Journal of Integer Sequences, Vol. 29 (2026),
Article 26.4.1

Meta-Automatic Sequences

John M. Campbell
Department of Mathematics and Statistics
Dalhousie University
Halifax, NS B3H 4R2
Canada
jh241966@dal.ca

Benoît Cloitre
19, rue Louise Michel
92300 Levallois-Perret
France
benoit7848c@yahoo.fr

Abstract

Nested (or *meta-Fibonacci*) recurrences, such as the recurrence used to define Hofstadter's Q-sequence, along with the *digit-based* recurrences that underlie automatic sequences are of interest from both number-theoretic and combinatorial points of view. In this direction, Allouche and Shallit showed how the frequency sequence of a variant of the Q-sequence is 2-automatic. This inspires us to introduce what may be seen as a natural combination of the recurrences for meta-Fibonacci and automatic sequences, by introducing the concept of a *meta-automatic sequence*. We exhibit two binary meta-automatic sequences $\mathcal{M}_1$ and $\mathcal{M}_2$ whose defining recurrences do not satisfy the Allouche-Shallit automaticity criterion directly, and this is formalized in our paper. For each of these integer sequences $\mathcal{M}_1$ and $\mathcal{M}_2$, we prove explicit DFAO evaluations, together with 4-uniform morphisms, and we also consider the factor complexities of these sequences.

1 Introduction

A *nested recurrence* is a recurrence in which a term $a(n)$ is defined using values of a at indices that themselves depend on earlier values of the sequence. The classical example is Hofstadter's Q-sequence [20, p. 137], indexed in the On-Line Encyclopedia of Integer Sequences [30] as A005185, defined by $Q(1) = Q(2) = 1$ and

$$Q(n) = Q(n - Q(n-1)) + Q(n - Q(n-2)) \tag{1}$$

for $n > 2$. Recurrences as in (1) are also known as *meta-Fibonacci recurrences*. Other examples of integer sequences satisfying nested recurrences include the sequence A244477 studied by Golomb [18]. Many problems concerning Hofstadter's Q-sequence and its generalizations remain open to this day. Notably, it is not known whether or not $Q(n)$ is even *defined* for all $n > 0$.

In contrast to nested recurrences, *digit-based recurrences* (also called *divide-and-conquer recurrences*) define $a(k^t n + j)$ for $j \in \{0, 1, \ldots, k^t - 1\}$ as a function of previously computed values $a(k^{t'} n' + j')$ with $t' < t$ and $j' \in \{0, 1, \ldots, k^{t'} - 1\}$. Digit-based recurrences can be used to define *automatic sequences* [3, 9], which appear in many areas of number theory, computer science, and combinatorics.

Allouche and Shallit [4] showed that the frequency sequence of a variant of Hofstadter's Q-sequence satisfies a digit-based recurrence, and, moreover, is 2-automatic. They also proved an automaticity criterion [4, Theorem 2.2] reviewed in Section 1.2 below. This inspires us to introduce the concept of a *meta-automatic sequence*, which may, informally, be thought of as "combining" the concepts of meta-Fibonacci and automatic sequences. This is formalized below. For example, in contrast to a rule such as $a(4n + 2) = a(2n)$ used to define an automatic sequence, we consider rules such as

$$a(4n + 2) = a(2n + 1 - a(n)), \tag{2}$$

in which the argument of a on the right-hand side depends on the value $a(n)$. A recurrence of the form suggested in (2) would not, in general, satisfy the Allouche-Shallit automaticity criterion, since the argument on the right is not of the required polynomial form.

The idea of combining automatic and meta-Fibonacci recurrences is motivated by past work on meta-Fibonacci sequences, in addition to the above referenced work by Allouche and Shallit [4]. We point to references on the Hofstadter Q-sequence [16, 17, 19, 27], on variants and generalizations of the Hofstadter Q-sequence [1, 2, 5, 12, 14, 32], and on meta-Fibonacci sequences [7, 8, 10, 11, 13, 15, 21, 22, 23, 28, 31].

1.1 Outline

A formal definition of a meta-automatic sequence is given in Section 2. A meta-automatic recurrence may, after simplification, reduce to a digit-based recurrence, in which case the Allouche-Shallit criterion applies directly. We formalize, in this paper, how meta-automatic sequences are not necessarily reducible in this way.

A binary sequence $a : \mathbb{N}_0 \to \{0, 1\}$ is said to be *balanced* if

$$a(2n) + a(2n + 1) = 1 \qquad (n \geq 0). \tag{3}$$

The binary sequences under consideration in this paper are, for the most part, balanced. In a related way, much of this paper makes use of the additive operation underlying the field $\mathbb{F}_2 = \{0, 1\}$, letting $\oplus$ denote this operation, which may be referred to as *XOR*, with $0 \oplus 0 = 1 \oplus 1 = 0$ and $0 \oplus 1 = 1 \oplus 0 = 1$. For balanced binary sequences, Lemma 5 gives the identity

$$a(2n + 1 - a(n)) = a(2n + 1) \oplus a(n).$$

We prove that balanced, binary, meta-automatic sequences are, for the cases covered in Lemma 6, ultimately periodic. To obtain non-periodic examples, we therefore work with base-4 recurrences. We consider:

- The integer sequence $\mathcal{M}_1$ (A392736), defined via a base-4 recurrence in which one of the rules in the recurrence is nested in the sense outlined above. After applying Lemma 5 below, the pair $(\mathcal{M}_1(n), b(n))$, for $b(n)$ defined in (8) below, yields a 4-state DFAO (Section 3).
- The integer sequence $\mathcal{M}_2$ (A391614), defined via a base-4 recurrence in which *both* of the rules in the recurrence are nested in the sense outlined above. After applying Lemma 5, the pair $\big(\mathcal{M}_2(n), \mathcal{M}_2(2n + \mathcal{M}_2(n))\big)$ obeys an affine rule over $\mathbb{F}_2^2$ with linear part $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$. The sequence satisfies $\mathcal{M}_2(n) = \mathbf{t}(q(n))$, where q is an explicit bit-masking operator and $\mathbf{t}$ is the Thue-Morse sequence (Theorem 26).

Although $\mathcal{M}_1$ and $\mathcal{M}_2$ both have 4-state DFAOs, we have that $\mathcal{M}_2$ admits the evaluation $\mathcal{M}_2(n) = \mathbf{t}(q(n))$, whereas $\mathcal{M}_1$ is described by means of the pair $(\mathcal{M}_1(n), b(n))$. Their factor-complexity functions, studied in Section 5, also differ. For each of the sequences $\mathcal{M}_1$ and $\mathcal{M}_2$, we give a base-4 DFAO (MSB-first) and an explicit 4-uniform morphism, and we also investigate the factor complexity functions for $\mathcal{M}_1$ and $\mathcal{M}_2$ in Section 5.

1.2 Preliminaries

For background on automatic sequences, we refer to the texts of Allouche and Shallit [3] and of Shallit [29]. We proceed with the following definition, which may be used, as below, to define automatic sequences.

Definition 1. Let $U = (U(n))_{n\geq 0}$ denote an infinite sequence. The *k-kernel* of $(U(n))_{n\geq 0}$ [3, p. 185] is

$$K_k(U) = \left\{ \left(U_{k^i n + j}\right)_{n\geq 0} : i \geq 0 \text{ and } 0 \leq j < k^i \right\}.$$

For $k \geq 2$, a *k-automatic sequence* is a sequence $U = (U(n))_{n\geq 0}$ such that $K_k(U)$ is finite. Equivalently, the sequence U is k-automatic if there exists a DFAO such that, upon inputting the base-k representation of n, this DFAO outputs $U(n)$. See Allouche and Shallit's text [3] for details. A prototypical instance of an automatic sequence is the *Thue-Morse sequence* $\mathbf{t} : \mathbb{N}_0 \to \{0, 1\}$ defined so that $\mathbf{t}(0) = 0$ and

$$\mathbf{t}(2n) = \mathbf{t}(n), \qquad \mathbf{t}(2n+1) = 1 - \mathbf{t}(n) \qquad (n \geq 0).$$

Equivalently, $\mathbf{t}(n) = s_2(n) \bmod 2$, where $s_2(n)$ is the sum of the binary digits of n, noting that the balanced condition in (3) holds. We introduce, in this paper, a variant of an automaticity criterion due to Allouche and Shallit [4], which we now recall.

Let $U = (U(n))_{n\geq 0}$ denote a sequence, and let α denote an integer. Adopting notation from Allouche and Shallit's work [4], for $n \geq -\alpha$, we write U^α in place of the sequence given by

$$U^\alpha := U(n + \alpha).$$

Similarly, letting q and i and j be positive integers, we write $U_{q,i,j}$ in place of the sequence defined by

$$U_{q,i,j}(n) = U\left(q^i n + j\right) \tag{4}$$

for $n \geq 0$. Moreover, for integers $q \geq 2$ and $t \geq 1$, we write $s(q, t) = \frac{1-q^{t+1}}{1-q}$ in place of the number of pairs (i, j) of integers i and j such that $0 \leq i \leq t$ and $0 \leq j \leq q^i - 1$. Fix an ordering of these pairs (i, j), with $(0, 0)$ appearing first, and write the corresponding sequences as $U_1 = U$, $U_2, \ldots, U_{s(q,t)}$. The choice of ordering plays no role in what follows. We also let $m(q, t) = q^{t+1} - 1$.

Theorem 2. *(Allouche and Shallit, 2012) Let the terms of $(U(n))_{n\geq 0}$ be in a finite set $\mathcal{A}$, and let $q \geq 2$ denote an integer. Then U is q-automatic if there are nonnegative integers t, a, b, and n_0 together with a family $\{f_j : j = 0, 1, \ldots, m(q, t)\}$ of functions from $\mathcal{A}^{a+b+s(q,t)}$ to $\mathcal{A}$ such that*

$$\begin{aligned} U\big(q^{t+1}n + j\big) = f_j\big(&U^{-a}(n), \ldots, U^{-1}(n), U^0(n), U^1(n), \ldots, \\ &U^b(n), U_2(n), U_3(n), \ldots, U_{s(q,t)}(n)\big) \end{aligned}$$

for all $n \geq n_0$ and for all $j \in \{0, 1, \ldots, m(q, t)\}$ [4].

2 Meta-automatic sequences

Let $q > 0$, $i_0 > 0$, and $j_0 \geq 0$ be integers. For some $v \geq 0$, and for indices $k \in \{1, 2, \ldots, v\}$, let c_k, $i_k > 0$, and $j_k \geq 0$ be integers, with $i_k < i_0$ for all k. By analogy with (4), and assuming that the codomain of U is contained in $\mathbb{N}_0$, we write

$$U_{q,i_0,j_0,c_1,i_1,j_1,\ldots,c_v,i_v,j_v}(n) = U\Big(q^{i_0}n + j_0 + c_1U(q^{i_1}n + j_1) + \cdots + c_vU(q^{i_v}n + j_v)\Big).$$

The following definition may be thought of as formalizing the intuition behind recurrences as in (2).

Definition 3. Suppose that each term of $(U(n))_{n\geq 0}$ is in a finite subset $\mathcal{A}$ of $\mathbb{N}_0$, and let $q \geq 2$ be an integer. The sequence U is *q-meta-automatic* if there exist nonnegative integers t, a, b, n_0 together with a family $\{f_j : j = 0, 1, \ldots, m(q,t)\}$ of functions to $\mathcal{A}$, with domains as below, such that, for all $n \geq n_0$ and all $j \in \{0, 1, \ldots, m(q,t)\}$, the value $U(q^{t+1}n + j)$ equals f_j evaluated at the tuple

$$\Big(U^{-a}(n), \ldots, U^{-1}(n), U^{0}(n), U^{1}(n), \ldots, U^{b}(n),$$
$$U_{q,i_0^{(1)},j_0^{(1)},c_1^{(1)},i_1^{(1)},j_1^{(1)},\ldots,c_{v_1}^{(1)},i_{v_1}^{(1)},j_{v_1}^{(1)}}(n), \ldots,$$
$$U_{q,i_0^{(w)},j_0^{(w)},c_1^{(w)},i_1^{(w)},j_1^{(w)},\ldots,c_{v_w}^{(w)},i_{v_w}^{(w)},j_{v_w}^{(w)}}(n)\Big)$$

for some $w \geq 0$ and integers $i_{k_1}^{(k_2)} \leq t$, $j_{k_1}^{(k_2)}$, and $c_{k_1}^{(k_2)}$, where the domain of f_j is $\mathcal{A}^{a+b+w+1}$.

Example 4. Define $(\mathcal{T}(n))_{n\geq 0}$ by $\mathcal{T}(0) = 0$ and

$$\mathcal{T}(2n) = \mathcal{T}(n - \mathcal{T}(n-1)) \text{ for } n \geq 1, \text{ and}$$
$$\mathcal{T}(2n+1) = 1 - \mathcal{T}(n) \text{ for } n \geq 0.$$

Then $(\mathcal{T}(n))_{n\geq 0}$ is 2-meta-automatic in the sense of Definition 3. For $n \geq 1$, this sequence agrees with the OEIS sequence A039982, which was defined without nested recurrences.

Lemma 5. *If a is balanced, then*

$$a(2n + a(n)) = a(2n) \oplus a(n) \quad \textit{and} \quad a(2n + 1 - a(n)) = a(2n+1) \oplus a(n),$$

for all $n \geq 0$.

Proof. The first identity follows from (3) by case analysis on $a(n) \in \{0, 1\}$, and similarly for the second. □

2.1 Base-4 recurrences

We now record a result that motivates the use of base-4 recurrences in the sequel. We let

$$a(2n) = a(n + a(n)) \qquad \text{or} \qquad a(2n) = a(n + 1 - a(n)),$$

and we let $a(2n+1) = 1 - a(2n)$. In both cases, we obtain ultimate periodicity, as we now show.

Lemma 6. *Let $a : \mathbb{N}_0 \to \{0, 1\}$ be balanced.*

(a) *If* $a(2n) = a(n + a(n))$ *for all* $n \geq 0$*, then* $a(0) = 0$ *and* a *is ultimately* 2*-periodic, with*

$$a(2n) = 0, \qquad a(2n+1) = 1 \qquad (n \geq 2).$$

(b) *If* $a(2n) = a(n + 1 - a(n))$ *for all* $n \geq 0$*, then* $a(0) = 1$ *and* a *is ultimately* 2*-periodic, with*

$$a(2n) = 1, \qquad a(2n+1) = 0 \qquad (n \geq 2).$$

Proof.

(a) Taking $n = 0$ gives $a(0) = a(a(0))$. If $a(0) = 1$, then balance gives us that $a(1) = 0$, and hence $a(0) = a(1) = 0$, a contradiction. Thus $a(0) = 0$ and $a(1) = 1$.

Write $b_m = a(2m)$, so that balance gives that $a(2m+1) = 1 - b_m$. Using $a(2n) = a(n + a(n))$ with $n = 2m$, we obtain

$$b_{2m} = a(4m) = a(2m + a(2m)) = a(2m + b_m) = \begin{cases} a(2m) = b_m, & \text{if } b_m = 0; \\ a(2m+1) = 1 - b_m = 0, & \text{if } b_m = 1. \end{cases}$$

So we have that $b_{2m} = 0$ for all $m \geq 0$.

Now, using $a(2n) = a(n + a(n))$ with $n = 2m + 1$ and $a(2m+1) = 1 - b_m$, we get

$$b_{2m+1} = a(4m+2) = a(2m + 1 + a(2m+1)) = a(2m + 2 - b_m) = \begin{cases} a(2m+2) = b_{m+1}, & \text{if } b_m = 0; \\ a(2m+1) = 0, & \text{if } b_m = 1. \end{cases}$$

We claim that $b_m = 0$ for all $m \geq 2$. This holds for $m = 2$, since $b_2 = 0$, and for $m = 3$ because the above identity with $m = 1$ gives $b_3 = b_2$ if $b_1 = 0$, and $b_3 = 0$ if $b_1 = 1$; in either case $b_3 = 0$ since $b_2 = 0$. Now let $m \geq 4$. If m is even then $b_m = b_{2r} = 0$. If m is odd, write $m = 2r + 1$ with $r \geq 1$. If $b_r = 1$ then $b_m = 0$; if $b_r = 0$ then $b_m = b_{r+1}$, and $r + 1 < m$, so by induction $b_{r+1} = 0$. Hence $b_m = 0$ for all $m \geq 2$, i.e., $a(2n) = 0$ for all $n \geq 2$. The balance condition in (3) then gives $a(2n+1) = 1$ for all $n \geq 2$.

(b) This follows from (a) by complementing: if a satisfies (b) then $\bar{a}(n) = 1 - a(n)$ satisfies (a), and vice versa.

$\square$

Lemma 6 leads us to consider base-4 recurrences such that the balance identity in (3) is preserved.

2.2 A non-nested case

Fix $a(0) = 0$ and $a(1) = 1$, and define

$$a(4n) = F(n), \quad a(4n+1) = 1 - F(n), a(4n+2) = G(n), \quad a(4n+3) = 1 - G(n), \tag{5}$$

so that $a(2n) + a(2n+1) = 1$ holds automatically. The sequences studied in this paper arise as specializations of (5). We first consider a sequence $(\mathcal{Q}(n))_{n\geq 0}$ whose recurrence is not nested, but which is used in the sequel.

Definition 7. The *Thue-Morse Quarto* sequence $\mathcal{Q} : \mathbb{N}_0 \to \{0,1\}$ (A298952, where we use the convention $\mathcal{Q}(0) = 0$ rather than $a(0) = 1$ as in the OEIS entry) is defined by $\mathcal{Q}(0) = 0$, $\mathcal{Q}(1) = 1$, and

$$\begin{aligned} \mathcal{Q}(4n) &= \mathcal{Q}(n), & \mathcal{Q}(4n+1) &= 1 - \mathcal{Q}(n), \\ \mathcal{Q}(4n+2) &= \mathcal{Q}(2n), & \mathcal{Q}(4n+3) &= 1 - \mathcal{Q}(2n). \end{aligned}$$

The sequence $\mathcal{Q}$ is 4-automatic. It is computed by the 3-state DFAO with states in $\{A, B, C\}$, initial state A, and output $\tau(A) = \tau(B) = 0$, $\tau(C) = 1$, shown in Figure 1. This DFAO corresponds to the 4-uniform morphism

$$A \mapsto ACBC, \quad B \mapsto BCCB, \quad C \mapsto CBBC,$$

along with a coding such that $\tau(A) = \tau(B) = 0$ and $\tau(C) = 1$.

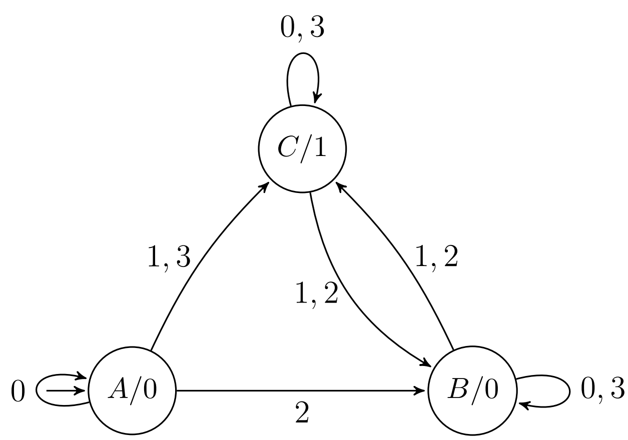

Figure 1: DFAO for $\mathcal{Q}$ (MSB-first). State labels are *state/output*. The initial state is A.

Proposition 8. *Define* $d : \mathbb{N}_0 \to \{0,1\}$ *so that* $d(0) = 0$ *and* $d(n) = \lfloor \log_2 n \rfloor \bmod 2$ *for* $n \geq 1$. *Then*

$$\mathcal{Q}(n) = \mathbf{t}(n) \oplus d(n) \qquad (n \geq 0).$$

Proof. The cases $n = 0, 1$ are immediate. For $n \geq 1$ and $r \in \{0, 1, 2, 3\}$, since $\lfloor \log_2(4n+r) \rfloor = 2 + \lfloor \log_2 n \rfloor$, we have $d(4n + r) = d(n)$ for all r. We also use $\mathbf{t}(4n + r) = \mathbf{t}(n) \oplus \mathbf{t}(r)$, and we proceed by induction.

Case $r = 0$: $\mathcal{Q}(4n) = \mathcal{Q}(n) = \mathbf{t}(n) \oplus d(n)$ and $\mathbf{t}(4n) \oplus d(4n) = \mathbf{t}(n) \oplus d(n)$.

Case $r = 1$: $\mathcal{Q}(4n+1) = 1 \oplus \mathcal{Q}(n) = 1 \oplus \mathbf{t}(n) \oplus d(n)$ and $\mathbf{t}(4n+1) \oplus d(4n+1) = (1 \oplus \mathbf{t}(n)) \oplus d(n)$.

Case $r = 2$: $\mathcal{Q}(4n+2) = \mathcal{Q}(2n)$. By induction, $\mathcal{Q}(2n) = \mathbf{t}(2n) \oplus d(2n) = \mathbf{t}(n) \oplus (d(n) \oplus 1) = 1 \oplus \mathbf{t}(n) \oplus d(n)$. Also $\mathbf{t}(4n+2) = \mathbf{t}(2n+1) = 1 \oplus \mathbf{t}(n)$, so $\mathbf{t}(4n+2) \oplus d(4n+2) = (1 \oplus \mathbf{t}(n)) \oplus d(n)$.

Case $r = 3$: $\mathcal{Q}(4n+3) = 1 \oplus \mathcal{Q}(2n) = \mathbf{t}(n) \oplus d(n)$, and $\mathbf{t}(4n+3) \oplus d(4n+3) = \mathbf{t}(n) \oplus d(n)$. □

3 The sequence $\mathcal{M}_1$

Definition 9. Define $\mathcal{M}_1(0) = 0$ and $\mathcal{M}_1(1) = 1$. For $n \geq 1$, set

$$\mathcal{M}_1(4n) = \mathcal{M}_1(n), \qquad \mathcal{M}_1(4n + 1) = 1 - \mathcal{M}_1(n). \tag{6}$$

For $n \geq 0$, set

$$\mathcal{M}_1(4n + 2) = \mathcal{M}_1(2n + 1 - \mathcal{M}_1(n)), \qquad \mathcal{M}_1(4n + 3) = 1 - \mathcal{M}_1(2n + 1 - \mathcal{M}_1(n)). \tag{7}$$

Lemma 10. *The sequence $\mathcal{M}_1(n)$ is well-defined for all $n \geq 0$ and satisfies $\mathcal{M}_1(2n) + \mathcal{M}_1(2n + 1) = 1$.*

Proof. Well-definedness follows by induction on m. For $m = 4n + r$, the right-hand sides in (6) reduce to indices $< m$; for (7), the nested argument $2n + 1 - \mathcal{M}_1(n) \in \{2n, 2n + 1\} < 4n + 2 \leq m$. Balance follows from (5), since $\mathcal{M}_1(4n + 1) = 1 - \mathcal{M}_1(4n)$ and $\mathcal{M}_1(4n + 3) = 1 - \mathcal{M}_1(4n + 2)$. □

Remark 11. Since $\mathcal{M}_1$ is balanced, Lemma 5 gives

$$\mathcal{M}_1(2n + 1 - \mathcal{M}_1(n)) = \mathcal{M}_1(2n + 1) \oplus \mathcal{M}_1(n),$$

so (7) is equivalent to the non-nested relation $\mathcal{M}_1(4n + 2) = \mathcal{M}_1(2n + 1) \oplus \mathcal{M}_1(n)$.

3.1 A 4-state DFAO

Define

$$b(n) := \mathcal{M}_1(2n + 1 - \mathcal{M}_1(n)), \tag{8}$$

and consider the pair $s(n) := (\mathcal{M}_1(n), b(n))$. Informally, the following result gives us that the pair $s(4n + r)$ depends only on $s(n)$, for $r \in \{0, 1, 2, 3\}$ and for $n \geq 1$, disregarding the trivial $n = 0$ case whereby $\mathcal{M}_1(0) = 0$.

Lemma 12. *Let $x = \mathcal{M}_1(n)$ and $y = b(n)$, and let $n \geq 1$. Then*

$$s(4n) = (x, y), \tag{9}$$
$$s(4n+1) = (1 \oplus x,\ x \oplus y), \tag{10}$$
$$s(4n+2) = (y,\ 1 \oplus x), \text{ and} \tag{11}$$
$$s(4n+3) = (1 \oplus y,\ x \oplus y \oplus 1). \tag{12}$$

Proof. By Lemma 10 and Lemma 5, we have $y = b(n) = \mathcal{M}_1(2n+1) \oplus \mathcal{M}_1(n) = (1 - \mathcal{M}_1(2n)) \oplus x$, so that

$$\mathcal{M}_1(2n) = 1 \oplus x \oplus y, \qquad \mathcal{M}_1(2n+1) = x \oplus y.$$

Case $r = 0$. The first coordinate is $\mathcal{M}_1(4n) = \mathcal{M}_1(n) = x$. The latter coordinate satisfies $b(4n) = \mathcal{M}_1(8n + 1 - \mathcal{M}_1(4n)) = \mathcal{M}_1(8n + 1 - x)$. If $x = 0$, then $b(4n) = \mathcal{M}_1(8n+1) = \mathcal{M}_1(4(2n)+1) = 1 - \mathcal{M}_1(2n) = 1 - (1 \oplus x \oplus y) = 1 - (1 \oplus y) = y$. If $x = 1$, then $b(4n) = \mathcal{M}_1(8n) = \mathcal{M}_1(4(2n)) = \mathcal{M}_1(2n) = 1 \oplus x \oplus y = y$. In both cases, we find that (9) holds.

Case $r = 1$. The first coordinate is $\mathcal{M}_1(4n+1) = 1 - x = 1 \oplus x$. The latter coordinate satisfies $b(4n+1) = \mathcal{M}_1(8n+2+x)$. If $x = 0$, then $b(4n+1) = \mathcal{M}_1(8n+2) = \mathcal{M}_1(4(2n)+2) = b(2n)$. Using the property that $b(2n) = \mathcal{M}_1(4n + 1 - \mathcal{M}_1(2n))$, together with $\mathcal{M}_1(2n) = 1 \oplus y$, this yields $b(2n) = \mathcal{M}_1(4n + y)$. So, if $y = 0$, then $\mathcal{M}_1(4n) = x = 0$, and if $y = 1$, then $\mathcal{M}_1(4n+1) = 1 - x = 1$, so that $b(4n+1) = y$. Similarly, if $x = 1$, then $b(4n+1) = \mathcal{M}_1(8n+3) = \mathcal{M}_1(4(2n)+3) = 1 - b(2n)$. With $\mathcal{M}_1(2n) = y$, $b(2n) = \mathcal{M}_1(4n+1-y)$, we find that if $y = 0$, then $1 - \mathcal{M}_1(4n+1) = x = 1$, and if $y = 1$, then $1 - \mathcal{M}_1(4n) = 1 - x = 0$, so that $b(4n+1) = 1 \oplus y$. Consequently, we obtain that (10) holds.

Case $r = 2$. The first coordinate is $\mathcal{M}_1(4n+2) = b(n) = y$. The latter coordinate satisfies $b(4n+2) = \mathcal{M}_1(8n+5-y)$. If $y = 0$, then $\mathcal{M}_1(8n+5) = \mathcal{M}_1(4(2n+1)+1) = 1 - \mathcal{M}_1(2n+1) = 1 \oplus (x \oplus y) = 1 \oplus x$. If $y = 1$, then $\mathcal{M}_1(8n+4) = \mathcal{M}_1(4(2n+1)) = \mathcal{M}_1(2n+1) = x \oplus y = 1 \oplus x$. In both cases, we see that (11) holds.

Case $r = 3$. The first coordinate is $\mathcal{M}_1(4n+3) = 1 - y = 1 \oplus y$. The latter coordinate satisfies $b(4n+3) = \mathcal{M}_1(8n+6+y)$. If $y = 0$, then $\mathcal{M}_1(8n+6) = \mathcal{M}_1(4(2n+1)+2) = b(2n+1)$. Since $\mathcal{M}_1(2n+1) = x$, $b(2n+1) = \mathcal{M}_1(4n+3-x)$. So, if $x = 0$, then $\mathcal{M}_1(4n+3) = 1 \oplus y = 1$; if $x = 1$, then $\mathcal{M}_1(4n+2) = y = 0$; so $b(4n+3) = x \oplus y \oplus 1$. If $y = 1$, then $\mathcal{M}_1(8n+7) = \mathcal{M}_1(4(2n+1)+3) = 1 - b(2n+1)$. Since $\mathcal{M}_1(2n+1) = x \oplus 1$, we find that $b(2n+1) = \mathcal{M}_1(4n+2+x)$. So, if $x = 0$, then $1 - \mathcal{M}_1(4n+2) = 1 - y = 0$; if $x = 1$, then $1 - \mathcal{M}_1(4n+3) = y = 1$, so that $b(4n+3) = x \oplus y \oplus 1$. This gives us the desired relation in (12). ∎

Theorem 13. *The sequence $\mathcal{M}_1$ is 4-automatic and is computed (MSB-first, base 4) by the DFAO with states $\{0, 1, 2, 3\}$, initial state 0, output map $\tau(0) = \tau(2) = 0$, $\tau(1) = \tau(3) = 1$, and transition table*

δ	0	1	2	3
0	0	1	1	2
1	1	2	3	0
2	2	3	0	1
3	3	0	2	3

Proof. Lemma 12 gives four $\mathbb{F}_2$-affine transition rules:

$$(x, y) \xrightarrow{0} (x,\ y), \qquad (x, y) \xrightarrow{1} (1 \oplus x,\ x \oplus y),$$
$$(x, y) \xrightarrow{2} (y,\ 1 \oplus x), \quad (x, y) \xrightarrow{3} (1 \oplus y,\ x \oplus y \oplus 1).$$

Encoding the states as $\mathsf{0} = (0, 1)$, $\mathsf{1} = (1, 1)$, $\mathsf{2} = (0, 0)$, $\mathsf{3} = (1, 0)$, these yield the stated transition table. The initial state is $s(0) = (\mathcal{M}_1(0), b(0)) = (0, \mathcal{M}_1(1)) = (0, 1)$, corresponding to state $\mathsf{0}$. The output reads the first coordinate $x = \mathcal{M}_1(n)$. □

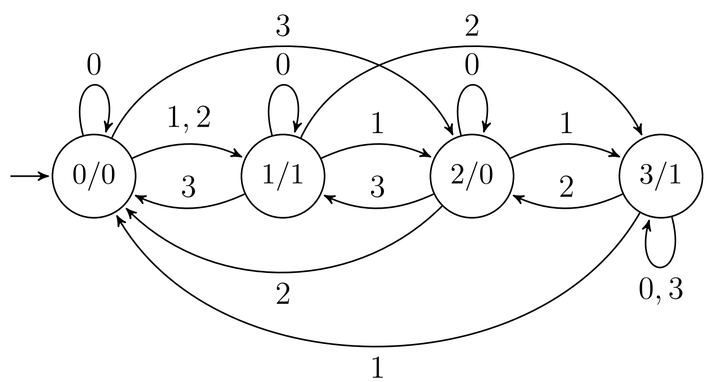

Figure 2: DFAO for $\mathcal{M}_1$ (4 states, base 4, MSB-first). State labels are *state/output*. The initial state is $\mathsf{0}$.

Remark 14. The four transition rules of the DFAO of $\mathcal{M}_1$ are $\mathbb{F}_2$-affine in (x, y); the same is true for $\mathcal{M}_2$ (Section 4). For each of the sequences $\mathcal{Q}$, $\mathcal{M}_1$, and $\mathcal{M}_2$, the affine form is a consequence of Lemma 5.

The defining recurrence of $\mathcal{M}_1$ is not directly an instance of Theorem 2, since the argument $2n + 1 - \mathcal{M}_1(n)$ in the nested recurrence in (7) depends on the value $\mathcal{M}_1(n)$. The b-sequence defined in (8) removes this dependence. Furthermore, Lemma 5 gives us that

$$\mathcal{M}_1(4n + 2) = \mathcal{M}_1(2n + 1) \oplus \mathcal{M}_1(n).$$

The XOR identity is a direct consequence of Lemma 5, since $\mathcal{M}_1$ is balanced.

Remark 15. By contrast, Allouche and Shallit [4] show that the frequency sequence $F(n)$ of the Hofstadter variant $V = Q_{1,4}$ satisfies a recurrence of the form $F(2a) = g(F(a-2), F(a-$

1), $F(a), F(a+1)$) with fixed shifts. This satisfies the required conditions in Theorem 2. In our setting, the sequences we consider *themselves*, as opposed to associated frequency sequences, are meta-automatic. In this direction, Lemma 5 provides a key tool, for our purposes.

Corollary 16. *The coding* $\tau\colon \mathsf{0},\mathsf{2} \mapsto 0$ *and* $\mathsf{1},\mathsf{3} \mapsto 1$ *applied to the fixed point of the* 4*-uniform morphism*

$$\mathsf{0} \to \mathsf{0112}, \quad \mathsf{1} \to \mathsf{1230}, \quad \mathsf{2} \to \mathsf{2301}, \quad \mathsf{3} \to \mathsf{3023}$$

starting from $\mathsf{0}$ *produces the sequence* $\mathcal{M}_1$.

Proof. For each state q, the morphism image is $\delta(q,0)\,\delta(q,1)\,\delta(q,2)\,\delta(q,3)$, read from the transition table of Theorem 13. ☐

4 The sequence $\mathcal{M}_2$

Definition 17. Define $\mathcal{M}_2(0) = 0$ and $\mathcal{M}_2(1) = 1$. For $n \geq 1$, set

$$\mathcal{M}_2(4n) = \mathcal{M}_2(2n + \mathcal{M}_2(n)), \qquad \mathcal{M}_2(4n+1) = 1 - \mathcal{M}_2(4n). \tag{13}$$

For $n \geq 0$, set

$$\mathcal{M}_2(4n+2) = \mathcal{M}_2(2n + 1 - \mathcal{M}_2(n)), \qquad \mathcal{M}_2(4n+3) = 1 - \mathcal{M}_2(4n+2). \tag{14}$$

Theorem 18. *The sequence* $\mathcal{M}_2$ *is well-defined and satisfies* (3)*.*

Proof. Well-definedness follows by induction as in Lemma 10. For the recurrence rules associated with the arguments $4n$ and $4n+1$, we have $n \geq 1$, so $2n + \mathcal{M}_2(n) \leq 2n+1 < 4n$. For the arguments $4n+2$ and $4n+3$, the nested argument satisfies $2n+1-\mathcal{M}_2(n) \leq 2n+1 < 4n+2$. Balance is immediate from the complementary definitions in (13)–(14). ☐

Define the pair

$$s(n) := \big(\mathcal{M}_2(n),\, \mathcal{M}_2(2n + \mathcal{M}_2(n))\big) \in \mathbb{F}_2^2. \tag{15}$$

Lemma 19. *If* $s(n) = (x,y)$*, then*

$$\mathcal{M}_2(2n) = x \oplus y, \qquad \mathcal{M}_2(2n+1) = x \oplus y \oplus 1.$$

Proof. If $x = 0$, then $y = \mathcal{M}_2(2n)$; if $x = 1$, then $y = \mathcal{M}_2(2n+1)$, and balance gives that $\mathcal{M}_2(2n) = 1 - y$. The identities follow. ☐

Theorem 20. *Writing* $s(n) = (x,y)$*, we have*

$$\begin{aligned} s(4n) &= (y,\ x \oplus y), \\ s(4n+1) &= (y \oplus 1,\ x \oplus y), \\ s(4n+2) &= (y \oplus 1,\ x \oplus y \oplus 1), \\ s(4n+3) &= (y,\ x \oplus y \oplus 1). \end{aligned}$$

Consequently $\mathcal{M}_2$ *is* 4*-automatic with at most* 4 *states.*

Proof. The case $n = 0$ is immediate from the initial values $\mathcal{M}_2(0) = 0$ and $\mathcal{M}_2(1) = 1$. We may therefore assume $n \geq 1$ whenever a rule for $4n$ or $4n+1$ is used. Write $s(n) = (x, y)$ so that $\mathcal{M}_2(2n) = x \oplus y$ and $\mathcal{M}_2(2n+1) = x \oplus y \oplus 1$.

Case $r = 0$. The first coordinate is given by $\mathcal{M}_2(4n) = \mathcal{M}_2(2n+x)$. By Lemma 19, we have that $\mathcal{M}_2(2n) = x \oplus y$ and $\mathcal{M}_2(2n+1) = x \oplus y \oplus 1$. In both cases, for $x = 0$ or $x = 1$, a direct check gives that $\mathcal{M}_2(2n+x) = y$. For the second coordinate $\mathcal{M}_2(8n+y)$, we see that if $y = 0$, then $\mathcal{M}_2(8n) = \mathcal{M}_2(4(2n)) = \mathcal{M}_2(2 \cdot 2n + \mathcal{M}_2(2n))$, and, since $\mathcal{M}_2(2n) = x \oplus y = x$, this equals $\mathcal{M}_2(4n+x)$. Simplifying both cases gives x. If $y = 1$, then $\mathcal{M}_2(8n+1) = 1 - \mathcal{M}_2(8n)$, and, by the same analysis as before, we get $x \oplus 1$. In both sub-cases the second coordinate equals $x \oplus y$. Therefore, we obtain

$$s(4n) = (y,\ x \oplus y).$$

Case $r = 1$. The first coordinate is given by $\mathcal{M}_2(4n+1) = 1 - \mathcal{M}_2(4n) = 1 - y = y \oplus 1$.

For the second coordinate of $s(4n+1)$, by definition, this is $\mathcal{M}_2(2(4n+1) + \mathcal{M}_2(4n+1)) = \mathcal{M}_2(8n+2+(y \oplus 1))$. When $y = 0$, we have $\mathcal{M}_2(8n+3) = \mathcal{M}_2(4(2n)+3) = 1 - \mathcal{M}_2(2(2n) + 1 - \mathcal{M}_2(2n))$. Since $\mathcal{M}_2(2n) = x$, this equals $1 - \mathcal{M}_2(4n+1-x) = x \oplus y$ for both values of x. When $y = 1$, we have $\mathcal{M}_2(8n+2) = \mathcal{M}_2(2(2n)+1-\mathcal{M}_2(2n))$. Since $\mathcal{M}_2(2n) = x \oplus 1$, this equals $\mathcal{M}_2(4n+x) = x \oplus y$ for both values of x. The second coordinate is therefore $x \oplus y$ in both cases. Therefore

$$s(4n+1) = (y \oplus 1,\ x \oplus y).$$

Case $r = 2$. For the first coordinate $\mathcal{M}_2(4n+2) = \mathcal{M}_2(2n+1-x)$, Lemma 5 gives $\mathcal{M}_2(2n+1) \oplus x = (x \oplus y \oplus 1) \oplus x = y \oplus 1$.

For the second coordinate of $s(4n+2)$, we find that $\mathcal{M}_2(2(4n+2) + \mathcal{M}_2(4n+2)) = \mathcal{M}_2(8n+4+(y \oplus 1))$. Since $\mathcal{M}_2(2n+1) = x \oplus 1$, we have $\mathcal{M}_2(8n+4) = \mathcal{M}_2(4n+3) = 1 - (y \oplus 1)$ for both values of x. When $y = 0$, this gives $\mathcal{M}_2(8n+4) = 0$, so $\mathcal{M}_2(8n+5) = 1 = x \oplus y \oplus 1$. When $y = 1$, since $\mathcal{M}_2(2n+1) = x$, we get $\mathcal{M}_2(8n+4) = \mathcal{M}_2(4n+2+x) = x \oplus y \oplus 1$ by a direct case check. In both cases the second coordinate is $x \oplus y \oplus 1$. It then follows that

$$s(4n+2) = (y \oplus 1,\ x \oplus y \oplus 1).$$

Case $r = 3$. The first coordinate is given by $\mathcal{M}_2(4n+3) = 1 - \mathcal{M}_2(4n+2) = 1 - (y \oplus 1) = y$.

For the second coordinate of $s(4n+3)$, we have that $\mathcal{M}_2(2(4n+3) + \mathcal{M}_2(4n+3)) = \mathcal{M}_2(8n+6+y)$. Since $\mathcal{M}_2(2n+1) = x \oplus 1$, the identity $\mathcal{M}_2(8n+6) = \mathcal{M}_2(2(2n+1)+1-\mathcal{M}_2(2n+1))$ gives $\mathcal{M}_2(8n+6) = x \oplus y \oplus 1$ by a direct two-case check on x; this handles $y = 0$. For $y = 1$, since $\mathcal{M}_2(2n+1) = x$, the same formula yields $1 - \mathcal{M}_2(4n+2+x) = x \oplus y \oplus 1$. The second coordinate is $x \oplus y \oplus 1$ in both cases. Consequently, we have that

$$s(4n+3) = (y,\ x \oplus y \oplus 1).$$

Each transition has the form $(x, y) \mapsto A\binom{x}{y} + \mathbf{v}(d)$ over $\mathbb{F}_2$, confirming the affine structure. $\square$

Remark 21. The linear part $(x, y) \mapsto (y, x \oplus y)$ is given by multiplication by the matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{F}_2),$$

which has order 3 (indeed $A^3 = I$). The translation vectors are $\mathbf{v}(0) = \binom{0}{0}$, $\mathbf{v}(1) = \binom{1}{0}$, $\mathbf{v}(2) = \binom{1}{1}$, $\mathbf{v}(3) = \binom{0}{1}$.

Theorem 22. *The sequence $\mathcal{M}_2$ is computed (MSB-first, base 4) by the DFAO with states $\{0, 1, 2, 3\}$, initial state 0, output $\tau(0) = \tau(2) = 0$, $\tau(1) = \tau(3) = 1$, and transition table*

δ	0	1	2	3
0	0	1	3	2
1	2	3	1	0
2	3	2	0	1
3	1	0	2	3

where the encoding is $\mathsf{0} = (0, 0)$, $\mathsf{1} = (1, 0)$, $\mathsf{2} = (0, 1)$, $\mathsf{3} = (1, 1)$.

Proof. This follows directly from the affine transitions in Theorem 20 applied to each encoded state. The initial state is $s(0) = (0, 0)$, corresponding to state $\mathsf{0}$. $\square$

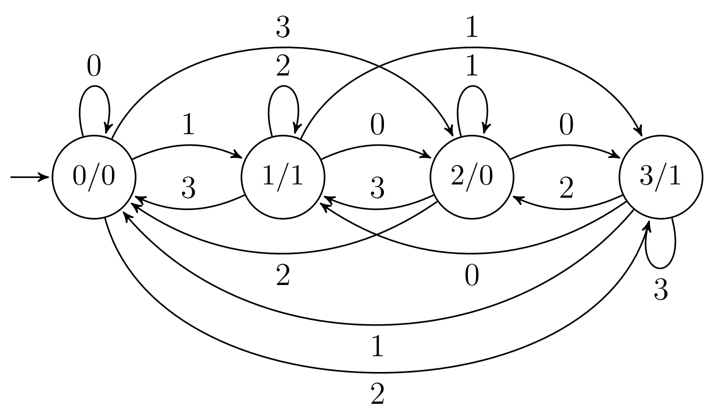

Figure 3: DFAO for $\mathcal{M}_2$ (4 states, base 4, MSB-first). State labels are *state/output*. The initial state is $\mathsf{0}$.

Corollary 23. *The coding $\tau \colon \mathsf{0}, \mathsf{2} \mapsto 0$ and $\mathsf{1}, \mathsf{3} \mapsto 1$ applied to the fixed point of the 4-uniform morphism*

$$\mathsf{0} \to \mathsf{0132}, \quad \mathsf{1} \to \mathsf{2310}, \quad \mathsf{2} \to \mathsf{3201}, \quad \mathsf{3} \to \mathsf{1023}$$

starting from $\mathsf{0}$ produces the sequence $\mathcal{M}_2$.

Remark 24. The four letters of the alphabet correspond to the states $(x, y) \in \mathbb{F}_2^2$ of the affine automaton (Theorem 20) via

$$\mathsf{0} = (0,0), \quad \mathsf{1} = (1,0), \quad \mathsf{2} = (0,1), \quad \mathsf{3} = (1,1).$$

With this identification, the image $\sigma_{\mathcal{M}_2}(\mathsf{r})$ lists the four successive states visited when the current state is r and the digits $d \in \{0,1,2,3\}$ are read in order, and the coding $\tau(\mathsf{r})$ is the first coordinate of the state. Under the relabeling $a = \mathsf{0}$, $b = \mathsf{2}$, $c = \mathsf{1}$, $d = \mathsf{3}$, the morphism reads $a \to acdb$, $b \to dbac$, $c \to bdca$, $d \to cabd$, with $\tau(a) = \tau(b) = 0$ and $\tau(c) = \tau(d) = 1$. This labeling is used in the right-special factor analysis below.

The defining recurrence of $\mathcal{M}_2$ is not directly an instance of Theorem 2, since the argument $2n + \mathcal{M}_2(n)$ in the nested recurrence in (13) and the argument $2n+1-\mathcal{M}_2(n)$ in (14) depend on the value $\mathcal{M}_2(n)$. This together with the sequence $\mathcal{M}_2(2n + \mathcal{M}_2(n))$ used in the pair (15) yields the system of recurrences in Theorem 20. Furthermore, Lemma 5 gives us that

$$\mathcal{M}_2(4n) = \mathcal{M}_2(2n) \oplus \mathcal{M}_2(n), \qquad \mathcal{M}_2(4n+2) = \mathcal{M}_2(2n+1) \oplus \mathcal{M}_2(n).$$

The XOR identities are direct consequences of Lemma 5, since $\mathcal{M}_2$ is balanced.

4.1 An explicit digit formula

Write the base-4 expansion $n = \sum_{k \geq 0} d_k 4^k$ with digits $d_k \in \{0,1,2,3\}$, and write $d_k = 2j_k + i_k$ with $i_k, j_k \in \{0,1\}$.

Theorem 25. *We have that*

$$\mathcal{M}_2(n) = \bigoplus_{k \equiv 0 \pmod 3} (i_k \oplus j_k) \;\oplus \bigoplus_{k \equiv 1 \pmod 3} j_k \;\oplus \bigoplus_{k \equiv 2 \pmod 3} i_k.$$

Proof. Let $d_{L-1} \cdots d_1 d_0$ be the base-4 expansion of n. Since the automaton reads digits MSB-first and the transition has the form $s \mapsto As + \mathbf{v}(d)$, starting from $s(0) = (0,0)$ gives

$$s(n) = \sum_{k=0}^{L-1} A^k \, \mathbf{v}(d_k) \pmod 2.$$

Since $A^3 = I$, the exponent depends only on $k \bmod 3$. Computing $A^k \mathbf{v}(d)$ for $k \in \{0,1,2\}$ and each digit d, we obtain

d	$A^0\mathbf{v}(d)$	$A^1\mathbf{v}(d)$	$A^2\mathbf{v}(d)$
0	$(0,0)$	$(0,0)$	$(0,0)$
1	$(1,0)$	$(0,1)$	$(1,1)$
2	$(1,1)$	$(1,0)$	$(0,1)$
3	$(0,1)$	$(1,1)$	$(1,0)$

Extracting the first coordinate and identifying $i_k = d_k \bmod 2$ and $j_k = \lfloor d_k/2 \rfloor$ yields the stated formula. ∎

Theorem 26. *Write the binary expansion* $n = \sum_{m \geq 0} b_m 2^m$ *with* $b_m \in \{0, 1\}$, *and define*

$$q(n) := \sum_{\substack{m \geq 0 \\ m \not\equiv 2 \pmod 3}} b_m 2^m, \tag{16}$$

so that $q(n)$ *is obtained from* n *by zeroing the binary digits in positions* $m \equiv 2 \pmod 3$. *Then*

$$\mathcal{M}_2(n) = \mathbf{t}(q(n)) \qquad (n \geq 0).$$

Equivalently, if $n = \sum_{k \geq 0} d_k 4^k$ *with* $d_k = 2j_k + i_k$, *then* $q(n) = \sum_{k \geq 0} d'_k 4^k$ *where*

$$d'_k = \begin{cases} d_k, & \textit{if } k \equiv 0 \pmod 3; \\ 2j_k, & \textit{if } k \equiv 1 \pmod 3; \\ i_k, & \textit{otherwise.} \end{cases}$$

and $\mathcal{M}_2(n) = \mathbf{t}(q(n))$.

Proof. In Theorem 25, each base-4 digit d_k contributes to $\mathcal{M}_2(n)$ in a way that depends only on $k \bmod 3$: For $k \equiv 0 \bmod 3$, we take $i_k \oplus j_k$, for $k \equiv 1 \bmod 3$, we take only j_k, and for $k \equiv 2 \bmod 3$, we take i_k. Since $i_k = b_{2k}$ and $j_k = b_{2k+1}$, the binary positions $m = 2k$ and $m = 2k+1$ are kept or omitted, according to the residue class $k \bmod 3$, in the following manner. For $k \equiv 0 \pmod 3$, both positions are kept. For $k \equiv 1 \pmod 3$, the position $m = 2k$ is omitted. For $k \equiv 2 \pmod 3$, the position $m = 2k + 1$ is omitted. In each of the cases where a position is omitted, the omitted position m satisfies $m \equiv 2 \pmod 3$. Conversely, every value m such that $m \equiv 2 \pmod 3$ falls into one of the two specified cases. Thus, the omitted positions are precisely those such that $m \equiv 2 \pmod 3$, and $\mathcal{M}_2(n)$ is the parity of the binary digit-sum of $q(n)$ defined in (16), and hence $\mathcal{M}_2(n) = \mathbf{t}(q(n))$. This gives us, in an equivalent way, the desired base-4 description of $q(n)$. $\square$

5 Factor complexity

Let $p_a(n)$ denote the number of distinct length-n factors in the infinite word $a(0)a(1)a(2)\cdots$. A length-n factor u of a is called *right-special* if both $u0$ and $u1$ are factors of a. We write $\mathrm{RS}_a(n)$ in place of the set of right-special factors of a of length n.

Lemma 27. *For every infinite binary word* a *and every* $n \geq 1$, *we have that*

$$p_a(n+1) - p_a(n) = \#\,\mathrm{RS}_a(n).$$

Proof. For each length-n factor u, let $E_R(u) \subseteq \{0, 1\}$ be the set of right extensions such that $u\alpha$ occurs in a. Then $p_a(n+1) = \sum_u |E_R(u)|$ and $p_a(n) = \sum_u 1$, so $p_a(n+1) - p_a(n) = \sum_u (|E_R(u)| - 1)$. In the binary case, $|E_R(u)| - 1$ equals 1 if u is right-special and 0 otherwise. $\square$

Lemma 28. *If a is k-automatic, then the function $n \mapsto \#\mathrm{RS}_a(n)$ is k-regular in the sense of Allouche-Shallit [3], and is effectively computable from any DFAO for a. Moreover, if $\#\mathrm{RS}_a(n)$ is bounded, then it is k-automatic.*

Proof. The predicate "u is a length-n factor of a" and the predicate "u is right-special" are first-order definable over $(\mathbb{N}, +, V_k)$ relative to a DFAO for a, so the set of encodings of such pairs is regular. Counting right-special factors of a given length reduces to a finite-automaton computation, giving a k-regular sequence [3, Ch. 16]. A bounded k-regular sequence is k-automatic [3, Theorem 16.1.5]. □

Several statements in this paper can be certified by the `Walnut` theorem prover [29], which decides first-order properties of automatic sequences in Presburger arithmetic. This applies to the defining recurrences and the balance of $\mathcal{Q}$, $\mathcal{M}_1$, and $\mathcal{M}_2$ (Sections 3 and 4), and (using Lemma 28) to the right-special factor structure of $\mathcal{Q}$ and $\mathcal{M}_2$ in the present section. The right-special factor verifications for $\mathcal{M}_2$ use the CCL(S) determinization algorithm of Nicol and Frohme [25]. We give arguments that do not require Walnut wherever this is convenient.

5.1 Factor complexity of $\mathcal{Q}$

We first state a property of the factors of $\mathbf{t}$.

Lemma 29. *The set $\mathrm{Fac}(\mathbf{t})$ is closed under bitwise complementation: If $w \in \mathrm{Fac}(\mathbf{t})$, then $\bar{w} \in \mathrm{Fac}(\mathbf{t})$.*

Proof. The Thue-Morse morphism $\mu : 0 \mapsto 01,\ 1 \mapsto 10$ satisfies $\overline{\mu(w)} = \mu(\bar{w})$ for every binary word w. Since $\mathbf{t} = \mu^\omega(0)$ and both 0 and $1 = \bar{0}$ appear in $\mathbf{t}$, every factor w of $\mu^\omega(0)$ yields $\bar{w} = \overline{\mu^\omega(0)}|_{\text{same position}} = \mu^\omega(1)|_{\text{same position}}$, which is a factor of $\mu^\omega(1) = 1\,\mathbf{t}[1\,\infty)$, hence a factor of $\mathbf{t}$. □

Proposition 30. *Let J_n denote the set of length-n factors of $\mathcal{Q}$ that are not factors of $\mathbf{t}$. Then*

$$p_{\mathcal{Q}}(n) = p_{\mathbf{t}}(n) + |J_n|.$$

Proof. By Proposition 8, on each interval $[2^k, 2^{k+1})$, the sequence $\mathcal{Q}$ coincides with $\mathbf{t}$ or $1-\mathbf{t}$, according to the parity of k. The cases give us that a factor of $\mathcal{Q}$ of length n is either a factor of $\mathbf{t}$ or the complement of a factor of $\mathbf{t}$, which, by Lemma 29, is again a factor of $\mathbf{t}$. In any case, such factors are in $\mathrm{Fac}(\mathbf{t})$.

For an interval of the form $[2^k, 2^{k+1})$, suppose that the indices associated with a factor of length n properly contains this interval. This can only occur when its left endpoint is at a position $< 2^k$ and its right endpoint is at a position $\geq 2^{k+1}$ for some k, which forces $n \geq 2^k$. In every case, such a factor has the form $u_0\bar{u}_1u_2\bar{u}_3\cdots$ obtained by concatenating consecutive pieces of $\mathbf{t}$ with alternating bitwise complementation, where the concatenation $u_0u_1u_2u_3\cdots$ is a factor of $\mathbf{t}$. By Lemma 29, the complementation pattern preserves membership in $\mathrm{Fac}(\mathbf{t})$ for the simplest one-boundary case $u_0\bar{u}_1$ if and only if $\bar{u}_0u_1 \in \mathrm{Fac}(\mathbf{t})$, and this need not hold in general. According to the above definition of J_n, we have that J_n as the set of length-n factors that have at least one index strictly less than 2^k and at least one index greater than or equal to 2^k and that are not in $\mathrm{Fac}(\mathbf{t})$.

Since every factor of $\mathbf{t}$ appears in $\mathcal{Q}$, the set of factors of $\mathcal{Q}$ of length n is exactly $\mathrm{Fac}(\mathbf{t}) \cap \{0,1\}^n \cup J_n$, giving us that $p_{\mathcal{Q}}(n) = p_{\mathbf{t}}(n) + |J_n|$. □

Theorem 31. *For all $k \geq 4$, we have that*

$$p_{\mathcal{Q}}(2^k) = 13 \cdot 2^{k-2} - 2.$$

Equivalently, $p_{\mathcal{Q}}(2^k) = p_{\mathbf{t}}(2^k) + 2^{k-2}$, so $|J_{2^k}| = 2^{k-2}$.

Proof. The value $p_{\mathbf{t}}(2^k) = 3 \cdot 2^k - 2$ is known [6, 24]. By Proposition 30, we have that $p_{\mathcal{Q}}(2^k) = p_{\mathbf{t}}(2^k) + |J_{2^k}|$. Also, Lemma 27 gives us that

$$\#\mathrm{RS}_{\mathcal{Q}}(n) = p_{\mathcal{Q}}(n+1) - p_{\mathcal{Q}}(n). \tag{17}$$

The right-special count in (17) takes values in $\{2,3,4\}$, and this is certified for all $n \geq 2$ by Walnut. The desired result follows from the recurrence

$$p_{\mathcal{Q}}(2^{k+2}) = 4\,p_{\mathcal{Q}}(2^k) + 6 \qquad (k \geq 4),$$

with $p_{\mathcal{Q}}(16) = 50$, which has the unique solution $p_{\mathcal{Q}}(2^k) = 13 \cdot 2^{k-2} - 2$, also verified computationally for $4 \leq k \leq 7$. □

Remark 32. The set J_{2^k} has size 2^{k-2}, so factors in J_{2^k} account for one quarter of the length-2^k factors of $\mathcal{Q}$ beyond those of $\mathbf{t}$.

Since $\#\mathrm{RS}_{\mathcal{Q}}(n) \in \{2,3,4\}$ for all $n \geq 1$, the values of $p_{\mathcal{Q}}$ are determined by the support of each slope.

Lemma 27 is implicit in the following Proposition, according to (18).

Proposition 33. *For $n \geq 16$, set $k = \lfloor \log_2 n \rfloor$ and write $n = 2^k + i$ with $0 \leq i < 2^k$. Then*

$$p_{\mathcal{Q}}(n) = \begin{cases} 13 \cdot 2^{k-2} - 2, & \text{if } i = 0; \\ 4n - 3 \cdot 2^{k-2} - 4, & \text{if } 1 \leq i \leq 2^{k-1}; \\ 3n + 3 \cdot 2^{k-2} - 3, & \text{if } 2^{k-1} < i \leq 3 \cdot 2^{k-2}; \\ 2n + 5 \cdot 2^{k-1} - 2, & \text{if } 3 \cdot 2^{k-2} < i < 2^k. \end{cases}$$

Moreover, we have that

$$\#\mathrm{RS}_{\mathcal{Q}}(n) = p_{\mathcal{Q}}(n+1) - p_{\mathcal{Q}}(n). \tag{18}$$

The left-hand side of (18) equals 4 for $n \in (2^k,\, 3 \cdot 2^{k-1}]$, and 3 for $n \in (3 \cdot 2^{k-1},\, 7 \cdot 2^{k-2}]$, and 2 for $n \in (7 \cdot 2^{k-2},\, 2^{k+1})$.

Proof. This is certified by Walnut. □

By Proposition 33, $n + 1 \leq p_{\mathcal{Q}}(n) = O(n)$ for all $n \geq 1$.

5.2 Factor complexity of $\mathcal{M}_1$ and $\mathcal{M}_2$

Proposition 34. *We have $p_{\mathcal{M}_1}(n) = \Theta(n)$ and $p_{\mathcal{M}_2}(n) = \Theta(n)$.*

Proof. The incidence matrices of the morphisms in Corollaries 16 and 23 are

$$M_{\mathcal{M}_1} = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \end{pmatrix}, \qquad M_{\mathcal{M}_2} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}. \tag{19}$$

The incidence matrices in (19) give us that both of the associated morphisms are primitive. A result of Pansiot [26] gives us that the fixed point of a primitive k-uniform morphism has factor complexity $O(n)$, and the same upper bound holds after applying a coding. Neither $\mathcal{M}_1$ nor $\mathcal{M}_2$ is ultimately periodic, as can be checked from the DFAOs of Theorems 13 and 22, or certified by Walnut. By the Morse–Hedlund theorem [3, Theorem 10.2.6], the factor complexity of a non-ultimately-periodic word is at least $n+1$ for every $n \geq 0$. Hence both complexities are $\Theta(n)$. $\square$

An explicit evaluation of the factor complexity of $\mathcal{M}_2$ is provided below.

Example 35. We obtain the initial values

$$p_{\mathcal{M}_1}(1 \ldots 15) = (2, 4, 6, 9, 12, 17, 22, 28, 34, 40, 46, 52, 58, 64, 70),$$

and

$$p_{\mathcal{M}_2}(1 \ldots 15) = (2, 4, 6, 10, 12, 14, 16, 18, 20, 24, 28, 32, 36, 40, 44).$$

Proposition 36. *For all $k \geq 3$, we have that*

$$p_{\mathcal{M}_2}(2^k) = \begin{cases} \frac{5}{2} \cdot 2^k - 2, & \textit{if } 3 \mid k; \\ 3 \cdot 2^k - 2, & \textit{otherwise.} \end{cases}$$

Moreover, the recurrence

$$p_{\mathcal{M}_2}(2^{k+3}) = 8\, p_{\mathcal{M}_2}(2^k) + 14 \qquad (k \geq 3)$$

holds true.

Proof. This is certified by Walnut. $\square$

Since $\#\mathrm{RS}_{\mathcal{M}_2}(n) \in \{2, 4\}$, the values of $p_{\mathcal{M}_2}$ are determined by the positions at which $\#\mathrm{RS}_{\mathcal{M}_2}(n) = 4$. The following proposition extends Proposition 36 to a complete evaluation.

Proposition 37. *We have* $p_{\mathcal{M}_2}(1) = 2$. *For* $n \geq 2$, *set* $k = \lfloor \log_2 n \rfloor$ *and* $r = k \bmod 3$.

Case $r = 0$: *On the interval* $[2^k, 2^{k+1}]$, *we have that*

$$p_{\mathcal{M}_2}(n) = \begin{cases} 2n + (2^{k-1} - 2), & \text{if } 2^k \leq n \leq 2^k + 1; \\ 4n - (3 \cdot 2^{k-1} + 4), & \text{if } 2^k + 1 \leq n \leq 7 \cdot 2^{k-2} + 1; \\ 2n + (2^{k+1} - 2), & \text{if } 7 \cdot 2^{k-2} + 1 \leq n \leq 2^{k+1}. \end{cases}$$

Case $r = 1$: *On the interval* $[2^k, 2^{k+1}]$, *we have that*

$$p_{\mathcal{M}_2}(n) = \begin{cases} 2n + (2^{k} - 2), & \text{if } 2^k \leq n \leq 2^k + 1; \\ 4n - (2^{k} + 4), & \text{if } 2^k + 1 \leq n \leq 3 \cdot 2^{k-1} + 1; \\ 2n + (2^{k+1} - 2), & \text{if } 3 \cdot 2^{k-1} + 1 \leq n \leq 2^{k+1}. \end{cases}$$

Case $r = 2$: *On the interval* $[2^k, 2^{k+1}]$, *we have that*

$$p_{\mathcal{M}_2}(n) = 2n + (2^k - 2).$$

Proof. This is certified by Walnut. ∎

6 Conclusion

Consider replacing $a(2n) + a(2n+1) = 1$ by $a(3n) + a(3n+1) + a(3n+2) = 2$, with a taking values in $\{0, 1\}$. This gives rise to expressions of the form $a(3n + r - a(n))$, and the analogue of Lemma 5 converts each such expression into a fixed-argument expression. A small number of initial configurations produce 3-automatic binary sequences with small DFAOs; we plan to treat this family separately.

7 Acknowledgments

The authors thank the anonymous referee for a careful reading of the manuscript and for helpful suggestions. The authors also thank Gandhar Joshi for independently testing the Walnut files. Claude (Anthropic) was used as a computational assistant, in particular for exploratory computations, for suggesting candidate formulas later checked by independent scripts, and for checking consistency between the manuscript, the Python script, and the Walnut files. The mathematical content, the proofs, and the final wording are the authors' own. The first author thanks Jean-Paul Allouche, Karl Dilcher, and Jeffrey Shallit for useful feedback related to this paper.

8 Minimal DFAO sizes

Minimal DFAO sizes for the meta-automatic sequences considered in this paper are listed in Table 1.

$F(n)$	$G(n)$	DFAO	Nested	Notes
$a(n)$	$a(n)$	2	0	Thue-Morse
$a(n)$	$a(2n)$	**3**	**0**	$\mathcal{Q}$ (Thue-Morse Quarto)
$a(n)$	$a(2n+1)$	2	0	Thue-Morse
$a(n)$	$a(2n+a(n))$	4	1	
$a(n)$	$a(2n+1-a(n))$	**4**	**1**	$\mathcal{M}_1$
$a(2n)$	$a(n)$	4	0	
$a(2n)$	$a(2n)$	—	0	ultimately periodic
$a(2n)$	$a(2n+1)$	2	0	Thue-Morse
$a(2n)$	$a(2n+a(n))$	3	1	
$a(2n)$	$a(2n+1-a(n))$	3	1	
$a(2n+1)$	$a(n)$	5	0	
$a(2n+1)$	$a(2n)$	3	0	complement of $\mathcal{Q}$
$a(2n+1)$	$a(2n+1)$	5	0	
$a(2n+1)$	$a(2n+a(n))$	5	1	
$a(2n+1)$	$a(2n+1-a(n))$	5	1	
$a(2n+a(n))$	$a(n)$	4	1	
$a(2n+a(n))$	$a(2n)$	4	1	
$a(2n+a(n))$	$a(2n+1)$	4	1	
$a(2n+a(n))$	$a(2n+a(n))$	4	2	
$a(2n+a(n))$	$a(2n+1-a(n))$	**4**	**2**	$\mathcal{M}_2$
$a(2n+1-a(n))$	$a(n)$	5	1	
$a(2n+1-a(n))$	$a(2n)$	4	1	
$a(2n+1-a(n))$	$a(2n+1)$	5	1	
$a(2n+1-a(n))$	$a(2n+a(n))$	5	2	
$a(2n+1-a(n))$	$a(2n+1-a(n))$	5	2	

Table 1: Minimal DFAO size for each specialization. The column "Nested" counts the number of branch rules (F and G) whose right-hand side contains an expression of the form $a(\,\cdot \pm a(n))$. Boldface marks the three sequences studied in this paper.

References

[1] A. Alkan, On a generalization of Hofstadter's Q-sequence: a family of chaotic generational structures, *Complexity* **2018** (2018), Article ID 8517125.

[2] A. Alkan, N. Fox, and O. O. Aybar, On Hofstadter heart sequences, *Complexity* (2017), Article ID 2614163.

[3] J.-P. Allouche and J. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press, 2003.

[4] J.-P. Allouche and J. Shallit, A variant of Hofstadter's sequence and finite automata, *J. Aust. Math. Soc.* **93** (2012), 1–8.

[5] B. Balamohan, A. Kuznetsov, and S. Tanny, On the behavior of a variant of Hofstadter's Q-sequence, *J. Integer Sequences* **10** (2007), Article 07.7.1.

[6] S. Brlek, Enumeration of factors in the Thue-Morse word, *Discrete Appl. Math.* **24** (1989), 83–96.

[7] M. Cai and S. M. Tanny, How the shift parameter affects the behavior of a family of meta-Fibonacci sequences, *J. Integer Sequences* **11** (2008), Article 08.3.6.

[8] J. Callaghan, J. J. Chew, III, and S. Tanny, On the behavior of a family of meta-Fibonacci sequences, *SIAM J. Discrete Math.* **18** (2005), 794–824.

[9] A. Cobham, Uniform tag sequences, *Math. Systems Theory* **6** (1972), 164–192.

[10] B. Dalton, M. Rahman, and S. Tanny, Spot-based generations for meta-Fibonacci sequences, *Exp. Math.* **20** (2011), 129–137.

[11] C. Deugau and F. Ruskey, Complete k-ary trees and generalized meta-Fibonacci sequences, *Fourth Colloquium on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities*, September 18–22, 2006, Institut Ilie Cartan, Nancy, France, *Discrete Mathematics and Theoretical Computer Science (DMTCS) Proceedings Series Volume* **AG** (2006), 203–214.

[12] F. M. Dekking, On Hofstadter's G-sequence, *J. Integer Sequences* **26** (2023), Article 23.9.2.

[13] N. D. Emerson, A family of meta-Fibonacci sequences defined by variable-order recursions, *J. Integer Sequences* **9** (2006), Article 06.1.8.

[14] N. Fox, The behavior of a three-term Hofstadter-like recurrence with linear initial conditions, *J. Integer Sequences* **27** (2024), Article 24.7.4.

[15] N. Fox, Linear recurrent subsequences of generalized meta-Fibonacci sequences, *J. Difference Equ. Appl.* **22** (2016), 1019–1026.

[16] N. Fox, A new approach to the Hofstadter Q-recurrence, *Integers* **20A** (2020), A8.

[17] N. Fox, Quasipolynomial solutions to the Hofstadter Q-recurrence, *Integers* **16** (2016), A68.

[18] S. W. Golomb, Discrete chaos: sequences satisfying "strange" recursions, unpublished manuscript, c. 1990. Cached copy available at https://oeis.org/A005185/a005185_1.pdf.

[19] R. J. Hendel, Quasi-periods for the Hofstadter Q function, *Fibonacci Quart.* **53** (2015), 112–123.

[20] D. R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*, Basic Books, 1979.

[21] A. Isgur and M. Rahman, On variants of Conway and Conolly's meta-Fibonacci recursions, *Electron. J. Combin.* **18** (2011), Paper 96.

[22] A. Isgur, D. Reiss, and S. Tanny, Trees and meta-Fibonacci sequences, *Electron. J. Combin.* **16** (2009), Research Paper 129, 40.

[23] B. Jackson and F. Ruskey, Meta-Fibonacci sequences, binary trees and extremal compact codes, *Electron. J. Combin.* **13** (2006), Research Paper 26, 13.

[24] A. de Luca and S. Varricchio, Some combinatorial properties of the Thue-Morse sequence and a problem in semigroups, *Theoret. Comput. Sci.* **63** (1989), 333–348.

[25] J. Nicol and M. Frohme, Deconstructing subset construction: reducing while determinizing, in *Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2026)*, Part II, Lect. Notes in Comp. Sci., Vol. 16506, Springer, 2026.

[26] J.-J. Pansiot, Complexité des facteurs des mots infinis engendrés par morphismes itérés, in *Automata, Languages and Programming: Proc. ICALP 1984*, Lect. Notes in Comp. Sci., Vol. 172, Springer, 1984, pp. 380–389.

[27] K. Pinn, Order and chaos in Hofstadter's $Q(n)$ sequence, *Complexity* **4** (1999), 41–46.

[28] F. Ruskey and C. Deugau, The combinatorics of certain k-ary meta-Fibonacci sequences, *J. Integer Sequences* **12** (2009), Article 09.4.3.

[29] J. Shallit, *The Logical Approach to Automatic Sequences*, London Math. Soc. Lecture Note Ser. Vol. 482, Cambridge University Press, 2022.

[30] N. J. A. Sloane et al., The On-Line Encyclopedia of Integer Sequences, 2025. Available at https://oeis.org.

[31] B. Sobolewski and M. Ulas, Solutions of certain meta-Fibonacci recurrences, *Acta Arith.* **209** (2023), 269–289.

[32] S. M. Tanny, A well-behaved cousin of the Hofstadter sequence, *Discrete Math.* **105** (1992), 227–239.

2020 *Mathematics Subject Classification*: Primary 11B85; Secondary 68Q45, 68R15.
Keywords: automatic sequence, nested recurrence, Thue-Morse sequence, DFAO, binary sequence.

(Concerned with sequences A005185, A039982, A244477, A298952, A391614, and A392736.)

Received March 2 2026; revised version received May 26 2026. Published in *Journal of Integer Sequences*, June 4 2026.

Return to Journal of Integer Sequences home page.