every strongly balanced binary sequence of length  $n$  with  $n \equiv 0 \pmod{4}$  and  $n \geq 92$  is an initial segment of either of  $Q_1, Q_2, Q_3$  or  $Q_4$ .*

Parts of the proof of this result can be found in Section 6.

## 2.2 The case $n \equiv 3 \pmod{4}$

This case is more complicated. Let  $R_1, \dots, R_{12}$  denote the following infinite eventually periodic binary sequences. Their initial segments of length  $n \equiv 3 \pmod{4}$  are all strongly balanced. Moreover, they account for all sufficiently long strongly balanced binary sequences, except for two more exotic ones in length  $n \equiv 3 \pmod{12}$  and  $n \equiv 7 \pmod{12}$ . For instance, one of these extra sequences for  $n \equiv 3 \pmod{12}$  is  $R_5[n-4] + -+-$ , that is, the initial segment of length  $n-4$  of  $R_5$  appended with the sequence  $+ - + -$ .

$$\begin{aligned}
R_1 &= ++-(+ - + + + - + + + -)^\infty, \\
R_2 &= +- - - - +(+ + - - + - + - + - -)^\infty, \\
R_3 &= +-+(+ + + - + - + + + -)^\infty, \\
R_4 &= +- + + + + - \\
&\quad (+ - + + - + - - - - + + + + - + - + - - - -)^\infty, \\
R_5 &= +- + + + + -(- + + - + + + - + + -)^\infty, \\
R_6 &= +- + - + - - (+ - + - + - - + + + -)^\infty, \\
R_7 &= +- + - + - - \\
&\quad (+ - + - - - - - - + - + - - + - + + - - - + -)^\infty, \\
R_8 &= +- +(- + - - - + - - + + -)^\infty, \\
R_9 &= - + +(+ + - + + + - + - +)^\infty, \\
R_{10} &= - + + + + -(- - + + + - - + - + -)^\infty, \\
R_{11} &= - - - - - +(+ - - + + + - - + - + -)^\infty, \\
R_{12} &= - - -(- - + - - + + + - -)^\infty.
\end{aligned}$$

**Theorem 2.2** *Let  $n \equiv 3 \pmod{4}$ . Then, the initial segment of length  $n$  of each of  $R_1, \dots, R_{12}$  is a strongly balanced binary sequence. Moreover, if  $n \geq 127$ , then every strongly balanced binary sequence of length  $n$  is an initial segment of one of  $R_1, \dots, R_{12}$ , with the following exceptions:*

- If  $n \equiv 3 \pmod{12}$ , there are two more strongly balanced binary sequences of length  $n$ , namely  $R_5[n-4] + -+-$  and  $R_8[n-4] + -++$ .
- If  $n \equiv 7 \pmod{12}$ , there are also two more strongly balanced binary sequences of length  $n$ , namely  $R_8[n-8] + -+++ + +$ , and either  $R_5[n-8] + -+ - - - -$  if  $n \equiv 7 \pmod{24}$ , or  $R_5[n-8] + -+ - - + - +$  if  $n \equiv 19 \pmod{24}$ .

The proof is similar to that of Theorem 2.1. See the last comment in Section 6.

Even though Theorems 2.1 and 2.2 achieve the complete description of all sufficiently long strongly balanced binary sequences, we should point out that there are other infinite families of (simply) balanced binary sequences. For example, for all  $n \equiv 3 \pmod{4}$ , the sequence  $Q_1[n]+$  happens to be balanced. Similarly, for all  $n \equiv 8 \pmod{12}$ , the sequence  $R_1[n]+ - - +$  is balanced as well. And of course, there are the sequences in [2] which originally solved the problem of Steinhaus. None of the presently discussed sequences are strongly balanced, though.

### 3 The method

We shall explain now the method by which we have obtained the results above, and shall also supply our specific *Mathematica* implementation of it.

The idea is quite simple. Assume  $X$  is a strongly balanced binary sequence of length  $n$ . An *extension* of  $X$  is any binary sequence  $Y$  containing  $X$  as an initial segment. Let  $Y$  be any one of the 16 possible extensions of  $X$  of length  $n + 4$ . Then,  *$Y$  is strongly balanced if and only if  $Y$  is balanced.* This holds because  $X$  itself is strongly balanced.

Consequently, if we know the set  $\mathcal{SB}(n)$  of all strongly balanced binary sequences of length  $n$ , and if  $\text{card}(\mathcal{SB}(n)) = t$ , then in order to construct the set  $\mathcal{SB}(n + 4)$ , it is enough to consider the  $16t$  extensions of length  $n + 4$  of all the elements in  $\mathcal{SB}(n)$ , and select those which are simply (hence strongly) balanced. This is a computational task of low complexity.

In summary, our method is a greedy algorithm, which aims to construct all strongly balanced sequences at increasing lengths. For lengths divisible by 4, the algorithm may start with the set  $\{\emptyset\}$  of (strongly) balanced sequences of length 0. In length  $3 \pmod{4}$ , it will start with the set  $\{++-, +-+, -++-, ---\}$  of all (strongly) balanced sequences of length 3.

Here are the very concise *Mathematica* functions which we have written to implement the method. The first four functions (`derive`, `triangle`, `weight` and `ext4`) take as argument an arbitrary finite binary

sequence  $s$ , e.g.  $s = \{1, 1, -1, 1\}$  in *Mathematica* syntax.

1. The function `derive[s]` outputs the derived sequence  $\partial s$  of  $s$ , that is, the sequence of pairwise products of consecutive terms in  $s$ .

```
derive[s_] := Table[s[[i]]s[[i+1]], {i, 1, Length[s] - 1}]
```

2. Then, the function `triangle[s]` outputs the derived triangle  $\Delta s$  of  $s$ , i.e. the list of all higher order derived sequences of  $s$ .

```
triangle[s_] := Block[{s1, tri}, s1 = s; tri = {s1};
  While[Length[s1] > 1, s1 = derive[s1];
    AppendTo[tri, s1]]; tri]
```

3. The function `weight[s]` outputs the sum of the entries in the derived triangle  $\Delta s$  of  $s$ .

```
weight[s_] := Apply[Plus, Flatten[triangle[s]]]
```

4. The function `ext4[s]` outputs the list of all *balanced* binary sequences containing  $s$  as an initial segment and 4 units longer. Note that, if  $s$  is strongly balanced, then `ext4[s]` outputs the list of all *strongly balanced* sequences containing  $s$  as an initial segment and 4 units longer.

```
ext4[s_] := Block[{l, sext}, l = {};
  Do[sext = Join[s, {x1, x2, x3, x4}];
    If[weight[sext] == 0, AppendTo[l, sext]],
    {x1, -1, 1, 2}, {x2, -1, 1, 2},
    {x3, -1, 1, 2}, {x4, -1, 1, 2}];
  l]
```

5. Finally, given a non-negative integer  $n \equiv 0$  or  $3 \pmod{4}$ , the function `strong[n]` successively builds all strongly balanced binary sequences of length  $m$  with  $m \leq n$  and  $m \equiv n \pmod{4}$ .

```
strong[n_] := strong[n] = (If[n == 0, Return[{}]]];
  If[n == 3,
    Return[{{1, 1, -1}, {1, -1, 1},
      {-1, 1, 1}, {-1, -1, -1}}]];
  Flatten[Map[ext4, strong[n - 4]], 1]]
```



For instance, the command `Sum[Length[strong[n]]*t^n, {n, 0, 88, 4}]` will output the polynomial  $f_0(t)$  of Theorem 1.3, where  $f_0(t) = \sum_{i=0}^{22} sb(4i)t^{4i}$  displays the numbers  $sb(n)$  for each length  $n = 4i \leq 88$ . This computation takes about 90 seconds on a standard PC with a Pentium 4m processor clocked at 1.6 GHZ.

## 4 Other possible strengthenings

We describe here two other attempts of strengthening the notion of balanced sequences. However, in contrast to strongly balanced sequences, these other strengthenings turn out to admit only *finitely many* complying binary sequences.

### 4.1 M-sequences

In our first attempt, we shall be seeking binary sequences  $X = (x_1, \dots, x_n)$  having the property **M** defined recursively as follows:  $X$  is balanced, and its middle segment  $(x_3, \dots, x_{n-2})$  of length  $n - 4$  is also balanced and satisfies property **M**. By convention, balanced binary sequences of length 0 or 3 satisfy property **M**. (Compare with the similar-looking recursive definition of strongly balanced sequences.) For brevity, sequences satisfying property **M** will be called **M**-sequences.

We shall restrict our attention to lengths  $n \equiv 0 \pmod{4}$ . As it turns out, there are binary **M**-sequences of length  $n$  for every  $n = 4, 8, \dots, 96$ . In length 96, there remain exactly two binary **M**-sequences. Quite surprisingly, none of these two sequences can be extended to a sequence of length 100 still satisfying property **M**. Consequently, *there are no binary M-sequences  $X$  of length  $n \equiv 0 \pmod{4}$  with  $n \geq 100$* . Thus, the generating function  $g_M(t) = \sum_X t^{l(X)}$ , where  $X$  runs over the set of all balanced binary **M**-sequences of even length, and where  $l(X)$  denotes the length of  $X$ , is a *polynomial* of degree 96, given by the following expression:

$$\begin{aligned} g_M(t) = & 2t^{96} + 8t^{92} + 10t^{88} + 14t^{84} + 22t^{80} + 22t^{76} + 30t^{72} \\ & + 48t^{68} + 76t^{64} + 88t^{60} + 108t^{56} + 130t^{52} + 174t^{48} \\ & + 226t^{44} + 222t^{40} + 198t^{36} + 172t^{32} + 144t^{28} \\ & + 138t^{24} + 94t^{20} + 60t^{16} + 40t^{12} + 20t^8 + 6t^4 + 1. \end{aligned}$$

For definiteness, here are the two binary **M**-sequences of length 96:

```

++++-----++++-----+-++-+-+--+--+--+
++-----+-----+---+---+---+---+---+---+---+---
+-+++++-----+---+---+---+---+---+---+---+---+
+++++---+---+---+---+---+---+---+---+---+---+---+
-++-+++++---+---+---+---+---+---+---+---+---+---+
--+-++-+-+---+---+---+---+---+---+---+---+---+

```

## 4.2 Universal balanced binary sequences

In our second attempt, we seek *universal* balanced binary sequences, *i.e.* balanced binary sequences  $X = (x_1, \dots, x_n)$  with the property that every initial segment  $(x_1, \dots, x_k)$ , with  $k \equiv 0$  or  $3 \pmod{4}$ , is also balanced. There are exactly 6 universal balanced binary sequences of length 11, namely  $+ - + + + - + - + +$ ,  $+ - + + + - - + + -$ ,  $+ - + + + - - + - +$ ,  $+ - + - + - - + - +$ ,  $+ - + - + - - + + +$  and  $+ - + - + - - + - -$ . As easily checked, by adding one more  $\pm$  sign at the end of each of these 6 sequences, we find that *there are no universal balanced binary sequences in length 12 or higher*.

## 5 Related open problems

We propose here a few open problems in the same spirit as that of Steinhaus.

**Problem 1** Are there infinitely many *symmetric* balanced binary sequences, such as  $X = ++-+-++$ ? More generally, what is the set of lengths of all such sequences? For instance, it may be shown that there exists no symmetric balanced binary sequences of length  $n \equiv 4 \pmod{8}$ .

**Problem 2** The balanced sequences  $X$  of length 12 and 20 given in [3] have the property that  $S(X) = 0$ , where  $S(X)$  is the sum of the entries in  $X$ . As a consequence, their derived sequences, of length 11 and 19, respectively, are also balanced. It would be of great interest to know, more generally, whether for every  $n$  divisible by 4, there exists a balanced binary sequence  $X$  of length  $n$  satisfying  $S(X) = 0$ . We did

find such sequences in every length  $n = 4k$  with  $n \leq 36$ . However, we do not know whether they exist in higher length. This problem was suggested by Michel Kervaire during a phone conversation with one of the authors.

**Problem 3** For every binary sequence  $X$  of length  $n \equiv 1$  or  $2 \pmod{4}$ , the sum  $S(\Delta X)$  of the entries of the derived triangle  $\Delta X$  of  $X$  is an odd number. It is natural to ask whether the value  $S(\Delta X) = 1$  (respectively  $S(\Delta X) = -1$ ) is attained for every  $n \equiv 1$  or  $2 \pmod{4}$ . More generally, given any integer  $v$ , are there infinitely many finite binary sequences  $X$  such that  $S(\Delta X) = v$ ? We know at least that the answer is positive for  $v = -3, -2, 1, 2, 4$  and  $5$ , by taking suitable initial segments of some of the  $Q_i$  and the  $R_i$ . The answer is also positive for  $v = -1$ , with the sequence  $Q_1[n] + -$  for every  $n \equiv 11 \pmod{12}$ . Still more generally, what can be said about the generating function  $G_n(t) = \sum_X t^{S(\Delta X)}$ , where  $X$  runs over the set of all binary sequences of length  $n$ ?

**Problem 4** The notion of balanced sequence makes sense not only with entries  $\pm 1$ , but more generally with entries taken in any (commutative) ring  $R$ . Indeed, let  $X = (x_1, \dots, x_n)$  be a sequence with entries  $x_i \in R$  for all  $i$ . The *derived sequence*  $\partial X = (y_1, \dots, y_{n-1})$  of  $X$  can still be defined by  $y_i = x_i x_{i+1}$  for all  $1 \leq i \leq n-1$ , and this gives rise again to the *derived triangle*  $\Delta X$  of  $X$ , namely the collection of the  $\partial^k X$ . Of course, the sequence  $X$  is said to be *balanced* if the sum of the entries in  $\Delta X$  is  $0 \in R$ . Are there interesting infinite families of balanced sequences in this more general setting?

For instance, let  $p$  be a prime number, let  $\zeta$  be a primitive  $p$ th root of unity, and let  $R = \mathbb{Z}[\zeta]$ . In a forthcoming note, we shall show that, for  $p = 3$ , the ring  $R$  contains infinitely many balanced sequences of powers of  $\zeta$ . We do not know whether this remains true for larger primes  $p$ .

The referee has suggested the following related problem. Let  $G$  be a finite group, even a non-abelian one. Are there infinitely many sequences  $X$  with entries in  $G$  whose derived triangle  $\Delta X$  contains the same number of occurrences of each group element?

**Problem 5** This is really a family of problems. We may consider higher-dimensional analogues of balanced sequences, such as balanced binary matrices, balanced binary 3-dimensional tensors, or balanced binary simplices for example. In general, the concept of a balanced object  $X$

will make sense whenever there is a suitable notion of derived object  $X \mapsto \partial X$ , with strictly decreasing sizes. The derived object should be constructed by taking the product of the neighbours for each suitable position in  $X$ , as is the case for sequences. A given object  $X$  will then be said to be balanced whenever the sum of the entries in the collection of its iterated derived objects  $\partial^k X$  is zero.

Consider, for example, the following notion of a balanced binary square matrix. If  $A = (a_{i,j})_{1 \leq i,j \leq n}$  is a binary matrix of order  $n$ , define  $\partial A$  as the binary matrix  $(b_{i,j})_{1 \leq i,j \leq n-1}$  of order  $n-1$ , where  $b_{i,j} = a_{i,j}a_{i,j+1}a_{i+1,j}a_{i+1,j+1}$ . The derived pyramid  $\Delta A$  is then defined as the collection of  $\partial^k A$  for  $0 \leq k \leq n-1$ . Note that again, the total number of binary entries in  $\Delta A$  is even if and only if  $n$  is congruent to 0 or 3 mod 4. Are there infinitely many balanced binary matrices?

**Problem 6** Let  $X$  be an arbitrary binary sequence of length  $n$ . Does there exist a *balanced* binary sequence  $Y$  having  $X$  as an initial segment? (This problem is due to Pierre Duchet.)

For instance, let  $J_n$  be the constant +1 sequence of length  $n$ . What is the length  $j(n)$  of a shortest possible balanced binary extension of  $J_n$ , if one exists at all? We know by construction that  $j(100) \leq 236$ .

## 6 Highlights of the proof of Theorem 2.1

We shall give here parts of the proof of Theorem 2.1. There are two things to prove. First, that the initial segments  $Q_i[n]$  are balanced, for every  $n \equiv 0 \pmod{4}$ . And second, that there are no other strongly balanced binary sequences of length  $n \equiv 0 \pmod{4}$ , provided  $n \geq 92$ .

We shall restrict our attention to  $Q_1$ . (The phenomena are similar for  $Q_2, Q_3, Q_4$ .) The fact that  $S(\Delta Q_1[n]) = 0$  for  $n \equiv 0 \pmod{4}$  will follow from a certain periodic structure of the derived triangle  $\Delta Q_1[n]$ . This structure then allows us to control which extensions of  $Q_1[n]$  remain strongly balanced, leading to the classification statement.

This is already quite tedious. Consequently, we shall not discuss Theorem 2.2 concerning sequences of length  $n \equiv 3 \pmod{4}$ . However, the phenomena are similar again, and it should become clear that a complete proof can be written in this case as well.

Figure 1: *Structure of the derived triangle of  $Q_1$ [52].*

Recall from subsection 2.1 that  $Q_1 = + - + + (+ + - + + - + - - - ++)^{\infty}$ . Let  $n \equiv 0 \pmod{4}$  be a given positive integer. We claim that  $\Delta Q_1[n]$  has a periodic structure, as illustrated in Figure 1. More specifically, we will prove that, if  $n \geq 16$ , there are nine types of NE/SW diagonal strips of width 4, denoted  $A_1, A_2, A_3, B_1, B_2, B_3, C_1, C_2, C_3$ , such that the derived triangle  $\Delta Q_1[n]$  is the periodic assembly of  $T_{16} = \Delta Q_1[16]$  and of the components  $A_i, B_i, C_i$ , as depicted in Figure 1. Note that the components  $A_1, B_1, C_1$  appear on the top of the derived triangle, the components  $A_3, B_3, C_3$  on its SW side, and  $A_2, B_2, C_2$  occupy the rest of the triangle (except  $T_{16}$ ). The sum of each component is as indicated (*e.g.*  $A_1$  has sum  $S(A_1) = 0$ ,  $B_1$  has sum  $S(B_1) = -4$ , and so on).

According to this structure of  $\Delta Q_1[n]$ , we see that each full NE/SW diagonal strip of width 4 on the right of  $T_{16}$  has sum zero, and therefore

$S(\Delta Q_1[n]) = 0$ , as claimed.

In order to establish this structure, we need to introduce a few notations selecting certain specific parts of these NE/SW diagonal strips.

**Notation.**

- $x_p^q$  denotes the  $p$ th digit in the  $q$ th row of  $\Delta Q_1[n]$ , for all  $1 \leq p \leq n$  and  $1 \leq q \leq n - p + 1$ . In particular, the first row of  $\Delta Q_1[n]$ , *i.e.*  $Q_1[n]$  itself, is constituted by the elements  $x_1^1, x_2^1, \dots, x_n^1$ , and the left side of the triangle  $\Delta Q_1[n]$  consists of  $x_1^1, x_1^2, \dots, x_1^n$ . The basic defining property of the triangle  $\Delta Q_1[n]$  thus reads  $x_p^{q+1} = x_p^q x_{p+1}^q$ .
- $d_i$  denotes the  $i$ th NE/SW diagonal of  $\Delta Q_1[n]$ , *i.e.*  $d_i$  is the right side of the triangle  $\Delta Q_1[i]$ , for all  $1 \leq i \leq n$ ;
- For  $i \equiv 1 \pmod{4}$  and  $j \equiv 1 \pmod{4}, 1 \leq j \leq i$ ,  $T_i^j$  denotes the trapezoid of figure 3.
- For  $i \equiv 1 \pmod{4}$ , we set  $S_i = T_i^i$ . This special trapezoid  $S_i$  corresponds to the last four NE/SW diagonals of  $\Delta Q_1[i+3]$ , and will be called a *strip*.
- For  $i \equiv 1 \pmod{4}$  and  $j \equiv 2 \pmod{4}, 2 \leq j \leq i$ ,  $P_i^j$  denotes the parallelogram of figure 2, of width 4 and length 12.

A few remarks are in order. First observe that, because of the basic property  $x_p^{q+1} = x_p^q x_{p+1}^q$ , the trapezoid  $T_i^j$  is completely determined by its top row and its left side, namely by  $x_i^1, x_{i+1}^1, x_{i+2}^1, x_{i+3}^1$  and  $x_i^1, x_{i-1}^2, \dots, x_{i+1-j}^j$ . Now, this left side of  $T_i^j$  is itself determined by  $x_i^1$  and by the right side of the adjacent trapezoid  $T_{i-4}^{j-4}$ . We record these observations as follows.

**Fact 1** The trapezoid  $T_i^j$  is completely determined by its top row and by the right side of  $T_{i-4}^{j-4}$ .

Similar remarks can be made about the parallelogram  $P_i^j$ , and we have:

**Fact 2** The parallelogram  $P_i^j$  is completely determined by the bottom of the quadrilateral just above it and by the right side of  $P_{i-4}^{j-4}$ .

Finally, given  $i \equiv 1 \pmod{4}$ , let  $j$  be the unique element in the set  $\{1, 5, 9\}$  which is congruent to  $i \pmod{12}$ . Clearly, with these notations,

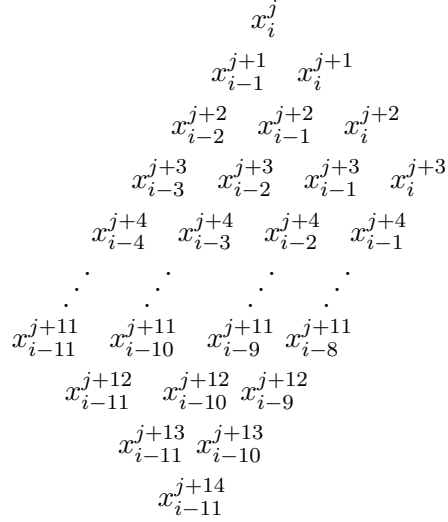


Figure 2: Parallelogram  $P_i^j$ .

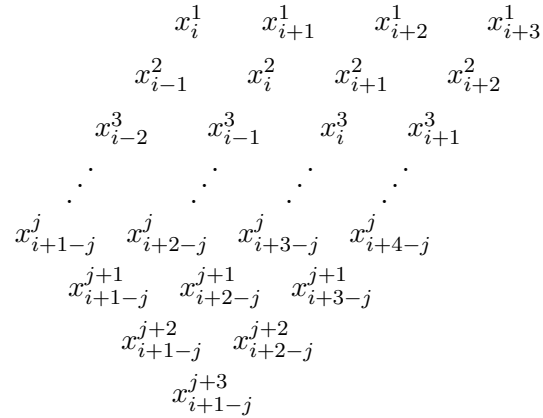


Figure 3: Trapezoid  $T_i^j$ .

the strip  $S_i$  is the concatenation, in the NE/SW direction, of the trapezoid  $T_i^j$  and of the  $(i-j)/12$  parallelograms  $P_i^{j+1}, P_i^{j+3}, \dots, P_i^{i-11}$ .

We will denote the NE/SW concatenation by the symbol  $+$ . With this notation, we have  $S_i = T_i^j + P_i^{j+1} + P_i^{j+3} + \dots + P_i^{i-11}$ .

We now define the 9 special components  $A_i, B_i, C_i$ , where  $A_1, B_1, C_1$  are trapezoids, whereas  $A_2, B_2, C_2, A_3, B_3, C_3$  are parallelograms:

$$\begin{array}{l}
 * A_1 := T_{17}^5 = \begin{array}{cccc} + & + & - & + \\ & + & + & - \\ + & + & - & + \\ - & + & - & - \\ - & - & - & + \\ + & + & - & \\ + & - & & \\ - & & & \end{array} & * A_2 := P_{29}^6 = \begin{array}{cccc} + & & & \\ & - & + & \\ & + & - & + \\ & + & - & - \\ + & - & + & + \\ - & - & - & + \\ + & + & + & - \\ - & + & + & - \\ + & - & + & - \\ + & - & - & - \\ - & - & - & + \\ + & + & - & \\ + & - & & \\ - & & & \end{array} \\
 \\
 * A_3 := P_{17}^6 = \begin{array}{cccc} + & & & \\ & - & + & \\ & + & - & + \\ + & - & - & - \\ + & - & + & + \\ - & - & - & + \\ + & + & + & - \\ - & + & + & - \\ + & - & + & - \\ + & - & - & - \\ + & - & + & + \\ + & - & - & + \\ - & + & - & \\ - & - & & \\ + & & & \end{array} & * B_1 := T_{21}^9 = \begin{array}{cccc} + & - & + & - \\ & + & - & - \\ & - & - & + \\ - & + & - & + \\ + & - & - & - \\ + & - & + & + \\ - & - & - & + \\ + & + & + & - \\ - & + & + & - \\ - & + & - & \\ - & - & & \\ + & & & \end{array}
 \end{array}$$





$$\begin{array}{cccc}
 * C_3 := P_{25}^{14} = & + & & \\
 & & - & - \\
 & & & + & + & + \\
 & & & - & + & + & - \\
 & & & & + & - & + & - \\
 & & & & & + & - & - & - \\
 & & & & & & + & - & + & + \\
 & & & & & & & - & - & - & + \\
 & & & & & & & & + & + & + & - \\
 & & & & & & & & - & + & + & - \\
 & & & & & & & & + & - & + & - \\
 & & & & & & & & - & - & - & - \\
 & & & & & & & & + & + & + \\
 & & & & & & & & & + & + \\
 & & & & & & & & & & +
 \end{array}$$

We shall need to observe some resemblances between some of these components, to be used with Facts 1 and 2.

- The SW edge of  $A_1$  (respectively  $B_1, C_1$ ) is equal to the SW edge of  $A_2$  (respectively  $B_2, C_2$ ).
- The 12-tuple composed by the last 12 digits of the right side of  $C_1$  is equal to the 12-tuple containing the digits of the right side of  $C_2$ .

We claim that the strips  $S_i$  come in 3 different types, depending on the class  $i \equiv 1, 5$  or  $9 \pmod{12}$ . Here is the general key formula we want to prove:

**Claim 1**

$$\begin{aligned}
 \forall k \in \mathbb{N}, k \geq 1, \quad S_{12k+5} &= A_1 + (k-1)A_2 + A_3, \\
 S_{12k+9} &= B_1 + (k-1)B_2 + B_3, \\
 S_{12(k+1)+1} &= C_1 + (k-1)C_2 + C_3.
 \end{aligned}$$

As we will see, this results from the structure of the 9 components  $A_i, B_i, C_i$  and Facts 1 and 2, and may be proved by induction on  $k$ .

To start the induction, one verifies the claim in  $\Delta Q_1[40]$  by direct observation.

Assume now that the claim is true for  $k = 1, 2$ . In particular, we know that  $S_{37} = C_1 + C_2 + C_3$ . We will show that  $S_{41} = A_1 + 2A_2 + A_3$ . By periodicity of the sequence  $Q_1$ , we know that the top of  $S_{41}$  is equal to the top of  $A_1$ . Thus, using Fact 1, we derive that the trapezoid  $T_{41}^{17}$  is equal to  $A_1 + A_2$ . Indeed, it is completely determined by the top of  $A_1$  and the right side of  $C_1$  (by the hypothesis for  $S_{37}$ ), and the same is true for  $A_1 + A_2$  in  $S_{29}$ , by the hypothesis for  $S_{29}$ . Thus, the parallelogram just under  $A_1 + A_2$  in  $S_{41}$  is completely determined by the bottom of  $A_2$  and the right side of  $C_2$ , which is equal to the last 12 digits of the right side of  $C_1$ . According to the verifications we have just made for the previous trapezoid, the same holds for  $A_2$ , whence Fact 2 implies:  $T_{41}^{29} = A_1 + A_2 + A_2$ .

Finally, similar arguments enable us to show that the last parallelogram of  $S_{41}$  is equal to the last parallelogram of  $S_{29}$ , namely  $A_3$ . Hence we get  $S_{41} = A_1 + 2A_2 + A_3$ , and we are done.

The case  $k \geq 3$  can be treated in the same way, by induction.

**Claim 2**  $\forall n \equiv 1 \pmod{4}, S(S_n) = 0$ .

Using Claim 1, it suffices to compute the sum of each of the 9 components  $A_i, B_i, C_i$  and of the first irregular strips.

For every  $n \leq 37$ , we check the equality by direct computations of sums in the triangle  $\Delta Q_1[40]$ .

For  $n \geq 41$ , we have to consider three possibilities, according to Claim 1:

- If  $n = 12k + 1, k \geq 2$ , then

$$\begin{aligned} S(S_n) &= S(C_1) + (k - 2)S(C_2) + S(C_3) \\ &= -4 + (k - 2) \times 0 + 4 \\ &= 0 ; \end{aligned}$$

- if  $n = 12k + 5, k \geq 2$ , then

$$\begin{aligned} S(S_n) &= S(A_1) + (k - 1)S(A_2) + S(A_3) \\ &= 0 + (k - 1) \times 0 + 0 \\ &= 0 ; \end{aligned}$$

- if  $n = 12k + 9, k \geq 2$ , then

$$S(S_n) = S(B_1) + (k - 1)S(B_2) + S(B_3)$$

$$\begin{aligned}
&= -4 + (k - 1) \times 0 + 4 \\
&= 0.
\end{aligned}$$

This proves Claim 2. It follows that  $S(\Delta Q_1[n]) = 0$ , *i.e.* that  $Q_1[n]$  is balanced, for every  $n \equiv 0 \pmod{4}$ .

We now turn to the proof of the second part of Theorem 2.1, namely that every strongly balanced binary sequence of length  $n$  with  $n = 4m \geq 92$  is equal to  $Q_i[n]$  for some  $1 \leq i \leq 4$ .

We do this by induction on  $n$ , starting at  $n = 92$ . In order to construct *all* strongly balanced binary sequences of length 92, we use the method of Section 3, implemented in the given *Mathematica* functions. For example, issuing the command `strong[92]` to *Mathematica* will output exactly four sequences, namely  $Q_1[92]$ ,  $Q_2[92]$ ,  $Q_3[92]$  and  $Q_4[92]$ . This computation uses exact integer arithmetic only. This establishes the case  $n = 92$ .

Let  $n \geq 92$  with  $n \equiv 0 \pmod{4}$ . It remains to show that, if  $X = Q_i[n]$  for some  $i \in \{1, 2, 3, 4\}$ , then there is a unique extension  $X'$  of  $X$ , of length  $n + 4$ , such that  $X'$  is (simply, hence strongly) balanced, and  $X' = Q_i[n + 4]$ . (In fact, this statement already holds true for  $n \geq 52$  if  $i = 1$  or 3, and for  $n \geq 64$  if  $i = 2$  or 4.)

Once again, we restrict our attention to  $Q_1$ , so  $X = Q_1[n]$ . We denote an arbitrary extension of length  $n + 4$  of  $X$  as the concatenation  $Y = Y(x_1, x_2, x_3, x_4) = Xx_1x_2x_3x_4$ , where  $x_1, x_2, x_3, x_4$  are unknown binary digits satisfying  $x_i^2 = 1$ . Our task is to determine those values of  $x_i \in \{\pm 1\}$  for which  $S(\Delta Y) = 0$ .

In order to do this, we need to determine the structure of the derived triangle  $\Delta Y(x_1, x_2, x_3, x_4)$  in terms of the unknown  $x_1, x_2, x_3, x_4$ .

**Claim 3** For every  $n \in \mathbb{N}$ ,  $n \equiv 0 \pmod{4}$ , the last strip  $S_{n+1}(x_1, x_2, x_3, x_4)$  of the triangle  $\Delta(Q_1[n]x_1x_2x_3x_4)$  has the following structure:

$$S_{n+1}(x_1, x_2, x_3, x_4) = \begin{cases} C'_1 + (k - 2)C'_2 + C'_3 & \text{if } n = 12k \\ A'_1 + (k - 1)A'_2 + A'_3 & \text{if } n = 12k + 4 \\ B'_1 + (k - 1)B'_2 + B'_3 & \text{if } n = 12k + 8, \end{cases}$$

where  $A'_1, B'_1, C'_1$  are trapezoids and  $A'_2, B'_2, C'_2, A'_3, B'_3, C'_3$  parallelograms. These components have the same size as the corresponding components  $A_i, B_i, C_i$ , and are depicted below.

Not surprisingly,  $A'_i, B'_i, C'_i$  share similar properties as  $A_i, B_i, C_i$ , *i.e.* the bottom of  $A'_1$  (respectively  $B'_1, C'_1$ ) is equal to the bottom of  $A'_2$  (respectively  $B'_2, C'_2$ ), and the 12-tuple composed by the last 12 digits of the right side of  $C'_1$  is equal to the 12-tuple containing the digits of the right side of  $C'_2$ . Thus, the proof of Claim 3 is similar to that of Claim 1. Here are  $A'_i, B'_i, C'_i$ , explicitly.

$$\begin{aligned}
 * A'_1 = & \begin{array}{cccc}
 x_1 & x_2 & x_3 & x_4 \\
 x_1 & x_1x_2 & x_2x_3 & x_3x_4 \\
 x_1 & x_2 & x_1x_3 & x_2x_4 \\
 -x_1 & x_1x_2 & x_1x_2x_3 & x_1x_2x_3x_4 \\
 -x_1 & -x_2 & x_3 & x_4 \\
 x_1 & x_1x_2 & -x_2x_3 & x_3x_4 \\
 -x_1 & x_2 & -x_1x_3 & -x_2x_4 \\
 x_1 & -x_1x_2 & -x_1x_2x_3 & x_1x_2x_3x_4 \\
 x_1 & -x_2 & x_3 & -x_4 \\
 x_1 & -x_1x_2 & -x_2x_3 & -x_3x_4 \\
 -x_1 & -x_2 & x_1x_3 & x_2x_4 \\
 x_1 & x_1x_2 & -x_1x_2x_3 & x_1x_2x_3x_4 \\
 -x_1 & x_2 & -x_3 & -x_4 \\
 x_1 & -x_1x_2 & -x_2x_3 & x_3x_4 \\
 x_1 & -x_2 & x_1x_3 & -x_2x_4 \\
 x_1 & -x_1x_2 & -x_1x_2x_3 & -x_1x_2x_3x_4 \\
 -x_1 & -x_2 & x_3 & x_4 \\
 x_1x_2 & -x_2x_3 & x_3x_4 & \\
 -x_1x_3 & -x_2x_4 & & \\
 x_1x_2x_3x_4 & & & 
 \end{array}
 \end{aligned}$$

$$\begin{aligned}
* A'_2 &= x_1 \\
&\quad -x_1 \quad x_2 \\
&\quad x_1 -x_1x_2 -x_1x_2x_3 \\
&\quad x_1 \quad -x_2 \quad x_3 \quad -x_4 \\
&\quad x_1 -x_1x_2 -x_2x_3 -x_3x_4 \\
&\quad -x_1 \quad -x_2 \quad x_1x_3 \quad x_2x_4 \\
&\quad x_1 \quad x_1x_2 -x_1x_2x_3 \quad x_1x_2x_3x_4 \\
&\quad -x_1 \quad x_2 \quad -x_3 \quad -x_4 \\
&\quad x_1 -x_1x_2 -x_2x_3 \quad x_3x_4 \\
&\quad x_1 \quad -x_2 \quad x_1x_3 -x_2x_4 \\
&\quad x_1 -x_1x_2 -x_1x_2x_3 -x_1x_2x_3x_4 \\
-x_1 &\quad -x_2 \quad x_3 \quad x_4 \\
&\quad x_1x_2 -x_2x_3 \quad x_3x_4 \\
&\quad -x_1x_3 \quad -x_2x_4 \\
&\quad x_1x_2x_3x_4
\end{aligned}$$

$$\begin{aligned}
* A'_3 &= x_1 \\
&\quad -x_1 \quad x_2 \\
&\quad x_1 -x_1x_2 -x_1x_2x_3 \\
&\quad x_1 \quad -x_2 \quad x_3 \quad -x_4 \\
&\quad x_1 -x_1x_2 -x_2x_3 -x_3x_4 \\
&\quad -x_1 \quad -x_2 \quad x_1x_3 \quad x_2x_4 \\
&\quad x_1 \quad x_1x_2 -x_1x_2x_3 \quad x_1x_2x_3x_4 \\
&\quad -x_1 \quad x_2 \quad -x_3 \quad -x_4 \\
&\quad x_1 -x_1x_2 -x_2x_3 \quad x_3x_4 \\
&\quad x_1 \quad -x_2 \quad x_1x_3 -x_2x_4 \\
&\quad x_1 -x_1x_2 -x_1x_2x_3 -x_1x_2x_3x_4 \\
x_1 &\quad -x_2 \quad x_3 \quad x_4 \\
&\quad -x_1x_2 -x_2x_3 \quad x_3x_4 \\
&\quad x_1x_3 \quad -x_2x_4 \\
&\quad -x_1x_2x_3x_4
\end{aligned}$$

$$\begin{aligned}
* B'_1 = & x_1 \quad x_2 \quad x_3 \quad x_4 \\
& x_1 \quad x_1x_2 \quad x_2x_3 \quad x_3x_4 \\
& -x_1 \quad x_2 \quad x_1x_3 \quad x_2x_4 \\
& -x_1 \quad -x_1x_2 \quad x_1x_2x_3 \quad x_1x_2x_3x_4 \\
& x_1 \quad x_2 \quad -x_3 \quad x_4 \\
& x_1 \quad x_1x_2 \quad -x_2x_3 \quad -x_3x_4 \\
& -x_1 \quad x_2 \quad -x_1x_3 \quad x_2x_4 \\
& x_1 \quad -x_1x_2 \quad -x_1x_2x_3 \quad -x_1x_2x_3x_4 \\
-x_1 \quad & -x_2 \quad x_3 \quad x_4 \\
& x_1x_2 \quad -x_2x_3 \quad x_3x_4 \\
& -x_1x_3 \quad -x_2x_4 \\
& x_1x_2x_3x_4
\end{aligned}$$

$$\begin{aligned}
* B'_2 = & x_1 \\
& x_1 \quad x_2 \\
& x_1 \quad x_1x_2 \quad -x_1x_2x_3 \\
& -x_1 \quad x_2 \quad -x_3 \quad -x_4 \\
& x_1 \quad -x_1x_2 \quad -x_2x_3 \quad x_3x_4 \\
& -x_1 \quad -x_2 \quad x_1x_3 \quad -x_2x_4 \\
& x_1 \quad x_1x_2 \quad -x_1x_2x_3 \quad -x_1x_2x_3x_4 \\
& x_1 \quad x_2 \quad -x_3 \quad x_4 \\
& x_1 \quad x_1x_2 \quad -x_2x_3 \quad -x_3x_4 \\
& -x_1 \quad x_2 \quad -x_1x_3 \quad x_2x_4 \\
& x_1 \quad -x_1x_2 \quad -x_1x_2x_3 \quad -x_1x_2x_3x_4 \\
-x_1 \quad & -x_2 \quad x_3 \quad x_4 \\
& x_1x_2 \quad -x_2x_3 \quad x_3x_4 \\
& -x_1x_3 \quad -x_2x_4 \\
& x_1x_2x_3x_4
\end{aligned}$$

$$\begin{aligned}
* B'_3 &= x_1 \\
&\quad x_1 \quad x_2 \\
&\quad\quad x_1 \quad x_1x_2 \quad -x_1x_2x_3 \\
&\quad\quad -x_1 \quad x_2 \quad -x_3 \quad -x_4 \\
&\quad\quad\quad x_1 \quad -x_1x_2 \quad -x_2x_3 \quad x_3x_4 \\
&\quad\quad\quad -x_1 \quad -x_2 \quad x_1x_3 \quad -x_2x_4 \\
&\quad\quad\quad\quad x_1 \quad x_1x_2 \quad -x_1x_2x_3 \quad -x_1x_2x_3x_4 \\
&\quad\quad\quad\quad x_1 \quad x_2 \quad -x_3 \quad x_4 \\
&\quad\quad\quad\quad\quad x_1 \quad x_1x_2 \quad -x_2x_3 \quad -x_3x_4 \\
&\quad\quad\quad\quad\quad -x_1 \quad x_2 \quad -x_1x_3 \quad x_2x_4 \\
&\quad\quad\quad\quad\quad\quad x_1 \quad -x_1x_2 \quad -x_1x_2x_3 \quad -x_1x_2x_3x_4 \\
&\quad\quad\quad\quad\quad\quad x_1 \quad -x_2 \quad x_3 \quad x_4 \\
&\quad\quad\quad\quad\quad\quad -x_1x_2 \quad -x_2x_3 \quad x_3x_4 \\
&\quad\quad\quad\quad\quad\quad\quad x_1x_3 \quad -x_2x_4 \\
&\quad\quad\quad\quad\quad\quad\quad -x_1x_2x_3x_4
\end{aligned}$$

$$\begin{aligned}
* C'_1 &= x_1 \quad x_2 \quad x_3 \quad x_4 \\
&\quad -x_1 \quad x_1x_2 \quad x_2x_3 \quad x_3x_4 \\
&\quad\quad x_1 \quad -x_2 \quad x_1x_3 \quad x_2x_4 \\
&\quad\quad\quad x_1 \quad -x_1x_2 \quad -x_1x_2x_3 \quad x_1x_2x_3x_4 \\
&\quad\quad\quad\quad x_1 \quad -x_2 \quad x_3 \quad -x_4 \\
&\quad\quad\quad\quad -x_1 \quad -x_1x_2 \quad -x_2x_3 \quad -x_3x_4 \\
&\quad\quad\quad\quad -x_1 \quad x_2 \quad x_1x_3 \quad x_2x_4 \\
&\quad\quad\quad\quad -x_1 \quad -x_1x_2 \quad x_1x_2x_3 \quad x_1x_2x_3x_4 \\
&\quad\quad\quad\quad\quad x_1 \quad x_2 \quad -x_3 \quad x_4 \\
&\quad\quad\quad\quad\quad -x_1 \quad x_1x_2 \quad -x_2x_3 \quad -x_3x_4 \\
&\quad\quad\quad\quad\quad\quad x_1 \quad -x_2 \quad -x_1x_3 \quad x_2x_4 \\
&\quad\quad\quad\quad\quad\quad -x_1 \quad -x_1x_2 \quad x_1x_2x_3 \quad -x_1x_2x_3x_4 \\
&\quad\quad\quad\quad\quad\quad -x_1 \quad x_2 \quad -x_3 \quad -x_4 \\
&\quad\quad\quad\quad\quad\quad -x_1x_2 \quad -x_2x_3 \quad x_3x_4 \\
&\quad\quad\quad\quad\quad\quad\quad x_1x_3 \quad -x_2x_4 \\
&\quad\quad\quad\quad\quad\quad\quad -x_1x_2x_3x_4
\end{aligned}$$



$$\begin{aligned}
* C'_2 = & -x_1 \\
& x_1 \quad x_2 \\
& -x_1 \quad x_1x_2 \quad x_1x_2x_3 \\
& x_1 \quad -x_2 \quad x_3 \quad -x_4 \\
& -x_1 \quad -x_1x_2 \quad -x_2x_3 \quad -x_3x_4 \\
& -x_1 \quad x_2 \quad x_1x_3 \quad x_2x_4 \\
& -x_1 \quad -x_1x_2 \quad x_1x_2x_3 \quad x_1x_2x_3x_4 \\
& x_1 \quad x_2 \quad -x_3 \quad x_4 \\
& -x_1 \quad x_1x_2 \quad -x_2x_3 \quad -x_3x_4 \\
& x_1 \quad -x_2 \quad -x_1x_3 \quad x_2x_4 \\
& -x_1 \quad -x_1x_2 \quad x_1x_2x_3 \quad -x_1x_2x_3x_4 \\
-x_1 & \quad x_2 \quad -x_3 \quad -x_4 \\
& -x_1x_2 \quad -x_2x_3 \quad x_3x_4 \\
& x_1x_3 \quad -x_2x_4 \\
& -x_1x_2x_3x_4
\end{aligned}$$

$$\begin{aligned}
* C'_3 = & -x_1 \\
& x_1 \quad x_2 \\
& -x_1 \quad x_1x_2 \quad x_1x_2x_3 \\
& x_1 \quad -x_2 \quad x_3 \quad -x_4 \\
& -x_1 \quad -x_1x_2 \quad -x_2x_3 \quad -x_3x_4 \\
& -x_1 \quad x_2 \quad x_1x_3 \quad x_2x_4 \\
& -x_1 \quad -x_1x_2 \quad x_1x_2x_3 \quad x_1x_2x_3x_4 \\
& x_1 \quad x_2 \quad -x_3 \quad x_4 \\
& -x_1 \quad x_1x_2 \quad -x_2x_3 \quad -x_3x_4 \\
& x_1 \quad -x_2 \quad -x_1x_3 \quad x_2x_4 \\
& -x_1 \quad -x_1x_2 \quad x_1x_2x_3 \quad -x_1x_2x_3x_4 \\
x_1 & \quad x_2 \quad -x_3 \quad -x_4 \\
& x_1x_2 \quad -x_2x_3 \quad x_3x_4 \\
& -x_1x_3 \quad -x_2x_4 \\
& x_1x_2x_3x_4
\end{aligned}$$

We are now in a position to determine the sum  $S(\Delta Q_1[n]x_1x_2x_3x_4)$  in terms of the  $x_i$ . Since we already know that  $S(\Delta Q_1[n]) = 0$ , it follows

that  $S(\Delta Q_1[n]x_1x_2x_3x_4) = S(S_{n+1}(x_1, x_2, x_3, x_4))$ . From this remark and Claim 3, we have, for all  $n \geq 36$ :

$$S(\Delta Q_1[n]x_1x_2x_3x_4) = \begin{cases} S(C'_1) + (k-2)S(C'_2) + S(C'_3) & \text{if } n = 12k \\ S(A'_1) + (k-1)S(A'_2) + S(A'_3) & \text{if } n = 12k + 4 \\ S(B'_1) + (k-1)S(B'_2) + S(B'_3) & \text{if } n = 12k + 8. \end{cases}$$

Writing  $n = 12k + r$  with  $r \in \{0, 4, 8\}$ , we shall use the notation

$$w_{k,r}(x_1, x_2, x_3, x_4) = S(\Delta Q_1[n]x_1x_2x_3x_4).$$

Computing explicitly  $S(A'_i), S(B'_i), S(C'_i)$  from the above figures, we get:

$$\begin{aligned} w_{k,0}(x_1, x_2, x_3, x_4) &= 5x_1 + x_2(-1 + x_1) \\ &\quad + x_3(1 - x_1 + x_2 - 2x_1x_2) \\ &\quad + x_4(1 + x_2 + x_3 + 3x_1x_2x_3) \\ &\quad + k[-4x_1 + 2x_2(1 - x_1) \\ &\quad \quad + x_3(-1 + x_1 - 3x_2 + 3x_1x_2) \\ &\quad \quad + x_4(-1 + x_2 - x_3 - x_1x_2x_3)], \end{aligned}$$

$$\begin{aligned} w_{k,4}(x_1, x_2, x_3, x_4) &= 3x_1 + x_2(1 + x_1) + x_3(2 + 2x_1 + x_1x_2) \\ &\quad + 2x_4(1 + x_3) \\ &\quad + k[4x_1 - 2x_2(1 + x_1) \\ &\quad \quad + x_3(1 + x_1 - 3x_2 - 3x_1x_2) \\ &\quad \quad + x_4(-1 - x_2 + x_3 + x_1x_2x_3)], \end{aligned}$$

$$\begin{aligned} w_{k,8}(x_1, x_2, x_3, x_4) &= 3x_1 + x_2(3 - x_1) + x_3(1 + x_1 - x_2) \\ &\quad + x_4(3 + x_2 + x_3 - x_1x_2x_3) \\ &\quad + k[4x_1 + 2x_2(1 + x_1) \\ &\quad \quad + x_3(-1 - x_1 - 3x_2 - 3x_1x_2) \\ &\quad \quad + x_4(1 - x_2 + x_3 - x_1x_2x_3)]. \end{aligned}$$

Successively replacing  $(x_1, x_2, x_3, x_4)$  by each of the 16 binary sequences of length 4, we obtain 48 polynomial functions of degree 1 in  $k$ . We must then determine the zeroes of these polynomials.

**Case 1** :  $r = 0$ , *i.e.* we consider sequences of the type  $Q_1[12k]x_1x_2x_3x_4$ . We obtain the following values of:

$$w_{k,0}(x_1, x_2, x_3, x_4) = S(\Delta Q_1[12k]x_1x_2x_3x_4) :$$

$$\begin{aligned}
w_{k,0}(1, 1, 1, 1) &= 10 - 6k \\
w_{k,0}(1, 1, 1, -1) &= -2 - 2k \\
w_{k,0}(1, 1, -1, 1) &= 4 - 2k \\
w_{k,0}(1, 1, -1, -1) &= 8 - 6k \\
w_{k,0}(1, -1, 1, 1) &= 4 - 6k \\
w_{k,0}(1, -1, 1, -1) &= 8 - 2k \\
w_{k,0}(1, -1, -1, 1) &= 6 - 6k \\
w_{k,0}(1, -1, -1, -1) &= 2 - 2k \\
w_{k,0}(-1, 1, 1, 1) &= -2 \\
w_{k,0}(-1, 1, 1, -1) &= -2 \\
w_{k,0}(-1, 1, -1, 1) &= -8 + 16k \\
w_{k,0}(-1, 1, -1, -1) &= -16 + 16k \\
w_{k,0}(-1, -1, 1, 1) &= 0 \\
w_{k,0}(-1, -1, 1, -1) &= -8 + 8k \\
w_{k,0}(-1, -1, -1, 1) &= -6 - 4k \\
w_{k,0}(-1, -1, -1, -1) &= 2 - 4k.
\end{aligned}$$

Given that  $w_{k,0}(-1, -1, 1, 1) = 0$ , independently of  $k$ , we see that the sequence  $Q_1[12k] - - + +$  is (simply, hence strongly) balanced. But, as easily checked,  $Q_1[12k] - - + + = Q_1[12k + 4]$ . The 15 other polynomials may vanish for small values of  $k$ , yielding "exotic" short strongly balanced sequences. However, direct inspection reveals that none of these other functions vanishes for  $k \geq 5$ .

Consequently,  $Q_1[12k + 4]$  is the unique balanced extension of length  $12k + 4$  of  $Q_1[12k]$ , provided  $k \geq 5$ . Note that, in the context of this proof, we have  $k \geq 7$  in fact, since we are assuming  $n \geq 92$ .

**Case 2 :**  $r = 4$ , *i.e.* we consider the binary sequences  $Q_1[12k + 4]x_1x_2x_3x_4$ . We obtain the following values of  $w_{k,4}(x_1, x_2, x_3, x_4)$  :

$$\begin{aligned}
w_{k,4}(1, 1, 1, 1) &= 14 - 4k \\
w_{k,4}(1, 1, 1, -1) &= 6 - 4k \\
w_{k,4}(1, 1, -1, 1) &= 0 \\
w_{k,4}(1, 1, -1, -1) &= 8k \\
w_{k,4}(1, -1, 1, 1) &= 8 + 16k \\
w_{k,4}(1, -1, 1, -1) &= 16k
\end{aligned}$$

$$\begin{aligned}
w_{k,4}(1, -1, -1, 1) &= -2 \\
w_{k,4}(1, -1, -1, -1) &= -2 \\
w_{k,4}(-1, 1, 1, 1) &= -6k \\
w_{k,4}(-1, 1, 1, -1) &= -8 - 2k \\
w_{k,4}(-1, 1, -1, 1) &= -2 - 6k \\
w_{k,4}(-1, 1, -1, -1) &= -2 - 2k \\
w_{k,4}(-1, -1, 1, 1) &= 2 - 2k \\
w_{k,4}(-1, -1, 1, -1) &= -6 - 6k \\
w_{k,4}(-1, -1, -1, 1) &= -4 - 6k \\
w_{k,4}(-1, -1, -1, -1) &= -4 - 2k.
\end{aligned}$$

Here we have  $w_{k,4}(1, 1, -1, 1) = 0$ , independently of  $k$ . Thus, the sequence  $Q_1[12k + 4] + + - +$  is (simply, hence strongly) balanced. Again, one easily checks that  $Q_1[12k + 4] + + - + = Q_1[12k + 8]$ . The other 15 functions do not vanish for  $k \geq 2$ .

Therefore,  $Q_1[12k + 8]$  is the unique balanced extension of length  $12k + 8$  of  $Q_1[12k + 4]$ , provided  $k \geq 2$ .

**Case 3 :**  $r = 8$ , i.e. we consider the binary sequences  $Q_1[12k + 8]x_1x_2x_3x_4$ . Here are the values of  $w_{k,8}(x_1, x_2, x_3, x_4)$  :

$$\begin{aligned}
w_{k,8}(1, 1, 1, 1) &= 10 \\
w_{k,8}(1, 1, 1, -1) &= 2 \\
w_{k,8}(1, 1, -1, 1) &= 8 + 16k \\
w_{k,8}(1, 1, -1, -1) &= 16k \\
w_{k,8}(1, -1, 1, 1) &= 8 + 8k \\
w_{k,8}(1, -1, 1, -1) &= 0 \\
w_{k,8}(1, -1, -1, 1) &= -2 - 4k \\
w_{k,8}(1, -1, -1, -1) &= -2 - 4k \\
w_{k,8}(-1, 1, 1, 1) &= 6 - 2k \\
w_{k,8}(-1, 1, 1, -1) &= -6 - 6k \\
w_{k,8}(-1, 1, -1, 1) &= 4 - 6k \\
w_{k,8}(-1, 1, -1, -1) &= -2k \\
w_{k,8}(-1, -1, 1, 1) &= -4 - 2k \\
w_{k,8}(-1, -1, 1, -1) &= -8 - 6k
\end{aligned}$$

$$\begin{aligned}w_{k,8}(-1, -1, -1, 1) &= -6 - 2k \\w_{k,8}(-1, -1, -1, -1) &= -10 - 6k.\end{aligned}$$

Again,  $w_{k,8}(1, -1, 1, -1) = 0$ , but none of the other 15 functions vanishes for  $k \geq 4$ . Moreover,  $Q_1[12k + 8] + - + - = Q_1[12k + 12]$ . Therefore,  $Q_1[12k + 12]$  is the unique balanced extension of length  $12k + 12$  of  $Q_1[12k + 8]$  for  $k \geq 4$ .

With the above three cases, we have verified that, for every  $n \equiv 0 \pmod{4}$  with  $n \geq 52$ , the sequence  $Q_1[n]$  admits a unique balanced binary extension of length  $n + 4$ , namely  $Q_1[n + 4]$ .

Similar phenomena as those described here for  $Q_1$  occur for the other sequences  $Q_2, Q_3, Q_4, R_1, \dots, R_{12}$ , and for the supplementary strongly balanced sequences described in Theorem 2.2. This explains why, after a somewhat chaotic initial behavior, the set  $\mathcal{SB}(n)$  of strongly balanced binary sequences of length  $n$  ultimately becomes periodic.

### Acknowledgement

We are grateful to Pierre Duchet for attracting our attention to the problem of Steinhaus. It is worthwhile to note that Duchet uses this problem with school pupils, by challenging them to find as large balanced binary sequences as possible. Interestingly, some pupils were able to discover such sequences in length as high as 123, or even 240 [1].

We also thank Michel Kervaire, Pierre de la Harpe and the anonymous referee for useful comments about this paper.

Shalom Eliahou  
*Laboratoire de Mathématiques  
Pures et Appliquées Joseph  
Liouville*  
Université du Littoral Côte  
d'Opale  
B.P. 699, 62228 Calais cedex,  
France.  
eliahou@lmpa.univ-littoral.fr

Delphine Hachez  
*Laboratoire de Mathématiques  
Pures et Appliquées Joseph  
Liouville*  
Université du Littoral Côte  
d'Opale  
B.P. 699, 62228 Calais cedex,  
France.  
hachez@lmpa.univ-littoral.fr

## References

- [1] Duchet P., *La règle des signes*, MATH.en.JEANS (1995), 139–140.
- [2] Harborth H., *Solution of Steinhaus's problem with plus and minus signs*, J. Comb. Th. (A) **12** (1972), 253–259.
- [3] Steinhaus H., *One Hundred Problems in Elementary Mathematics*, Pergamon, Elinsford, N.Y., 1963.