hexagon:2611.00027v1

Powers of a Prime in the Image of Summatory Valuation Functions

Frank Vega

Information Physics Institute, 840 W 67th St, Hialeah, FL 33012, USA

Abstract

J.-L. Nicolas asked how often a power of 2 belongs to the image of $c(n) = \sum_{k \le n} 2^{v_2(k)}$, where v_2 is the 2-adic valuation, and showed that both the number $A(x)$ of exponents $a \le x$ with 2^a in the image and the number $B(x)$ of the other exponents tend to infinity. We prove the identity $c(n) = n + \frac{1}{2}\sum_i i\,2^i$, the sum running over the positions of the binary digits 1 of n, so that 2^a lies in the image of c if and only if 2^{a+3} is a sum of distinct terms of the superincreasing sequence $(j2^j)_{j\ge 1}$. For every exponent there is at most one possible largest term K, given by $K + \lceil \log_2 K \rceil = a + 3$, and no such K exists exactly when $a = 2^r + r - 2$. The remaining exponents are decided by an explicit recursion on pairs of integers bounded by $a + 3$, in at most $a + 1$ steps. We give the closed formula $c(2^{K-2} + 2^{K-s-2}) = 2^{K+h-3}$ for $s = 2^k$, $h \equiv k \pmod{2^{k+1}}$ and $K = (2^{h+s} + s)/(2^s + 1)$, and infinite explicit families of exponents a for which 2^a is not in the image, indexed by $s = 3$ and $s = 2^i$; the families given by Nicolas are the cases $s = 1$ and $s = 2$. As a consequence, $A(x) \ge (2 + o(1)) \log_2 x$ and $B(x) \ge (\frac{5}{4} + \frac{1}{2}\sum_{i\ge 1} \frac{1}{2^i-1} + o(1)) \log_2 x = (2.0533\ldots + o(1)) \log_2 x$. We prove that the states of the recursion form a forest, rooted at the states $(J, 0)$ and at the failing states, in which every state (J', w') has exactly $v_2(J' + 1 + w')$ children; that the sets S_J of successful states are symmetric; that exactly half of them lie in the tree of the root $(1, 0)$; that $N(2k + 1) = N(2k)$; that $4 \mid N(J)$ for $J \ge 4$; and that S_J satisfies a recursion from the smallest elements upwards, with period $J/2$ for even J. The symmetry gives further explicit families, among them $c(2^{2^r} - 1) = 2^{2^r+r-1}$, and the improved bounds $A(x) \ge (3.9655\ldots + o(1)) \log_2 x$ and $B(x) \ge (2.8178\ldots + o(1)) \log_2 x$. For every prime p, the function $c_p(n) = \sum_{k\le n} p^{v_p(k)}$ counts the factor $1 - X$ in $(1 - X)\cdots(1 - X^n)$ over $\mathbb{F}_p$; we extend to it the reduction, the unique candidate, the recursions and the symmetry, give explicit families, and describe completely the exponents rejected at the first step, which occur only for odd p. We also report experiments in exact integer arithmetic: every statement is checked by three independent methods, the values of $A(x)$ reproduce the table of Nicolas, the sets S_J are computed for $J \le 131072$, and $A(x)/x$ lies between 0.5132 and 0.8750 for $x = 1000, 2000, \ldots, 131000$.

Keywords: Arithmetic functions, 2-adic valuation, Binary expansion, Superincreasing sequences, Polynomials over $\mathbb{F}_2$, Open questions of Nicolas
2020 MSC: 11A63, 11A25, 11B83, 11N37

1. Introduction

For a positive integer k let $v_2(k)$ denote the exponent of 2 in k. In a joint paper with J. Dixmier, J.-L. Nicolas considered the function

$$c(n) = \sum_{k=1}^{n} 2^{v_2(k)}, \qquad n \ge 1, \tag{1}$$

which is the exponent of $1 + X$ in the factorization of $(1 + X)(1 + X^2)\cdots(1 + X^n)$ into irreducible factors over $\mathbb{F}_2[X]$, and in his list of open questions [1] he asked how often a power of 2 belongs to the image of c. The sequence $(c(n))_{n\ge 1} = 1, 3, 4, 8, 9, 11, 12, 20, \ldots$ of the partial sums of the highest power of 2 dividing k is the sequence M2344 of the encyclopedia of Sloane and Plouffe [2]. It satisfies the recurrence $c(2n) = 2c(n) + n$ and the formulas

Email address: vega.frank@gmail.com (Frank Vega)

$c(2^r) = 2^{r-1}(r+2)$ and $c(2^r-1) = r2^{r-1}$ [1], and the recurrence splits at half, a class of recurrences whose identities and periodic oscillations are studied by Hwang, Janson and Tsai [3]. With

$$\mathcal{A} = \{a \geq 0 :\ 2^a \in c(\mathbb{N})\}, \qquad A(x) = \#\{a \in \mathcal{A} :\ a \leq x\}, \qquad B(x) = \#\{a \in \mathbb{Z} \cap [0,x] :\ a \notin \mathcal{A}\}, \tag{2}$$

Nicolas showed that $A(x) \to \infty$ and $B(x) \to \infty$, and asked whether the behaviour of $A(x)$ can be made precise; his table gives, for instance, the values 635 for $x = 1000$ and 5963 for $x = 10000$ [1], which are the numbers of exponents $1 \leq a \leq x$ in $\mathcal{A}$, that is $A(x) - 1$ in the notation (2), since $c(1) = 2^0$ (Section 10). He obtained these limits from three families: $2^a \notin c(\mathbb{N})$ when $a = 2^r + r - 2$; $a + (2^a - 11)/3 \in \mathcal{A}$ for odd a; and $a + (2^a - 10)/3 \notin \mathcal{A}$ for even a [1].

The same list contains a question on a family of fractions, namely the size of the largest numerator $A(N)$ of $\prod_{n \leq N}((n+1)/n)^{\varepsilon(n)}$ with $\varepsilon(n) \in \{-1, 0, 1\}$; for that question de la Bretèche, Pomerance and Tenenbaum proved $\log A(N) = \{\kappa + o(1)\}N \log N$ with an explicit constant $\kappa \approx 0.107005$ [4]. The present paper is concerned with the question on $c(n)$ only.

Contributions. We prove the following results.

- The identity $c(n) = n + \frac{1}{2}\sum_{i \in \mathcal{B}(n)} i\, 2^i$, where $\mathcal{B}(n)$ is the set of positions of the binary digits 1 of n, equivalently $c(n) = n + \sum_{k \geq 1} 2^{k-1}\lfloor n/2^k \rfloor$, and the equivalence $a \in \mathcal{A} \iff 2^{a+3} = \sum_{j \in T} j\, 2^j$ for a finite set T of positive integers (Lemma 2.1, Theorem 2.2).

- The numbers $j2^j$ form a superincreasing sequence. Consequently, for every $a \geq 0$ there is at most one integer $K \geq 2$ that can be the largest element of T, namely the solution of $K + \lceil \log_2 K \rceil = a + 3$, and such a K fails to exist exactly for $a = 2^r + r - 2$, $r \geq 1$ (Theorem 3.3). This recovers the first family of Nicolas and shows that it is the only obstruction of this kind.

- A recursion on pairs of integers bounded by $a + 3$ that decides whether $a \in \mathcal{A}$ in at most $a + 1$ steps, without handling integers of size 2^a (Theorem 4.1, Corollary 4.2).

- Infinite explicit families of members of $\mathcal{A}$, given by the closed formula $c(2^{K-2} + 2^{K-s-2}) = 2^{K+h-3}$ for $s = 2^k$, $k \geq 0$, $h \equiv k \pmod{2^{k+1}}$ and $K = (2^{h+s} + s)/(2^s + 1)$ (Theorem 5.1), and infinite explicit families of non-members, indexed by $s = 3$ and by $s = 2^i$, $i \geq 1$ (Theorem 6.1). The second and third families of Nicolas are the cases $s = 1$ of Theorem 5.1 and $s = 2$ of Theorem 6.1. These families are complete in the following sense: they are exactly the exponents that Algorithm 1 decides in at most one iteration (Theorem 6.2).

- Explicit lower bounds for both counting functions (Theorem 7.1):

$$A(x) \geq (2 + o(1)) \log_2 x, \qquad B(x) \geq \Big(\frac{5}{4} + \frac{E}{2} + o(1)\Big) \log_2 x, \qquad E = \sum_{i \geq 1} \frac{1}{2^i - 1} = 1.6066951\ldots,$$

  so that $\frac{5}{4} + \frac{E}{2} = 2.0533475\ldots$.

- The states of the recursion form a forest: the roots are the states $(J, 0)$, one at each level, and the failing states, whose number up to level L is given by an explicit formula; each state (J', w') has exactly $v_2(J' + 1 + w')$ children; and the initial states are the states with $J + 1 + w$ a power of 2 (Theorem 8.4). The root of the tree of a successful state is determined by the least element of its representing set, and exactly half of the successful states of each level lie in the tree of the root $(1, 0)$ (Proposition 8.5). The sets S_J satisfy the recursion $S_J = \{c/2 :\ c \in S_{J-1} \text{ even}\} \cup \{(c + J)/2 :\ c \in S_{J-1} \text{ odd}\}$ for odd J and $S_J = \{c/2,\ c/2 + J/2 :\ c \in S_{J-1} \text{ even}\}$ for even J; in particular S_J is periodic with period $J/2$ for even J (Theorem 8.6). The symmetry yields three further explicit families, among them $c(2^{2^r} - 1) = 2^{2^r + r - 1}$ (Theorem 8.7), and the improved bounds $A(x) \geq (3.9655\ldots + o(1)) \log_2 x$ and $B(x) \geq (2.8178\ldots + o(1)) \log_2 x$ (Theorem 8.8). The sets S_J of successful states, the states in the trees of the roots $(J, 0)$: the number $N(J) = |S_J|$ counts the sets $T \subseteq \{1, \ldots, J\}$ with $\sum_{j \in T} j2^j \equiv 0 \pmod{2^{J+1}}$, S_J is symmetric under $w \mapsto J - 1 - w$, $N(2k + 1) = N(2k)$, and $4 \mid N(J)$ for $J \geq 4$ (Section 8).

- The analogue for every prime p (Section 9), with the complete description of the immediate failures, the bottom-up recursion and the symmetry $w \mapsto (p-1)K - w$ (Theorems 9.6 and 9.7): $c_p(n) = \sum_{k \le n} p^{v_p(k)}$ is the exponent of $1 - X$ in $(1-X)\cdots(1-X^n)$ over $\mathbb{F}_p[X]$; the reduction to digit sums of the numbers $p^i((p-1)i + p)$, the unique candidate, the exponents $(p^g - 1)/(p-1) + g - 1$ without candidate, the recursion, and explicit families, among them a family of non-members with no analogue for $p = 2$ (Theorems 9.3, 9.4 and 9.5).

- Experiments (Section 10) with exact integer arithmetic: a verification of the statements above by three independent methods, among them a brute-force search over $n \le 2^{24}$; the values of $A(x)$ and $B(x)$ for $x \le 20000$, which agree with the table of Nicolas; the sets S_J for $J \le 65536$, from which $A(x)$ is obtained for $x \le 65000$; and the proportion of exponents in $\mathcal{A}$ among the candidates $K \in (2^{h-1}, 2^h]$ for $h \le 16$.

Section 2 proves the identity and the equivalence, Section 3 treats the largest element of T, Section 4 gives the recursion, Sections 5 and 6 give the explicit families, Section 7 proves the counting bounds, Section 8 studies the successful states, Section 9 treats an arbitrary prime, Section 10 reports the experiments and Section 11 concludes.

Notation. $\mathbb{N} = \{1, 2, 3, \dots\}$. For $n \ge 1$, $\mathcal{B}(n) \subset \{0, 1, 2, \dots\}$ is the set of positions of the digits 1 in the binary expansion of n, so that $n = \sum_{i \in \mathcal{B}(n)} 2^i$. For $J \ge 0$ we put

$$\mathcal{E}_J = \Big\{ \sum_{j \in T} j\, 2^j : \ T \subseteq \{1, \dots, J\} \Big\}, \qquad \mathcal{E} = \bigcup_{J \ge 0} \mathcal{E}_J, \tag{3}$$

with $\mathcal{E}_0 = \{0\}$ (empty sum).

2. The function $c(n)$ and sums of $j2^j$

Lemma 2.1. *For every $r \ge 0$, $c(2^r) = 2^{r-1}(r+2)$, and for every $n \ge 1$,*

$$c(n) = \sum_{i \in \mathcal{B}(n)} 2^{i-1}(i+2) = n + \frac{1}{2} \sum_{i \in \mathcal{B}(n)} i\, 2^i. \tag{4}$$

Proof. For $r \ge 0$, the number of $k \in [1, 2^r - 1]$ with $v_2(k) = v$ is 2^{r-1-v} for $0 \le v \le r-1$, so $c(2^r - 1) = \sum_{v=0}^{r-1} 2^v 2^{r-1-v} = r2^{r-1}$ and $c(2^r) = r2^{r-1} + 2^r = 2^{r-1}(r+2)$. If $m \ge 0$ and $1 \le s < 2^m$, then $v_2(2^m + k) = v_2(k)$ for $1 \le k \le s$, hence $c(2^m + s) = c(2^m) + c(s)$. Writing $n = 2^{n_1} + \cdots + 2^{n_t}$ with $n_1 > \cdots > n_t$ and applying this relation $t - 1$ times gives $c(n) = \sum_{l=1}^{t} c(2^{n_l}) = \sum_{i \in \mathcal{B}(n)} 2^{i-1}(i+2)$. Finally $2^{i-1}(i+2) = \frac{1}{2} i 2^i + 2^i$ and $\sum_{i \in \mathcal{B}(n)} 2^i = n$. □

The first equality of (4) is formula (7.4) of [1]. The second form is equivalent to $c(n) = n + \sum_{k \ge 1} 2^{k-1} \lfloor n/2^k \rfloor$: since $\lfloor n/2^k \rfloor = \sum_{i \in \mathcal{B}(n),\, i \ge k} 2^{i-k}$,

$$\sum_{k \ge 1} 2^{k-1} \Big\lfloor \frac{n}{2^k} \Big\rfloor = \sum_{i \in \mathcal{B}(n)} \sum_{k=1}^{i} 2^{k-1} 2^{i-k} = \sum_{i \in \mathcal{B}(n)} i\, 2^{i-1}.$$

The binary form $\frac{1}{2} \sum_{i \in \mathcal{B}(n)} i 2^i$ is the one used below.

Theorem 2.2. *Let $a \ge 0$. Then $a \in \mathcal{A}$ if and only if $2^{a+3} \in \mathcal{E}$. More precisely, if T is a finite set of positive integers with $\sum_{j \in T} j2^j = 2^{a+3}$, then $1 \notin T$, and $n = \sum_{j \in T} 2^{j-2}$ satisfies $c(n) = 2^a$; conversely, if $c(n) = 2^a$, then $T = \mathcal{B}(n) + 2$ satisfies $\sum_{j \in T} j2^j = 2^{a+3}$.*

Proof. By (4), $8c(n) = \sum_{i \in \mathcal{B}(n)} (i+2) 2^{i+2} = \sum_{j \in \mathcal{B}(n)+2} j2^j$, which gives the converse part. For the direct part, every $j \ge 2$ gives $j2^j \equiv 0 \pmod 8$, while $1 \cdot 2^1 = 2$; since $2^{a+3} \equiv 0 \pmod 8$, $1 \notin T$. As $2^{a+3} > 0$, $T \ne \emptyset$, so $n = \sum_{j \in T} 2^{j-2}$ is a positive integer with $\mathcal{B}(n) = T - 2$, and $8c(n) = \sum_{j \in T} j2^j = 2^{a+3}$. □

3. The largest element of T

Lemma 3.1. *For every $J \geq 0$, $\sum_{j=1}^{J} j2^j = (J-1)2^{J+1} + 2$. In particular $\max \mathcal{E}_J = (J-1)2^{J+1} + 2 < (J+1)2^{J+1}$.*

Proof. The identity holds for $J = 0$, and if it holds for J, then $(J-1)2^{J+1} + 2 + (J+1)2^{J+1} = J2^{J+2} + 2$. The inequality is $2 < 2^{J+2}$. □

Lemma 3.2. *Let $J \geq 1$ and $V \in \mathcal{E}_J$ with $V > 0$, and let $T \subseteq \{1, \dots, J\}$ with $\sum_{i \in T} i2^i = V$. Then $\max T$ is the largest $j \in \{1, \dots, J\}$ with $j2^j \leq V$, and T is unique. Consequently, if $V > 0$ and j is the largest element of $\{1, \dots, J\}$ with $j2^j \leq V$, then $V \in \mathcal{E}_J$ if and only if $V - j2^j \in \mathcal{E}_{j-1}$.*

Proof. Let $m = \max T$, so $m2^m \leq V$. If some j with $m < j \leq J$ satisfied $j2^j \leq V$, Lemma 3.1 would give $V \leq \sum_{i=1}^{m} i2^i < (m+1)2^{m+1} \leq j2^j \leq V$, a contradiction. Hence m is determined by V, and uniqueness of T follows by induction on $|T|$ applied to $V - m2^m \in \mathcal{E}_{m-1}$. For the last statement, if $V \in \mathcal{E}_J$ the representing set has largest element j, so $V - j2^j \in \mathcal{E}_{j-1}$; conversely, adding j to a set representing $V - j2^j$ in $\mathcal{E}_{j-1}$ represents V in $\mathcal{E}_J$. □

For $K \geq 2$ put

$$h(K) = \lceil \log_2 K \rceil, \qquad \beta(K) = K + h(K). \tag{5}$$

Theorem 3.3. *Let $b \geq 3$.*

1. *If $2^b \in \mathcal{E}$ and T is the set representing 2^b, then $K = \max T$ satisfies $K \geq 2$ and $\beta(K) = b$.*
2. *The map β is strictly increasing on $\{2, 3, \dots\}$, and $\{b \geq 3\} \setminus \beta(\{2, 3, \dots\}) = \{2^r + r + 1 : r \geq 1\}$.*
3. *If $b = \beta(K)$ and $h = h(K)$, then $0 \leq 2^h - K \leq K - 2$, and $2^b \in \mathcal{E}$ if and only if $2^K(2^h - K) \in \mathcal{E}_{K-1}$.*

In particular, for $a \geq 0$: if $a = 2^r + r - 2$ with $r \geq 1$, then $a \notin \mathcal{A}$; otherwise there is exactly one $K \geq 2$ with $\beta(K) = a + 3$, and $a \in \mathcal{A}$ if and only if $2^K(2^{h(K)} - K) \in \mathcal{E}_{K-1}$.

Proof. (1) By Lemma 3.1, $K2^K \leq 2^b \leq (K-1)2^{K+1} + 2$. If $K = 1$ this gives $2^b = 2$, contradicting $b \geq 3$; so $K \geq 2$. Put $g = b - K$. Then $2^g \geq K \geq 2$, so $g \geq 1$, and $2^g \leq 2K - 2 + 2^{1-K} < 2K - 1$, hence $2^g \leq 2K - 2$. Thus $K \leq 2^g < 2K$, that is $g = \lceil \log_2 K \rceil$ and $b = \beta(K)$.

(2) $\beta(K+1) - \beta(K) = 1 + \lceil \log_2(K+1) \rceil - \lceil \log_2 K \rceil$, and $\lceil \log_2(K+1) \rceil > \lceil \log_2 K \rceil$ exactly when K is a power of 2, in which case the difference of the ceilings is 1. Hence β increases by 1 at each step except from $K = 2^r$ to $K = 2^r + 1$, where $\beta(2^r) = 2^r + r$ and $\beta(2^r + 1) = 2^r + r + 2$; since $\beta(2) = 3$, the values $b \geq 3$ not taken are exactly $2^r + r + 1$, $r \geq 1$.

(3) Since $2^{h-1} < K \leq 2^h$, we have $K \geq 2^{h-1} + 1$, so $2^h \leq 2K - 2$, which gives $0 \leq 2^h - K \leq K - 2$. If $2^b \in \mathcal{E}$, part (1) and part (2) show that the largest element of the representing set is K, and Lemma 3.2 gives $2^b - K2^K = 2^K(2^h - K) \in \mathcal{E}_{K-1}$. Conversely, if $2^K(2^h - K) \in \mathcal{E}_{K-1}$, adding K to the representing set represents 2^b.

The final statement follows from Theorem 2.2 with $b = a + 3$. □

The exponents $a = 2^r + r - 2$ are those of the first family of Nicolas [1]. Theorem 3.3 shows that every other a has exactly one candidate leading term, and that the candidates K for $a \leq x$ are the integers $2 \leq K \leq K_{\max}(x)$, where $K_{\max}(x)$ is the largest K with $\beta(K) \leq x + 3$.

4. A recursion on small integers

By Theorem 3.3(3), the membership of a in $\mathcal{A}$ is the question whether $2^{J+1}w \in \mathcal{E}_J$ for the pair $(J, w) = (K-1, 2^h - K)$, which satisfies $J \geq 1$ and $0 \leq w \leq J - 1$. We call such a pair a *state*, and we call it *successful* if $2^{J+1}w \in \mathcal{E}_J$.

Theorem 4.1. *Let (J, w) be a state, that is $J \geq 1$ and $0 \leq w \leq J - 1$.*

1. *If $w = 0$, the state is successful.*
2. *If $w \geq 1$, then $J \geq 2$; let s be the least integer $\sigma \geq 1$ with $2^\sigma w + \sigma \geq J + 1$. Then $1 \leq s \leq J - 1$, and:*
   (a) *if $2^{s-1}w + s = J + 1$, the state is not successful;*

(b) otherwise $w' = 2^s w + s - J - 1$ *satisfies* $0 \le w' \le J - s - 1$*, the pair* $(J - s, w')$ *is a state, and* (J, w) *is successful if and only if* $(J - s, w')$ *is successful.*

Proof. (1) $0 \in \mathcal{E}_J$. (2) Since $1 \le w \le J - 1$, $J \ge 2$. For $\sigma = J - 1$ we have $2^{J-1}w + J - 1 \ge 2 + J - 1 = J + 1$, so s exists and $s \le J - 1$. Put $V = 2^{J+1}w > 0$ and $j = J + 1 - s$, so $2 \le j \le J$. For $\sigma \in \{1, \dots, J\}$ and $i = J + 1 - \sigma$, the inequality $i2^i \le V$ is equivalent to $J + 1 - \sigma \le 2^\sigma w$; hence the largest $i \le J$ with $i2^i \le V$ corresponds to the least admissible σ, that is $i = j$. By Lemma 3.2, $V \in \mathcal{E}_J$ if and only if $V - j2^j = 2^j(2^s w - j) = 2^j w' \in \mathcal{E}_{j-1}$, where $w' = 2^s w - j = 2^s w + s - J - 1 \ge 0$.

We bound w' from above. If $s = 1$, then $w' = 2w - J \le J - 2 = j - 2$. If $s \ge 2$, the minimality of s gives $2^{s-1}w + s - 1 \le J$, hence $2^s w \le 2J + 2 - 2s$ and $w' \le J + 1 - s = j$. Moreover $w' = j - 1$ would mean $2^s w = 2j - 1$, which is impossible because $2^s w$ is even. Therefore either $w' \le j - 2$ or $w' = j$, and $w' = j$ is equivalent to $2^s w = 2j$, that is $2^{s-1}w + s = J + 1$ (which cannot occur for $s = 1$ since $w \le J - 1$).

In case (a), $w' = j$ and $2^j w' = j2^j > (j - 2)2^j + 2 = \max \mathcal{E}_{j-1}$ by Lemma 3.1, since $2^{j+1} > 2$; hence $V \notin \mathcal{E}_J$. In case (b), $w' \le j - 2 = (J - s) - 1$ and $J - s = j - 1 \ge 1$, so $(J - s, w')$ is a state, and $2^j w' = 2^{(J-s)+1}w'$, so $V \in \mathcal{E}_J$ if and only if $(J - s, w')$ is successful. □

Algorithm 1 Membership of a in $\mathcal{A}$

Require: an integer $a \ge 0$
$b \leftarrow a + 3$
if there is no $K \ge 2$ with $K + \lceil \log_2 K \rceil = b$ **then**
  return $a \notin \mathcal{A}$
end if
$h \leftarrow \lceil \log_2 K \rceil$; $J \leftarrow K - 1$; $w \leftarrow 2^h - K$
while $w \ge 1$ **do**
  $s \leftarrow$ least $\sigma \ge 1$ with $2^\sigma w + \sigma \ge J + 1$
  if $2^{s-1}w + s = J + 1$ **then**
    return $a \notin \mathcal{A}$
  end if
  $w \leftarrow 2^s w + s - J - 1$; $J \leftarrow J - s$
end while
return $a \in \mathcal{A}$

Corollary 4.2. *Algorithm 1 decides whether* $a \in \mathcal{A}$*. It performs at most* $K - 2 \le a + 1$ *iterations of the loop, and every integer it handles, other than the bound* $2^h \le 2K - 2$ *of the initial step, is at most* $2K \le 2a + 6$*. When it returns* $a \in \mathcal{A}$*, the indices* $j = J + 1 - s$ *produced at the successive iterations, together with* K*, form the set* T *of Theorem 2.2, and* $n = \sum_{j \in T} 2^{j-2}$ *satisfies* $c(n) = 2^a$*.*

Proof. Correctness follows from Theorems 3.3 and 4.1. Each iteration decreases J by $s \ge 1$ and keeps $J \ge 1$, starting from $J = K - 1$, so there are at most $K - 2$ iterations; $K \le \beta(K) - 1 = a + 2$. Throughout, $0 \le w \le J - 1 \le K - 2$, and in the computation of s only $\sigma \le s$ are tested, with $2^\sigma w + \sigma \le 2^s w + s \le 2J + 2 - s \le 2K$ for $s \ge 2$ by the proof of Theorem 4.1, and $2w + 1 \le 2K$ for $s = 1$. The last statement follows from the proof of Theorem 4.1 and Lemma 3.2. □

Example 4.3. For $a = 5$, $b = 8 = \beta(5)$, $h = 3$ and the initial state is $(4, 3)$. The loop gives $s = 1$ and $(3, 2)$, then $s = 1$ and $(2, 1)$, then $s = 1$ and $(1, 0)$, so $5 \in \mathcal{A}$ with $T = \{5, 4, 3, 2\}$ and $n = 2^3 + 2^2 + 2 + 1 = 15$; indeed $c(15) = 32$ [1]. For $a = 6$, $b = 9 = \beta(6)$, $h = 3$ and the initial state is $(5, 2)$; here $s = 2$ and $2^1 \cdot 2 + 2 = 6 = J + 1$, so $6 \notin \mathcal{A}$.

5. Explicit members of $\mathcal{A}$

Theorem 5.1. *1. For every* $r \ge 1$*,* $c(2^{2^r - 2}) = 2^{2^{2r+r-3}}$*; thus* $2^r + r - 3 \in \mathcal{A}$*.*

2. *Let $k \geq 0$, $s = 2^k$, and let $h \geq k+2$ with $h \equiv k \pmod{2^{k+1}}$. Then*

$$K = \frac{2^{h+s} + s}{2^s + 1} \tag{6}$$

is an integer with $2^{h-1} < K < 2^h$ and $K - s \geq 2$, and

$$c(2^{K-2} + 2^{K-s-2}) = 2^{K+h-3}; \tag{7}$$

thus $K + h - 3 = \beta(K) - 3 \in \mathcal{A}$.

Proof. (1) By Lemma 2.1 with $m = 2^r - 2$, $c(2^m) = 2^{m-1}(m+2) = 2^{2^r-3}2^r$.

(2) Modulo $2^s + 1$ we have $2^s \equiv -1$, hence $2^{2s} \equiv 1$. Writing $h = k + 2sq$ with $q \geq 0$ we get $2^h \equiv 2^k = s$, and therefore $2^{h+s} + s \equiv -2^h + s \equiv 0 \pmod{2^s+1}$, so K is an integer. Next, $K = 2^h - (2^h - s)/(2^s+1)$. Since $s = 2^k < 2^h$, $K < 2^h$; since $2^s + 1 \geq 3$, $(2^h - s)/(2^s+1) < 2^{h-1}$, so $K > 2^{h-1}$. As $h \geq k+2$, $s \leq 2^{h-2}$ and $K - s \geq 2^{h-1} + 1 - 2^{h-2} \geq 2$. By Lemma 2.1, with $K - 2 > K - s - 2 \geq 0$,

$$c(2^{K-2} + 2^{K-s-2}) = 2^{K-3}K + 2^{K-s-3}(K-s) = 2^{K-s-3}(K(2^s+1) - s) = 2^{K-s-3}2^{h+s} = 2^{K+h-3}.$$

Since $2^{h-1} < K < 2^h$, $h = h(K)$ and $K + h = \beta(K)$. □

For $k = 0$ ($s = 1$) the condition is that $h \geq 2$ is even, $K = (2^{h+1}+1)/3$, and the exponent is $K + h - 3 = (h+1) + (2^{h+1} - 11)/3$; with $a = h+1$ this is the family $a + (2^a - 11)/3$, a odd, of Nicolas [1]. For $k = 1$ ($s = 2$), $h \equiv 1 \pmod 4$; for instance $h = 5$ gives $K = 26$ and $c(2^{24} + 2^{22}) = 2^{28}$.

6. Explicit non-members of $\mathcal{A}$

Theorem 6.1. *Let $s \geq 2$ and $h \geq 1$ with $s < 2^h$ and $2^h \equiv s \pmod{2^{s-1}+1}$. Then*

$$K = \frac{2^{h+s-1} + s}{2^{s-1} + 1} \tag{8}$$

is an integer with $2^{h-1} < K < 2^h$, and $K + h - 3 = \beta(K) - 3 \notin \mathcal{A}$. The hypotheses hold in the following cases:

1. *$s = 2^i$ with $i \geq 1$, and $h > i$ with $h \equiv i \pmod{2(2^i - 1)}$;*
2. *$s = 3$, and $h \equiv 3 \pmod 4$.*

Proof. Put $M = 2^{s-1} + 1$. Since $2^{s-1} \equiv -1 \pmod M$, $2^{h+s-1} + s \equiv -2^h + s \equiv 0 \pmod M$, so K is an integer. As $K = 2^h - (2^h - s)/M$ with $0 < 2^h - s$ and $M \geq 3$, we get $2^{h-1} < K < 2^h$, hence $h = h(K)$ and $K + h = \beta(K)$. By Theorem 2.2 and Theorem 3.3, $K + h - 3 \in \mathcal{A}$ if and only if the state $(J, w) = (K-1, 2^h - K)$ is successful. Here $w \geq 1$, and $K \geq 3$ because $s \geq 2$ and $s < 2^h$ force $h \geq 2$; so $J \geq 2$. The definition of K is equivalent to $K(2^{s-1}+1) = 2^{h+s-1} + s$, that is

$$2^{s-1}w + s = K = J + 1.$$

We check that s is the least $\sigma \geq 1$ with $2^\sigma w + \sigma \geq J + 1$. The function $\sigma \mapsto 2^\sigma w + \sigma$ is increasing; at $\sigma = s - 1 \geq 1$ its value is J, and at $\sigma = s$ it is $2(K - s) + s = 2K - s$, which is at least K because $K - s = (2^h - s)(1 - 1/M) > 0$. Hence Theorem 4.1(2a) applies, and the state is not successful.

Case (1): $M = 2^{s-1} + 1$ with $s = 2^i$, so $2^{2(s-1)} \equiv 1 \pmod M$ and $h = i + 2(s-1)q$ gives $2^h \equiv 2^i = s$; and $s = 2^i < 2^h$ because $h > i$. Case (2): $M = 5$, $2^4 \equiv 1$ and $2^3 \equiv 3 \pmod 5$, so $h \equiv 3 \pmod 4$ gives $2^h \equiv 3 = s$; and $3 < 2^h$ because $h \geq 3$. □

For $s = 2$ the condition is that $h \geq 3$ is odd, $K = (2^{h+1}+2)/3$, and the exponent is $K + h - 3 = (h+1) + (2^{h+1} - 10)/3$; with $a = h+1$ this is the family $a + (2^a - 10)/3$, a even, of Nicolas [1]. For $s = 3$ and $h = 3$, $K = 7$ and $7 \notin \mathcal{A}$; for $s = 4$ and $h = 8$, $K = 228$ and $233 \notin \mathcal{A}$.

The families of Theorems 5.1 and 6.1 are exactly the exponents decided by at most one iteration of Algorithm 1.

Theorem 6.2. *Let $a \geq 0$ have a candidate K (Theorem 3.3), let $h = h(K)$, and let $(J, w) = (K-1, 2^h - K)$ be the initial state.*

1. *Algorithm 1 returns $a \in \mathcal{A}$ without any iteration if and only if $K = 2^h$, that is, a is an exponent of Theorem 5.1(1).*
2. *Algorithm 1 returns $a \in \mathcal{A}$ after exactly one iteration if and only if a is an exponent of Theorem 5.1(2), that is, $K = (2^{h+s} + s)/(2^s + 1)$ with $s = 2^k$, $k \geq 0$, $h \geq k+2$ and $h \equiv k \pmod{2^{k+1}}$.*
3. *Algorithm 1 returns $a \notin \mathcal{A}$ at its first iteration if and only if a is an exponent of Theorem 6.1(1) or (2), that is, $K = (2^{h+s-1} + s)/(2^{s-1} + 1)$ with either $s = 2^i$, $i \geq 1$, $h > i$ and $h \equiv i \pmod{2(2^i - 1)}$, or $s = 3$ and $h \equiv 3 \pmod 4$.*

Proof. (1) The loop is not entered if and only if $w = 0$, that is $K = 2^h$.

We use the following fact on residues. Let $m \geq 1$ and $N = 2^m + 1$. Since $2^m \equiv -1 \pmod N$, the residues of the powers 2^e, $e \geq 0$, are 2^l and $N - 2^l$ for $0 \leq l \leq m-1$; these $2m$ residues are distinct, because $N - 2^l \geq 2^{m-1} + 1 > 2^{m-1} \geq 2^{l'}$, so 2 has order $2m$ modulo N and $2^e \equiv 2^l \pmod N$ with $0 \leq l \leq m-1$ holds if and only if $e \equiv l \pmod{2m}$.

(2) Suppose $w \geq 1$. The first iteration ends with $w' = 0$, without failure, if and only if $2^s w + s = J + 1 = K$ for the least admissible s. Conversely, if $2^\sigma w + \sigma = K$ for some $\sigma \geq 1$, then σ is the least admissible value, because $2^{\sigma-1} w + \sigma - 1 = (K + \sigma - 2)/2 < K$ (as $\sigma < K$), and the failure condition $2^{\sigma-1} w + \sigma = (K + \sigma)/2 = K$ would give $K = \sigma$, which is impossible since $K = 2^\sigma w + \sigma > \sigma$. Hence one-iteration success is equivalent to the existence of $s \geq 1$ with $2^s(2^h - K) + s = K$, that is $K(2^s + 1) = 2^{h+s} + s$, which requires $2^{h+s} + s \equiv 0$, that is $2^h \equiv s \pmod{2^s + 1}$. Apply the fact with $m = s$: $s < 2^s + 1$, and $s = 2^s + 1 - 2^l$ with $l \leq s - 1$ would give $s \geq 2^{s-1} + 1 > s$; hence $s = 2^l$ for some l, which we call k, with $k \leq s - 1$, and $2^h \equiv 2^k$ gives $h \equiv k \pmod{2^{k+1}}$. Moreover $w \geq 1$ means $K < 2^h$, which is $s < 2^h$, that is $h > k$; with $h \equiv k \pmod{2^{k+1}}$ this gives $h \geq k + 2^{k+1} \geq k + 2$. Conversely, under these conditions Theorem 5.1(2) gives $K(2^s + 1) = 2^{h+s} + s$ with $1 \leq w$.

(3) Suppose $w \geq 1$. By Theorem 4.1, the first iteration fails if and only if $2^{s-1} w + s = J + 1 = K$ for the least admissible s, and then $s \geq 2$. Conversely, if $2^{\sigma-1} w + \sigma = K$ for some $\sigma \geq 2$, the proof of Theorem 6.1 shows that σ is the least admissible value. Hence first-iteration failure is equivalent to the existence of $s \geq 2$ with $K(2^{s-1} + 1) = 2^{h+s-1} + s$, which requires $2^h \equiv s \pmod M$, $M = 2^{s-1} + 1$, and $w \geq 1$ is $s < 2^h$. Apply the fact with $m = s - 1$; note $2 \leq s < M$. If $s = 2^l$ with $l \leq s - 2$, then $l \geq 2$, and $h \equiv l \pmod{2(s-1)}$, which with $s < 2^h$ is case (1) of Theorem 6.1 with $i = l \geq 2$. Otherwise $s = M - 2^l$ with $0 \leq l \leq s - 2$, that is $2^{s-1} + 1 - s = 2^l$. For $s \geq 4$ this is impossible, since $2^{s-2} < 2^{s-1} + 1 - s < 2^{s-1}$. For $s = 2$, $l = 0$ and $2^h \equiv M - 1 = 2 \pmod 3$, that is h odd, and $s < 2^h$ gives $h \geq 3$: case (1) with $i = 1$. For $s = 3$, $l = 1$ and $2^h \equiv 3 \pmod 5$, that is $h \equiv 3 \pmod 4$: case (2). Conversely, Theorem 6.1 shows that its exponents fail at the first iteration. □

Consequently, every exponent a that is not of the form $2^r + r - 2$ and is not covered by Theorems 5.1 and 6.1 requires at least two iterations of Algorithm 1.

7. Counting functions

Theorem 7.1. *As $x \to \infty$,*

$$A(x) \geq 2 \log_2 x - O(\log \log x), \qquad B(x) \geq \Big(\frac{5}{4} + \frac{E}{2}\Big) \log_2 x - O(\log \log x),$$

where $E = \sum_{i \geq 1} 1/(2^i - 1) = 1.6066951\ldots$, so that $\frac{5}{4} + \frac{E}{2} = 2.0533475\ldots$.

Proof. Let $x \geq 16$ and $H = \lfloor \log_2 x \rfloor - 1 \geq 3$. If $1 \leq h \leq H$ and $2 \leq K \leq 2^h$, then $\beta(K) - 3 \leq 2^h + h - 3 \leq x/2 + \log_2 x \leq x$. All exponents produced by Theorems 5.1(2) and 6.1 are of the form $\beta(K) - 3$ with $2^{h-1} < K < 2^h$, and β is injective (Theorem 3.3(2)), so distinct values of K give distinct exponents. For fixed h, the integer K in (6) and in (8) equals $2^h - (2^h - s)/(2^s + 1)$ and $2^h - (2^h - s)/(2^{s-1} + 1)$ respectively, both strictly increasing in s, so distinct values of s in the same theorem give distinct values of K; a member and a non-member are distinct.

Members. Theorem 5.1(1) gives the H exponents $\beta(2^h) - 3$, $1 \leq h \leq H$, which are distinct from those of part (2) since there $K < 2^h$. For $0 \leq k \leq \kappa = \lfloor \log_2 H \rfloor$, the integers $h = k + 2^{k+1} t$ with $1 \leq t \leq \lfloor (H - k)/2^{k+1} \rfloor$ satisfy

the hypotheses of Theorem 5.1(2) and $h \le H$; their number is at least $(H-k)/2^{k+1} - 1$. Summing over k and using $\sum_{k\ge 0} k/2^{k+1} = 1$ and $2^{-\kappa-1}H < 1$,

$$\sum_{k=0}^{\kappa}\Big(\frac{H-k}{2^{k+1}} - 1\Big) \ge H(1-2^{-\kappa-1}) - 1 - (\kappa+1) \ge H - \log_2 H - 3.$$

Hence $A(x) \ge 2H - \log_2 H - 3 = 2\log_2 x - O(\log\log x)$.

Non-members. Theorem 3.3 gives the H exponents $2^r + r - 2 \le x$, $1 \le r \le H$, which are not of the form $\beta(K) - 3$. For $1 \le i \le \kappa$, the integers $h = i + 2(2^i - 1)t$ with $1 \le t \le \lfloor (H-i)/(2(2^i-1))\rfloor$ satisfy Theorem 6.1(1) and $h \le H$; their number is at least $(H-i)/(2(2^i-1)) - 1$. Since $\sum_{i>\kappa} 1/(2(2^i-1)) \le \sum_{i>\kappa} 2^{-i} = 2^{-\kappa} < 2/H$ and $\sum_{i\ge 1} i/(2(2^i-1)) < \infty$, the sum over i is at least $\frac{E}{2}H - O(\log H)$. Theorem 6.1(2) gives the integers $3 \le h \le H$ with $h \equiv 3 \pmod 4$, at least $H/4 - 1$ of them. Hence $B(x) \ge (1 + \frac{1}{4} + \frac{E}{2})H - O(\log H)$, and $H = \log_2 x + O(1)$. □

8. Successful states

For $J \ge 1$ let

$$S_J = \{w :\ 0 \le w \le J-1,\ 2^{J+1}w \in \mathcal{E}_J\}, \qquad N(J) = |S_J|,$$

so that S_J is the set of w for which the state (J, w) of Section 4 is successful. By Theorems 2.2 and 3.3, an exponent a with candidate K and $h = h(K)$ belongs to $\mathcal{A}$ if and only if $2^h - K \in S_{K-1}$.

Proposition 8.1. *1. For $J \ge 1$, $N(J) = \#\{T \subseteq \{1,\dots,J\} :\ \sum_{j\in T} j2^j \equiv 0 \pmod{2^{J+1}}\}$.*

2. For $h \ge 1$, the number of $a \in \mathcal{A}$ whose candidate K lies in $(2^{h-1}, 2^h]$ equals $\#\{K \in (2^{h-1}, 2^h] :\ 2^h - K \in S_{K-1}\}$.

Proof. (1) By Lemma 3.2, the map $T \mapsto \sum_{j\in T} j2^j$ is injective on the subsets of $\{1,\dots,J\}$. If $\sum_{j\in T} j2^j \equiv 0 \pmod{2^{J+1}}$, the sum equals $2^{J+1}w$ with $w \ge 0$, and $2^{J+1}w \le (J-1)2^{J+1} + 2$ by Lemma 3.1, so $w \le J-1$ and $w \in S_J$; conversely every $w \in S_J$ is obtained from exactly one T. (2) The map $K \mapsto \beta(K) - 3$ is injective (Theorem 3.3), and $\beta(K) - 3 \in \mathcal{A}$ if and only if $2^h - K \in S_{K-1}$. □

Proposition 8.2. *For every $J \ge 1$, $w \in S_J$ if and only if $J - 1 - w \in S_J$.*

Proof. For $J = 1$, $S_1 = \{0\}$. Let $J \ge 2$, $w \in S_J$, and let $T \subseteq \{1,\dots,J\}$ represent $V = 2^{J+1}w$. Since $J + 1 \ge 3$, $V \equiv 0 \pmod 8$, and as in the proof of Theorem 2.2, $1 \notin T$. By Lemma 3.1, $\sum_{j=2}^{J} j2^j = (J-1)2^{J+1}$, so the set $\{2,\dots,J\} \setminus T$ represents $(J-1)2^{J+1} - 2^{J+1}w = 2^{J+1}(J-1-w)$, and $J-1-w \in S_J$. Applying this to $J-1-w$ gives the converse. □

Proposition 8.3. *For every $k \ge 1$, $N(2k+1) = N(2k)$, and $4 \mid N(J)$ for every $J \ge 4$.*

Proof. Let $k \ge 1$. For $T \subseteq \{1,\dots,2k+1\}$ put $U = T \setminus \{2k+1\}$ and $\sigma_U = \sum_{j\in U} j2^j$. Since $(2k+1)2^{2k+1} \equiv 2^{2k+1} \pmod{2^{2k+2}}$, the condition $\sum_{j\in T} j2^j \equiv 0 \pmod{2^{2k+2}}$ is $\sigma_U \equiv -[2k+1 \in T]\,2^{2k+1} \pmod{2^{2k+2}}$. If $\sigma_U \equiv 0 \pmod{2^{2k+1}}$, then σ_U is congruent to 0 or to 2^{2k+1} modulo 2^{2k+2}, and exactly one of the two choices for $2k+1 \in T$ satisfies the condition; otherwise none does. By Proposition 8.1(1), $N(2k+1) = \#\{U \subseteq \{1,\dots,2k\} :\ \sigma_U \equiv 0 \pmod{2^{2k+1}}\} = N(2k)$.

Let $k \ge 2$. Since $2k \cdot 2^{2k} = k2^{2k+1} \equiv 0 \pmod{2^{2k+1}}$, the element $2k$ does not affect the condition of Proposition 8.1(1) for $J = 2k$, so $N(2k) = 2Z$ with $Z = \#\{U \subseteq \{1,\dots,2k-1\} :\ \sigma_U \equiv 0 \pmod{2^{2k+1}}\}$. For such U, $1 \notin U$, because $2k+1 \ge 5$ and $\sigma_U \equiv 2 \pmod 8$ when $1 \in U$. By Lemma 3.1, $\sum_{j=2}^{2k-1} j2^j = (2k-2)2^{2k} = (k-1)2^{2k+1} \equiv 0 \pmod{2^{2k+1}}$, so $U \mapsto \{2,\dots,2k-1\} \setminus U$ maps the sets counted by Z to themselves; it is an involution without fixed points because $\{2,\dots,2k-1\} \ne \emptyset$. Hence Z is even and $4 \mid N(2k)$; with $N(2k+1) = N(2k)$ this gives $4 \mid N(J)$ for all $J \ge 4$. □

The recursion of Theorem 4.1 organizes all states into a forest. For a state (J, w) with $w \ge 1$ that does not fail, let $\Phi(J, w) = (J - s,\ 2^s w + s - J - 1)$, with s as in Theorem 4.1; by that theorem, $\Phi(J, w)$ is a state. We call (J, w) a *root* if $w = 0$ or if (J, w) fails. For $J \ge 1$ let $F(J)$ be the number of states (J, w) that fail.

Theorem 8.4. *1. A state (J, w) with $w \ge 1$ fails if and only if $w = (J + 1 - \sigma)/2^{\sigma-1}$ for some integer $\sigma \ge 2$ with $2^{\sigma-1} \mid J + 1 - \sigma$ and $J + 1 - \sigma \ge 2^{\sigma-1}$. Distinct σ give distinct w, so $F(J)$ is the number of such σ, and for every $L \ge 1$,*

$$\sum_{J=1}^{L} F(J) = \sum_{\sigma=2}^{L}\Big\lfloor \frac{L+1-\sigma}{2^{\sigma-1}}\Big\rfloor.$$

2. *Let (J', w') be a state and $m = v_2(J' + 1 + w')$. The states (J, w) with $w \geq 1$ that do not fail and satisfy $\Phi(J, w) = (J', w')$ are exactly the m states $(J' + \sigma,\ (J' + 1 + w')/2^\sigma)$, $1 \leq \sigma \leq m$.*

3. *The map Φ makes the set of states a forest whose roots are the states $(J, 0)$, $J \geq 1$, and the failing states; a state is successful if and only if the root of its tree is a state $(J, 0)$. For every $J \geq 2$,*

$$N(J) = 1 + \sum_{\sigma=1}^{J-1} \#\{w' \in S_{J-\sigma} :\ 2^\sigma \mid J-\sigma+1+w'\}, \qquad J = 1 + F(J) + \sum_{\sigma=1}^{J-1} \#\{0 \leq w' \leq J-\sigma-1 :\ 2^\sigma \mid J-\sigma+1+w'\}.$$

4. *The initial states $(K - 1, 2^h - K)$, $K \in (2^{h-1}, 2^h]$, are exactly the states (J, w) with $J + 1 + w = 2^h$.*

Proof. (1) By Theorem 4.1, (J, w) fails if and only if $2^{s-1}w + s = J + 1$ for the least admissible s. Conversely, suppose $2^{\sigma-1}w + \sigma = J + 1$ for some $\sigma \geq 2$. The function $\tau \mapsto 2^\tau w + \tau$ is increasing; at $\tau = \sigma - 1$ its value is $J < J + 1$, and at $\tau = \sigma$ it is $2(J + 1 - \sigma) + \sigma = 2J + 2 - \sigma \geq J + 1$, because $J + 1 = 2^{\sigma-1}w + \sigma > \sigma$. Hence σ is the least admissible value and the state fails. Therefore (J, w) with $w \geq 1$ fails if and only if $w = (J + 1 - \sigma)/2^{\sigma-1}$ for some $\sigma \geq 2$, this quotient being an integer at least 1; such a w satisfies $w \leq (J - 1)/2 \leq J - 1$, so (J, w) is a state. The quotient is strictly decreasing in σ, so distinct σ give distinct w. Finally, for fixed $\sigma \geq 2$, the integers $J \leq L$ with $J + 1 - \sigma = 2^{\sigma-1}t$, $t \geq 1$, correspond to $1 \leq t \leq (L + 1 - \sigma)/2^{\sigma-1}$, and there are none when $\sigma > L$; summing over σ gives the formula.

(2) Put $u = J' + 1 + w'$; since $0 \leq w' \leq J' - 1$, $J' + 1 \leq u \leq 2J'$. If $\Phi(J, w) = (J', w')$ with step s, then $J = J' + s$ and $2^s w = w' + J + 1 - s = u$, so $2^s \mid u$, $s \leq m$ and $w = u/2^s$. Conversely, let $1 \leq \sigma \leq m$, $J = J' + \sigma$ and $w = u/2^\sigma$. Then $w \geq 1$, and $w \leq u/2 \leq J' \leq J - 1$, so (J, w) is a state. We have $2^\sigma w + \sigma = u + \sigma = J + 1 + w' \geq J + 1$, and $2^{\sigma-1}w + \sigma - 1 = u/2 + \sigma - 1 < J + 1$, since this is equivalent to $w' < J' + 3$; so σ is the least admissible value. The failure condition $2^{\sigma-1}w + \sigma = J + 1$ would mean $u = 2J' + 2$, that is $w' = J' + 1$, which is impossible. Finally $2^\sigma w + \sigma - J - 1 = u + \sigma - J' - \sigma - 1 = w'$, so $\Phi(J, w) = (J', w')$.

(3) Φ decreases J, so the orbit of every state ends at a root after finitely many steps, and every state that is not a root has exactly one image; this is a forest with the stated roots. By Theorem 4.1, a state is successful if and only if its orbit ends at a state $(J, 0)$. The states of level J are $(J, 0)$, the $F(J)$ failing states, and the states (J, w) with $w \geq 1$ that do not fail; by part (2), the latter correspond bijectively to the pairs (σ, w') with $1 \leq \sigma \leq J - 1$, $(J - \sigma, w')$ a state and $2^\sigma \mid J - \sigma + 1 + w'$, through $(J, w) \mapsto (s, \Phi(J, w))$. Such a state is successful if and only if $w' \in S_{J-\sigma}$. Counting the successful states and all states of level J gives the two identities.

(4) If $J = K - 1$ and $w = 2^h - K$, then $J + 1 + w = 2^h$. Conversely, if $J + 1 + w = 2^h$ with $0 \leq w \leq J - 1$, put $K = J + 1$; then $K = 2^h - w \leq 2^h$ and $2^h \leq 2J < 2K$, so $K \in (2^{h-1}, 2^h]$ and $w = 2^h - K$. □

By Theorem 8.4(1), the number of failing states up to level L is at most $\sum_{\sigma \geq 2} L/2^{\sigma-1} = L$, and there is exactly one root $(J, 0)$ at each level. By Theorem 8.4(2), the number of children of a state (J', w') is $v_2(J' + 1 + w')$. Consequently, by parts (3) and (4), an exponent a with candidate K belongs to $\mathcal{A}$ if and only if the state $(K - 1, 2^{h(K)} - K)$, the unique state of level $K - 1$ with $J + 1 + w$ a power of 2, lies in the tree of a root $(J, 0)$.

Proposition 8.5. *Let (J, w) be a successful state with $w \geq 1$, and let $T \subseteq \{1, \ldots, J\}$ be the set representing $2^{J+1}w$. Then $1 \notin T$, and the root of the tree of (J, w) is $(\min T - 1, 0)$. For every $J \geq 2$, exactly $N(J)/2$ of the successful states of level J lie in the tree of the root $(1, 0)$; in particular $N(J)$ is even.*

Proof. Since $J + 1 \geq 3$, $2^{J+1}w \equiv 0 \pmod 8$, so $1 \notin T$, as in the proof of Theorem 2.2. By the proof of Theorem 4.1 and Lemma 3.2, if $j = \max T$ then $\Phi(J, w) = (j - 1, w')$, where $T \setminus \{j\}$ represents $2^j w'$. Applying this repeatedly removes the elements of T from the largest to the smallest; when the smallest element $\min T$ is removed, the remaining set is empty, the state reached is $(\min T - 1, 0)$, and $\min T - 1 \geq 1$. Hence the root is $(\min T - 1, 0)$. Now let $J \geq 2$. The successful state $(J, 0)$ is represented by $T = \emptyset$ and is its own root, different from $(1, 0)$. A successful state of level J therefore lies in the tree of $(1, 0)$ if and only if its representing set T satisfies $\min T = 2$, that is $2 \in T$, since $1 \notin T$. The map $T \mapsto \{2, \ldots, J\} \setminus T$ of the proof of Proposition 8.2 is a bijection of the set of representing sets of the states in S_J onto itself which exchanges the sets containing 2 and the sets not containing 2. Hence exactly half of the $N(J)$ successful states have $2 \in T$. □

The sets S_J can also be built from the smallest elements of T upwards.

Theorem 8.6. *We have $S_1 = \{0\}$, and for every $J \geq 2$:*

1. *if J is odd, $S_J = \{c/2 : c \in S_{J-1},\ c \text{ even}\} \cup \{(c+J)/2 : c \in S_{J-1},\ c \text{ odd}\}$, and $c \mapsto w$ is a bijection from S_{J-1} onto S_J;*

2. *if J is even, $S_J = \{c/2 : c \in S_{J-1},\ c \text{ even}\} \cup \{c/2 + J/2 : c \in S_{J-1},\ c \text{ even}\}$, and the two sets are disjoint.*

Consequently, $N(J) = N(J-1)$ for odd J, $N(J) = 2\#\{c \in S_{J-1} : c \text{ even}\}$ for even J, and for even J and $0 \leq w < J/2$, $w \in S_J$ if and only if $w + J/2 \in S_J$.

Proof. $S_1 = \{0\}$ because $\mathcal{E}_1 = \{0, 2\}$ and $4 \nmid 2$. Let $J \geq 2$. By Proposition 8.1(1) and its proof, $w \in S_J$ if and only if $2^{J+1}w = \sum_{j\in T} j2^j$ for a set $T \subseteq \{1, \dots, J\}$, which is then unique. Write $x = 1$ if $J \in T$ and $x = 0$ otherwise, and $U = T \setminus \{J\}$. If $w \in S_J$, then $\sum_{j\in U} j2^j = 2^{J+1}w - xJ2^J \equiv 0 \pmod{2^J}$, so $c = 2^{-J}\sum_{j\in U} j2^j \in S_{J-1}$ and

$$w = \frac{c + xJ}{2}.$$

Conversely, if $c \in S_{J-1}$, $x \in \{0, 1\}$ and $c + xJ$ is even, then $U \cup \{J\}$ (if $x = 1$) or U (if $x = 0$), where U represents $2^J c$, represents $2^{J+1}(c + xJ)/2$, so $(c + xJ)/2 \in S_J$. If J is odd, $c + xJ$ is even exactly for $x \equiv c \pmod 2$, which gives (1); the map $c \mapsto w$ is injective because w determines T, hence U and c. If J is even, $c + xJ$ is even exactly when c is even, with both $x = 0$ and $x = 1$, which gives (2); since $c \leq J - 2$, $c/2 < J/2 \leq c'/2 + J/2$ for all even $c, c' \in S_{J-1}$, so the two sets are disjoint. The consequences follow at once: for even J, $S_J = B \cup (B + J/2)$ with $B = \{c/2 : c \in S_{J-1},\ c \text{ even}\} \subseteq [0, J/2)$. □

Theorem 8.6 gives a second proof of the equality $N(2k+1) = N(2k)$ of Proposition 8.3. In terms of Theorem 8.6, an exponent a with candidate K belongs to $\mathcal{A}$ if and only if $2^{h(K)} - K \in S_{K-1}$, where the sets S_J are obtained from $S_1 = \{0\}$ by the maps $c \mapsto c/2$ and $c \mapsto (c + J)/2$ of parts (1) and (2).

The symmetry of Proposition 8.2 transports the states decided by one step of the recursion to further initial states. The first family below is the formula $c(2^k - 1) = k2^{k-1}$ of [1] at $k = 2^r$.

Theorem 8.7. 1. *For every $r \geq 0$, $c(2^{2^r} - 1) = 2^{2^r+r-1}$; the candidate of the exponent $2^r + r - 1$ is $K = 2^r + 1$.*

2. *Let $i \geq 1$, $\sigma = 2^{i-1} + 1$, and let $h \geq \sigma + 2$ with $h \equiv i \pmod{\sigma + 1}$. Then $w = (2^h - 2\sigma + 2)/(2^{\sigma+1} - 1)$ is an even positive integer, $K = 2^{h-1} + 1 + w/2$ satisfies $2^{h-1} + 1 < K < 2^h$, and $K + h - 3 \in \mathcal{A}$.*

3. *Let $i \geq 2$, $\sigma = 2^{i-1} + 1$, and let $h \geq \sigma + 1$ with $h \equiv i \pmod{\sigma}$. Then $w = (2^h - 2\sigma + 2)/(2^\sigma - 1)$ is an even positive integer, $K = 2^{h-1} + 1 + w/2$ satisfies $2^{h-1} + 1 < K < 2^h$, and $K + h - 3 \notin \mathcal{A}$.*

Proof. (1) By the proof of Lemma 2.1, $c(2^k - 1) = k2^{k-1}$; take $k = 2^r$. Then $\mathcal{B}(2^k - 1) + 2 = \{2, \dots, k + 1\}$, whose largest element is $K = 2^r + 1$.

(2) Put $M = 2^{\sigma+1} - 1$. Since $2^{\sigma+1} \equiv 1 \pmod M$ and $h \equiv i \pmod{\sigma + 1}$, $2^h \equiv 2^i = 2\sigma - 2 \pmod M$, so w is an integer; and $w \geq 1$ because $2^h \geq 2^{\sigma+2} = 2M + 2 > M + 2\sigma - 2$. Put $J' = 2^\sigma w - 1 \geq 1$ and $J = J' + \sigma$. Since $v_2(J' + 1) \geq \sigma$, Theorem 8.4(2) shows that (J, w) is a state that does not fail, with $\Phi(J, w) = (J', 0)$; hence $w \in S_J$, and $J - 1 - w \in S_J$ by Proposition 8.2. Moreover $2J - w = 2^{\sigma+1}w + 2\sigma - 2 - w = Mw + 2\sigma - 2 = 2^h$, so w is even and, with $K = J + 1$, $K = 2^{h-1} + 1 + w/2$ and $2^h - K = J - 1 - w \in S_{K-1}$. Since $w \geq 2$ and $w < 2^h - 2$, $2^{h-1} + 1 < K < 2^h$, so $h = h(K)$, and Theorem 3.3 gives $K + h - 3 \in \mathcal{A}$.

(3) Put $M' = 2^\sigma - 1$. As in (2), $2^h \equiv 2^i = 2\sigma - 2 \pmod{M'}$, so w is an integer, and $w \geq 1$ because $2^h \geq 2^{\sigma+1} = 2M' + 2 \geq M' + 2\sigma - 2$ for $\sigma \geq 3$. Put $J = 2^{\sigma-1}w + \sigma - 1$. Then $J + 1 - \sigma = 2^{\sigma-1}w \geq 2^{\sigma-1}$ and $w = (J + 1 - \sigma)/2^{\sigma-1}$, so the state (J, w) fails by Theorem 8.4(1), and $w \notin S_J$; by Proposition 8.2, $J - 1 - w \notin S_J$. Moreover $2J - w = 2^\sigma w + 2\sigma - 2 - w = M'w + 2\sigma - 2 = 2^h$, and the conclusion follows as in (2). □

Theorem 8.8. *As $x \to \infty$,*

$$A(x) \geq (3 + C_1)\log_2 x - O(\log\log x), \qquad B(x) \geq \Big(\frac{5}{4} + \frac{E}{2} + C_2\Big)\log_2 x - O(\log\log x),$$

where $C_1 = \sum_{k\geq 0} 1/(2^k + 2) = 0.9655832\dots$, $C_2 = \sum_{k\geq 1} 1/(2^k + 1) = 0.7644997\dots$ and E is as in Theorem 7.1; thus $3 + C_1 = 3.9655832\dots$ and $\frac{5}{4} + \frac{E}{2} + C_2 = 2.8178473\dots$.

Proof. We keep the notation $H = \lfloor \log_2 x \rfloor - 1$ of the proof of Theorem 7.1, whose families we use together with those of Theorem 8.7. All the exponents are of the form $\beta(K) - 3$ with $K \le 2^h$, $h \le H$, hence at most x, and distinct values of K give distinct exponents. For $h \ge 4$, the integers K of Theorem 8.7(2) and (3) satisfy $K = 2^{h-1} + 1 + w/2$ with $w \le 2^h/7$, hence $K < \frac{2}{3}\,2^h$ because $2^{h-1}(\frac{1}{3} - \frac{1}{7}) > 1$; the integers K of Theorems 5.1(2) and 6.1 satisfy $K > 2^h - 2^h/3 = \frac{2}{3}\,2^h$, since there the denominators $2^s + 1$ and $2^{s-1} + 1$ are at least 3. In Theorem 8.7(2) and (3), $w/2$ is strictly decreasing in σ, and $K > 2^{h-1} + 1$; and for $h \ge 3$ the integer $2^{h-1} + 1$ of Theorem 8.7(1) differs from 2^h and from the integers of Theorem 5.1(2), which exceed $\frac{2}{3}\,2^h$. Hence, for $h \ge 4$, all the integers K produced in the block $(2^{h-1}, 2^h]$ are distinct.

Members. Besides the $2H - \log_2 H - 3$ exponents of the proof of Theorem 7.1, Theorem 8.7(1) gives one exponent for each $4 \le h \le H$, and Theorem 8.7(2) gives, for each $1 \le i \le \kappa = \lfloor \log_2 H \rfloor$, at least $(H - \sigma - 1)/(\sigma + 1) - 1$ integers $h \in [\sigma + 2, H]$ with $h \equiv i \pmod{\sigma + 1}$, where $\sigma = 2^{i-1} + 1$. Since $\sum_{i>\kappa} 1/(2^{i-1} + 2) \le 2^{1-\kappa} < 4/H$ and $\sigma \le H/2 + 1$ for $i \le \kappa$, the sum over i is at least $C_1 H - O(\log H)$.

Non-members. Besides the exponents of the proof of Theorem 7.1, Theorem 8.7(3) gives, for each $2 \le i \le \kappa$, at least $(H - \sigma)/\sigma - 1$ integers $h \in [\sigma + 1, H]$ with $h \equiv i \pmod{\sigma}$; as above, the sum over i is at least $(\sum_{i \ge 2} 1/(2^{i-1} + 1))H - O(\log H) = C_2 H - O(\log H)$. Since $H = \log_2 x + O(1)$, the bounds follow. □

9. The analogue for an arbitrary prime

Let p be a prime and v_p the p-adic valuation. Put

$$c_p(n) = \sum_{k=1}^{n} p^{v_p(k)}, \qquad \mathcal{A}_p = \{a \ge 0 : \ p^a \in c_p(\mathbb{N})\}, \qquad A_p(x) = \#\{a \in \mathcal{A}_p : \ a \le x\},$$

and $B_p(x) = \#\{a \in \mathbb{Z} \cap [0, x] : \ a \notin \mathcal{A}_p\}$, so that $c_2 = c$, $\mathcal{A}_2 = \mathcal{A}$, $A_2 = A$ and $B_2 = B$. The interpretation of c recalled in Section 1 extends to every prime.

Proposition 9.1. *For every $n \ge 1$, $c_p(n)$ is the exponent of $1 - X$ in the factorization of $(1 - X)(1 - X^2) \cdots (1 - X^n)$ into irreducible factors over $\mathbb{F}_p[X]$.*

Proof. Write $k = p^v k'$ with $p \nmid k'$. In characteristic p, $1 - X^k = (1 - X^{k'})^{p^v}$, and $1 - X^{k'} = (1 - X)(1 + X + \cdots + X^{k'-1})$, where the second factor takes the value $k' \ne 0$ at $X = 1$; hence $1 - X$ divides $1 - X^k$ exactly $p^v = p^{v_p(k)}$ times, and the exponents add up over $k \le n$. □

Lemma 9.2. *1. For every $n \ge 1$, $c_p(pn) = p\,c_p(n) + (p - 1)n$.*

2. For every $m \ge 0$, $c_p(p^m - 1) = m(p - 1)p^{m-1}$ and $c_p(p^m) = p^{m-1}((p - 1)m + p)$.

3. If $n = \sum_{i \ge 0} d_i p^i$ is the expansion of n in base p, then $c_p(n) = \sum_{i \ge 0} d_i\, p^{i-1}((p - 1)i + p)$.

Proof. (1) Among $k \le pn$, the $(p - 1)n$ integers prime to p contribute 1 each, and $k = pk_1$ with $k_1 \le n$ contributes $p \cdot p^{v_p(k_1)}$. (2) For $0 \le v \le m - 1$, exactly $(p - 1)p^{m-1-v}$ integers $k \in [1, p^m - 1]$ have $v_p(k) = v$, which gives $c_p(p^m - 1)$, and $c_p(p^m) = c_p(p^m - 1) + p^m$. (3) Let $m \ge 0$, $1 \le d \le p - 1$ and $0 \le s < p^m$. For $0 \le t \le d - 1$ and $1 \le k \le p^m - 1$, $v_p(tp^m + k) = v_p(k)$, and $v_p(tp^m) = m$ for $1 \le t \le d$; hence $c_p(dp^m + s) = d(c_p(p^m - 1) + p^m) + c_p(s) = d\,c_p(p^m) + c_p(s)$, with $c_p(0) = 0$. Induction on the number of nonzero digits and (2) give (3). □

For $p = 2$, part (1) is the recurrence $c(2n) = 2c(n) + n$ of [1], and part (3) is Lemma 2.1. Put $t_i = p\,c_p(p^i) = p^i((p - 1)i + p)$ for $i \ge 0$, and let $\mathcal{E}_K^{(p)}$ be the set of sums $\sum_{0 \le i < K} d_i t_i$ with $0 \le d_i \le p - 1$.

Theorem 9.3. *1. For $a \ge 0$, $a \in \mathcal{A}_p$ if and only if $p^{a+1} = \sum_{i \ge 0} d_i t_i$ with $0 \le d_i \le p - 1$, not all zero; then $n = \sum_i d_i p^i$ satisfies $c_p(n) = p^a$.*

2. For every $m \ge 0$, $\sum_{0 \le i < m} (p - 1)t_i = m(p - 1)p^m < t_m$. Hence a representation $V = \sum_i d_i t_i$ with digits $0 \le d_i \le p - 1$ is unique, and its largest index K with $d_K \ne 0$ satisfies $d_K = \lfloor V/t_K \rfloor$.

3. *For $K \geq 0$ let $g(K)$ be the least integer $g \geq 1$ with $p^g \geq (p-1)(K+1)+1$. If $p^b = \sum_i d_i t_i$ has largest index K, then $b = K + g(K)$ and $d_K = \lfloor p^{g(K)}/((p-1)K+p) \rfloor \in [1, p-1]$. The map $K \mapsto K + g(K)$ is strictly increasing on $\{0, 1, 2, \dots\}$, and the integers $b \geq 1$ outside its image are exactly $b = (p^g-1)/(p-1) + g$, $g \geq 1$. Consequently, $a \notin \mathcal{A}_p$ for $a = (p^g-1)/(p-1) + g - 1$, $g \geq 1$, and every other $a \geq 0$ has a unique candidate K with $K + g(K) = a + 1$.*

Proof. (1) By Lemma 9.2(3), $p\, c_p(n) = \sum_i d_i t_i$, and every digit vector with digits in $[0, p-1]$, not all zero, is the expansion of a unique $n \geq 1$.

(2) By Lemma 9.2(2), $\sum_{i<m}(p-1)t_i = p\, c_p(p^m - 1) = m(p-1)p^m < p^m((p-1)m + p) = t_m$. If K is the largest index with $d_K \neq 0$, then $0 \leq V - d_K t_K \leq \sum_{i<K}(p-1)t_i < t_K$, so $d_K = \lfloor V/t_K \rfloor$; uniqueness follows by induction.

(3) Put $u = (p-1)(K+1)$ and $\gamma = b - K$. From $t_K \leq p^b \leq \sum_{i \leq K}(p-1)t_i = (K+1)(p-1)p^{K+1}$ we get $u + 1 = (p-1)K + p \leq p^\gamma \leq pu$. The interval $[u+1, pu]$ contains exactly one power of p: if g is least with $p^g \geq u+1$, then $g \geq 1$ because $u + 1 \geq p$, and $p^{g-1} \leq u$, so $p^g \leq pu$; and $p^{g+1} > pu$. Hence $\gamma = g(K)$. Since $p^{g(K)} \leq pu < p((p-1)K+p)$ and $p^{g(K)} \geq (p-1)K + p$, part (2) gives $1 \leq d_K \leq p-1$. Next, u increases by $p-1$ when K increases by 1, so $g(K+1) - g(K) \in \{0, 1\}$, and $g(K+1) = g(K) + 1$ exactly when $p^{g(K)} < (p-1)(K+2) + 1$, that is $p^{g(K)} - 1 \in [u, u+p-2]$. Both $p^g - 1$ and u are multiples of $p-1$, so this means $p^{g(K)} - 1 = u$, that is $K+1 = (p^g-1)/(p-1)$ with $g = g(K)$. Hence $K \mapsto K + g(K)$ increases by 1 at each step except from $K = (p^g-1)/(p-1) - 1$, where $g(K) = g$, to $K+1$, where it increases by 2 and skips $b = K + g + 1 = (p^g-1)/(p-1) + g$; every $g \geq 1$ occurs, and $0 + g(0) = 1$. The last statement follows from (1) with $b = a + 1$. □

For $p = 2$, the change of index $j = i + 2$ turns t_i into $j2^j/4$; the candidate of Theorem 9.3 is the candidate of Theorem 3.3 minus 2, and the excluded exponents are $2^g + g - 2$. The recursion of Theorem 4.1 extends to every prime as follows. A *p-state* is a pair (K, w) of integers with $K \geq 0$ and $0 \leq w \leq (p-1)K$; it is *successful* if $p^K w \in \mathcal{E}_K^{(p)}$.

Theorem 9.4. 1. *Let $a \geq 0$ have the candidate K, put $g = a + 1 - K$, $u = (p-1)K + p$, $d = \lfloor p^g/u \rfloor$ and $w_0 = p^g - du$. Then $a \in \mathcal{A}_p$ if and only if $w_0 \leq (p-1)K$ and the p-state (K, w_0) is successful.*

2. *A p-state $(K, 0)$ is successful. Let (K, w) be a p-state with $w \geq 1$, so $K \geq 1$; let s be the least integer $\sigma \geq 1$ with $p^\sigma w \geq (p-1)(K-\sigma) + p$, put $i = K - s$, $u_i = (p-1)i + p$, $d = \lfloor p^s w/u_i \rfloor$ and $w' = p^s w - du_i$. Then $1 \leq s \leq K$, and (K, w) is successful if and only if $d \leq p-1$, $w' \leq (p-1)i$, and the p-state (i, w') is successful.*

Proof. (1) By Theorem 9.3, $a \in \mathcal{A}_p$ if and only if $p^{a+1} = dt_K + R$ with $R \in \mathcal{E}_K^{(p)}$, where $d = \lfloor p^{a+1}/t_K \rfloor$; here $R = p^K(p^g - du) = p^K w_0$, and $R \leq \max \mathcal{E}_K^{(p)} = (p-1)Kp^K$ means $w_0 \leq (p-1)K$.

(2) Let $V = p^K w \geq 1$. For $0 \leq i' < K$, $t_{i'} \leq V$ is equivalent to $(p-1)i' + p \leq p^{K-i'}w$, so $i = K - s$ is the largest index below K with $t_i \leq V$; $s \leq K$ because $p^K w \geq p$. By Theorem 9.3(2), a representation of V in $\mathcal{E}_K^{(p)}$ has largest index i and digit $\lfloor V/t_i \rfloor = d$: indeed its largest index m satisfies $t_m \leq V$, so $m \leq i$, and $m < i$ is impossible because $V \leq \sum_{i' \leq m}(p-1)t_{i'} < t_{m+1} \leq t_i \leq V$. Hence $V \in \mathcal{E}_K^{(p)}$ if and only if $d \leq p-1$ and $V - dt_i = p^i w' \in \mathcal{E}_i^{(p)}$; the latter requires $p^i w' \leq (p-1)ip^i$, that is $w' \leq (p-1)i$. □

Since $s \geq 1$, the index decreases at each step, so the recursion ends after at most K steps; by the minimality of s, $p^{s-1}w < (p-1)(i+1) + p$, so all the integers involved are $O(pK)$. For an odd prime, the condition $w_0 \leq (p-1)K$ in part (1) can fail, which never happens for $p = 2$; this gives a family of non-members with no analogue for $p = 2$.

Theorem 9.5. *Let p be a prime.*

1. *For every $g \geq 1$, $K = (p^g - p)/(p-1)$ satisfies $c_p(p^K) = p^{K+g-1}$; thus $K + g - 1 \in \mathcal{A}_p$.*

2. *If p is odd, then $c_p(p^m - 1)$ is not a power of p for any $m \geq 1$.*

3. *If p is odd and $g \geq 2$ is even, then $K = (p^g - 2p + 1)/(2(p-1))$ is a nonnegative integer with candidate exponent $K + g - 1$, and $K + g - 1 \notin \mathcal{A}_p$.*

Consequently, for every odd prime p, $A_p(x) \geq \log_p x - O(1)$ and $B_p(x) \geq \frac{3}{2}\log_p x - O(1)$.

Proof. (1) $(p-1)K+p=p^g$, so $c_p(p^K)=p^{K-1}p^g$ by Lemma 9.2(2). (2) $c_p(p^m-1)=m(p-1)p^{m-1}$, and $m(p-1)$ is not a power of p because $p-1\ge 2$ divides it and is prime to p. (3) Since $p^2-1=(p-1)(p+1)$ and $p+1$ is even, $p^2\equiv 1 \pmod{2(p-1)}$, so $p^g\equiv 1$ and $p^g-2p+1\equiv 2-2p\equiv 0 \pmod{2(p-1)}$; and $p^g\ge p^2>2p-1$, so $K\ge 0$. Put $u=(p-1)K+p=(p^g+1)/2$. Then $u=(p-1)(K+1)+1\le p^g$ and $p^g\le p(u-1)=p(p^g-1)/2$ because $p\ge 3$, so $g(K)=g$ by the proof of Theorem 9.3(3), and $K+g-1$ has candidate K. Moreover $d=\lfloor p^g/u\rfloor=\lfloor 2p^g/(p^g+1)\rfloor=1$ and $w_0=p^g-u=u-1=(p-1)K+p-1>(p-1)K$, so $K+g-1\notin\mathcal{A}_p$ by Theorem 9.4(1).

For the counting statement, the exponents of (1), of (3) and of Theorem 9.3(3) are at most $p^g/(p-1)+g$ for the corresponding g, so those with $g\le \log_p x-1$ are at most x for large x; there are $\log_p x-O(1)$ values of g in (1) and in Theorem 9.3(3), and $\frac{1}{2}\log_p x-O(1)$ even values in (3). Members and non-members are distinct, the exponents of Theorem 9.3(3) have no candidate, and distinct g give distinct K and hence, by Theorem 9.3(3), distinct exponents. □

For $p=3$, Theorem 9.5 gives $c_3(27)=81$, so $4\in\mathcal{A}_3$, and $2\notin\mathcal{A}_3$; Theorem 9.3(3) gives $1\notin\mathcal{A}_3$; and $c_3(13)=c_3(1)+c_3(3)+c_3(9)=1+5+21=27$, so $3\in\mathcal{A}_3$.

The exponents rejected by the condition $w_0\le(p-1)K$ of Theorem 9.4(1), which we call *immediate failures*, admit a complete description.

Theorem 9.6. *Let $a\ge 0$ have the candidate K, and put $g=a+1-K$ and $u=(p-1)K+p$. Then a is an immediate failure if and only if $u\mid p^g+j$ for some $j\in\{1,\dots,p-1\}$; in that case $e=(p^g+j)/u$ satisfies $2\le e\le p-1$. Consequently, there are no immediate failures for $p=2$; for $p=3$ the immediate failures are exactly the exponents of Theorem 9.5(3); and for every prime p the immediate failures are the exponents $K+g-1$ obtained from the triples (g,e,j) with $2\le e\le p-1$, $1\le j\le p-1$, $e\mid p^g+j$, $u=(p^g+j)/e\equiv 1 \pmod{p-1}$, $u\ge p$ and $u\le p^g\le p(u-1)$, through $K=(u-p)/(p-1)$. In particular, there are at most $(p-1)(p-2)$ immediate failures with a given g.*

Proof. By the proof of Theorem 9.3(3), $u\le p^g\le p(u-1)$, $d=\lfloor p^g/u\rfloor\in[1,p-1]$ and $0\le w_0=p^g-du<u$. Hence $w_0>(p-1)K=u-p$ if and only if $w_0=u-j$ with $1\le j\le p-1$, that is $p^g+j=(d+1)u$, which is $u\mid p^g+j$ because $0\le w_0<u$ determines d. Then $e=d+1\ge 2$, and $e\le p-1$ because $p^g+j\le p(u-1)+p-1<pu$. For $p=2$ the range $2\le e\le p-1$ is empty. Conversely, given a triple as in the statement, $K=(u-p)/(p-1)$ is a nonnegative integer with $(p-1)K+p=u$, the inequalities $u\le p^g\le p(u-1)$ say that $g=g(K)$, so $K+g-1$ has candidate K, and $p^g=eu-j$ with $1\le j\le p-1<u$ gives $d=e-1$ and $w_0=u-j>u-p$. For $p=3$, $e=2$ and $j\in\{1,2\}$; $2\mid 3^g+j$ forces $j=1$, and $u=(3^g+1)/2$ is odd, that is $u\equiv 1 \pmod 2$, exactly when $3^g\equiv 1 \pmod 4$, that is g even; this is Theorem 9.5(3). For given g, each pair (e,j) gives at most one u, hence at most one exponent. □

The sets of successful p-states satisfy the analogues of Proposition 8.2 and Theorem 8.6. For $K\ge 0$ let $S_K^{(p)}$ be the set of w such that (K,w) is a successful p-state, and $N_p(K)=|S_K^{(p)}|$; let $u_K=(p-1)K+p$.

Theorem 9.7. *1. $S_0^{(p)}=\{0\}$, and for every $K\ge 0$, $S_{K+1}^{(p)}=\{(w+du_K)/p:\ w\in S_K^{(p)},\ 0\le d\le p-1,\ p\mid w+du_K\}$, the elements obtained from distinct pairs (w,d) being distinct.*

2. If $p\nmid K$, then for each $w\in S_K^{(p)}$ exactly one $d\in\{0,\dots,p-1\}$ satisfies $p\mid w+du_K$, and $N_p(K+1)=N_p(K)$.

3. If $p\mid K$, then $S_{K+1}^{(p)}=\bigcup_{d=0}^{p-1}(B+d\,u_K/p)$ with $B=\{w/p:\ w\in S_K^{(p)},\ p\mid w\}$, the p sets being disjoint; thus $N_p(K+1)=p\,\#\{w\in S_K^{(p)}:\ p\mid w\}$.

4. For every $K\ge 0$, $w\in S_K^{(p)}$ if and only if $(p-1)K-w\in S_K^{(p)}$.

An exponent a with candidate K belongs to $\mathcal{A}_p$ if and only if $w_0\le(p-1)K$ and $w_0\in S_K^{(p)}$, with w_0 as in Theorem 9.4(1).

Proof. (1) $\mathcal{E}_0^{(p)}=\{0\}$ gives $S_0^{(p)}=\{0\}$. By Theorem 9.3(2), $p^{K+1}w'\in\mathcal{E}_{K+1}^{(p)}$ if and only if $p^{K+1}w'=dt_K+p^Kw$ with $0\le d\le p-1$ and $p^Kw\in\mathcal{E}_K^{(p)}$; since $t_K=p^Ku_K$, this is $pw'=w+du_K$, and $0\le w\le(p-1)K$ holds for every element p^Kw of $\mathcal{E}_K^{(p)}$. Distinct pairs (w,d) give distinct representations, hence distinct w' by Theorem 9.3(2). (2) $u_K\equiv(p-1)K\equiv -K \pmod p$ is invertible modulo p when $p\nmid K$. (3) If $p\mid K$, then $p\mid u_K$, and $p\mid w+du_K$ is $p\mid w$ for all d; the sets are disjoint because $B\subseteq[0,(p-1)K/p]$ and $u_K/p=(p-1)K/p+1$. (4) By Lemma 9.2(2), $\sum_{i<K}(p-1)t_i=(p-1)Kp^K$; replacing every digit d_i by $p-1-d_i$ maps a representation of p^Kw to one of $p^K((p-1)K-w)$. The last statement is Theorem 9.4(1). □

For $p = 2$, Theorem 9.7 is Theorem 8.6 and Proposition 8.2 after the change of index of Section 9.

10. Experiments

All computations of this section were made by the script `powers_of_two_c.py`, which was written with Claude Opus 5.5 and run by the author with Python 3.12.3, with the parameters 2^{20} for the range of the check of Lemma 2.1, 2^{24} for the brute-force search, 4000 for the big-integer check, $x \le 20000$ for the counting functions, $h \le 300$ for the closed-form checks of the families, $h \le 18$ for the runs of Algorithm 1 on the families, $J \le 65536$ for the sets S_J computed with the recursion of Theorem 4.1, and $J \le 131072$ for the sets S_J computed with the recursion of Theorem 8.6, with NumPy 2.4.6, $a \le 3000$ for the recursion of Theorem 9.4 with the primes $p = 3, 5, 7, 11$, $a \le 100000$ for the immediate failures and $K \le 10000$ for the recursion of Theorem 9.7; the whole run took 222.5 seconds. The script requires Python 3.12 and the Python package NumPy; the run reported here used Python 3.12.3 and NumPy 2.4.6. NumPy is used only for the sets S_J (Sections 10.4–10.6); without it the script runs in pure Python with $J \le 8192$ and $J \le 16384$ for the two recursions. The script uses exact integer arithmetic only. It decides the membership of an exponent a in $\mathcal{A}$ by three independent methods: a brute-force search computing $c(n)$ from the definition (1); the greedy algorithm applied to the integers 2^{a+3} and $j2^j$ themselves; and Algorithm 1. The script performs 121 checks, and all of them passed.

10.1. Verification of the proved statements

The identity (4) holds for $1 \le n \le 2^{20}$, the formula $c(2^r) = 2^{r-1}(r+2)$ holds for $0 \le r < 200$, and $c(1), \dots, c(16)$ agree with the table of Section 7 of [1]. The brute-force search over $1 \le n \le 2^{24}$ reaches $c(2^{24}) = 218103808$, so it decides the membership of every exponent $0 \le a \le 27$. Table 1 lists the exponents found, with the integer n such that $c(n) = 2^a$. For all $0 \le a \le 27$, the brute-force search, the big-integer greedy algorithm and Algorithm 1 return the same answer, and for the exponents in $\mathcal{A}$ the integer $n = \sum_{j \in T} 2^{j-2}$ of Corollary 4.2 equals the integer found by the search.

Table 1: The exponents $0 \le a \le 27$ with $2^a \in c(\mathbb{N})$, and the integer n with $c(n) = 2^a$, found by the brute-force search over $n \le 2^{24}$. The other exponents $a \le 27$ are $1, 4, 6, 7, 9, 18, 21, 23, 24, 26, 27$.

a	n	a	n	a	n	a	n
0	1	10	255	14	2624	19	65535
2	3	11	447	15	4864	20	122111
3	4	12	768	16	8960	22	433407
5	15	13	1471	17	16384	25	3023616
8	64						

For $0 \le a \le 4000$, Algorithm 1 and the big-integer greedy algorithm return the same answer, the set T produced by Algorithm 1 is the set produced by the greedy algorithm, and for the exponents $a \le 1500$ in $\mathcal{A}$ the integer n of Corollary 4.2 satisfies $c(n) = 2^a$. The values $3 \le b \le 5000$ not taken by β are $4, 7, 12, 21, 38, 71, 136, 265, 522, 1035, 2060, 4109$, that is $2^r + r + 1$ for $1 \le r \le 12$, as stated in Theorem 3.3, and $0 \le 2^h - K \le K - 2$ holds for $2 \le K < 2^{16}$. For $0 \le a \le 20000$, Algorithm 1 performs at most $K - 2$ iterations and handles integers at most $2K$ in its loop, as stated in Corollary 4.2; the largest number of iterations is 19781.

For Theorem 5.1(2), the 295 pairs (k, h) with $h \le 300$ give integers K with $2^{h-1} < K < 2^h$, $K - s \ge 2$ and $K(2^s + 1) - s = 2^{h+s}$, which is the identity used in the proof of (7); the equality (7) itself was evaluated with Lemma 2.1 for $h \le 16$. For Theorem 6.1, the 310 pairs (s, h) with $h \le 300$ give integers K with $2^{h-1} < K < 2^h$, the least σ of Theorem 4.1 equals s, and $2^{s-1}w + s = J + 1$. Table 2 lists the pairs with $h \le 18$; on each of them Algorithm 1 returns the answer of the theorem, a member for Theorem 5.1 and a failure at the first iteration for Theorem 6.1, and the big-integer greedy algorithm agrees for $a \le 4000$. The exponents $a + (2^a - 11)/3$ for odd $3 \le a \le 15$ are in $\mathcal{A}$, the exponents $a + (2^a - 10)/3$ for even $4 \le a \le 14$ are not, and the exponents $2^r + r - 2$, $1 \le r < 40$, have no candidate K.

Table 2: The families of Theorems 5.1(2) and 6.1 with $h \le 18$: the parameter s, the integer K and the exponent $a = K + h - 3$.

Members (Theorem 5.1)				Non-members (Theorem 6.1)			
s	h	K	a	s	h	K	a
1	2	3	2	2	3	6	6
1	4	11	12	2	5	22	24
1	6	43	46	2	7	86	90
1	8	171	176	2	9	342	348
1	10	683	690	2	11	1366	1374
1	12	2731	2740	2	13	5462	5472
1	14	10923	10934	2	15	21846	21858
1	16	43691	43704	2	17	87382	87396
1	18	174763	174778	4	8	228	233
2	5	26	28	4	14	14564	14575
2	9	410	416	8	17	130056	130070
2	13	6554	6564	3	3	7	7
2	17	104858	104872	3	7	103	107
4	10	964	971	3	11	1639	1647
4	18	246724	246739	3	15	26215	26227

10.2. The counting functions

Table 3 gives $A(x)$ and $B(x)$ for $x = 1000, 2000, \ldots, 20000$, computed with Algorithm 1. The exponent $a = 0$ belongs to $\mathcal{A}$ because $c(1) = 1 = 2^0$, and the number $A(x) - 1$ of exponents $1 \le a \le x$ in $\mathcal{A}$ coincides with the table of Nicolas [1] at the ten values $x = 1000, \ldots, 10000$. In the range $x \le 20000$ the ratio $A(x)/x$ takes values between 0.5132, at $x = 6000$, and 0.7418, at $x = 20000$, and $B(x)$ increases by 42 between $x = 15000$ and $x = 20000$.

Table 3: The counting functions $A(x)$ and $B(x)$ of (2), and the values of the table of Nicolas [1], which count the exponents $1 \le a \le x$.

x	$A(x)$	$B(x)$	$A(x)/x$	Nicolas	x	$A(x)$	$B(x)$	$A(x)/x$
1000	636	365	0.6360	635	11000	6664	4337	0.6058
2000	1375	626	0.6875	1374	12000	7332	4669	0.6110
3000	1788	1213	0.5960	1787	13000	8133	4868	0.6256
4000	2406	1595	0.6015	2405	14000	8939	5062	0.6385
5000	2749	2252	0.5498	2748	15000	9878	5123	0.6585
6000	3079	2922	0.5132	3078	16000	10867	5134	0.6792
7000	3647	3354	0.5210	3646	17000	11855	5146	0.6974
8000	4381	3620	0.5476	4380	18000	12847	5154	0.7137
9000	5228	3773	0.5809	5227	19000	13842	5159	0.7285
10000	5964	4037	0.5964	5963	20000	14836	5165	0.7418

Among the 5165 exponents $0 \le a \le 20000$ outside $\mathcal{A}$, 14 have no candidate K (the exponents $2^r + r - 2$, $1 \le r \le 14$), 11 are rejected by Algorithm 1 at its first iteration, and 5140 at a later iteration. The script checks that the 11 exponents rejected at the first iteration are exactly the 11 exponents $a \le 20000$ given by Theorem 6.1, all listed in Table 2: $6, 24, 90, 348, 1374, 5472$ ($s = 2$), $233, 14575$ ($s = 4$) and $7, 107, 1647$ ($s = 3$). It also checks that the 14 exponents $a \le 20000$ in $\mathcal{A}$ decided without any iteration, $0, 3, 8, 17, 34, 67, 132, 261, 518, 1031, 2056, 4105, 8202, 16395$, are exactly those of Theorem 5.1(1), and that the 11 exponents $a \le 20000$ in $\mathcal{A}$ decided after exactly one iteration, $2, 12, 28, 46, 176, 416, 690, 971, 2740, 6564, 10934$, are exactly those of Theorem 5.1(2). These three checks confirm Theorem 6.2 for $a \le 20000$.

Table 4 gives the number of iterations of Algorithm 1 for the 19987 exponents $0 \le a \le 20000$ that have a candidate K. Among them, 14751 members and 5072 non-members require at least 10 iterations, that is 99.2% of them, and the

largest number of iterations is 19781.

Table 4: Number of iterations of Algorithm 1 for the exponents $0 \le a \le 20000$ with a candidate K; the last row counts the exponents with at least 10 iterations.

iterations	members	non-members	iterations	members	non-members
0	14	0	5	6	8
1	11	11	6	8	8
2	7	9	7	12	9
3	6	4	8	7	10
4	9	10	9	5	10
			≥ 10	14751	5072

Table 5 gives, for each $h \le 14$, the number of candidates $K \in (2^{h-1}, 2^h]$, which is 2^{h-1}, and the number of them for which $\beta(K) - 3 \in \mathcal{A}$. For $5 \le h \le 14$ the proportion takes values between 0.5000, at $h = 6$, and 0.8198, at $h = 14$.

Table 5: Candidates $K \in (2^{h-1}, 2^h]$ and the number of them with $\beta(K) - 3 \in \mathcal{A}$.

h	candidates	members	proportion	h	candidates	members	proportion
1	1	1	1.0000	8	128	103	0.8047
2	2	2	1.0000	9	256	192	0.7500
3	4	2	0.5000	10	512	277	0.5410
4	8	8	1.0000	11	1024	744	0.7266
5	16	10	0.6250	12	2048	1054	0.5146
6	32	16	0.5000	13	4096	2087	0.5095
7	64	45	0.7031	14	8192	6716	0.8198

10.3. The bounds of Theorem 7.1

Table 6 compares, for $H = \lfloor \log_2 x \rfloor - 1$, the number of exponents produced by the families in the proof of Theorem 7.1 with the explicit lower bound $2H - \log_2 H - 3$ for the members, and with $A(x)$ and $B(x)$ when $x \le 20000$. For every x of the table the family exponents are at most x, the member and non-member families are disjoint, the number of member exponents is at least $2H - \log_2 H - 3$, and for $x \le 20000$ every member exponent is in $\mathcal{A}$ and every non-member exponent is not. For $x \le 10000$ the families account for at most 21 of the $A(x)$ members and 21 of the $B(x)$ non-members.

Table 6: Number of family exponents up to x (members of Theorem 5.1, non-members of Theorems 3.3 and 6.1), the bound $2H - \log_2 H - 3$ of the proof of Theorem 7.1, and the exact $A(x)$ and $B(x)$ for $x \le 10000$.

x	H	members (families)	$2H - \log_2 H - 3$	$A(x)$	non-members (families)	$B(x)$
16	3	4	1.42	12	5	5
64	5	8	4.68	38	8	27
256	7	11	8.19	182	12	75
1000	8	13	10.00	636	14	365
4096	11	19	15.54	2449	20	1648
10000	12	21	17.42	5964	21	4037
65536	15	26	23.09	–	28	–
10^6	18	33	28.83	–	33	–
10^9	28	52	48.19	–	53	–
10^{12}	38	73	67.75	–	74	–

10.4. The successful states

The script computes the sets S_J for all $J \le 65536$, row by row, from Theorem 4.1: $w \in S_J$ if and only if $w = 0$, or the iteration from (J, w) does not fail and its result $(J - s, w')$ satisfies $w' \in S_{J-s}$. The values of $N(J)$ obtained in this way agree with Algorithm 1 for $J < 200$ and for ten random J in $[200, 1500]$, and Algorithm 1 agrees with the big-integer greedy algorithm on all states with $J < 120$. The symmetry of Proposition 8.2 holds for every $J \le 65536$. The values of $N(J)$ printed by the script, for $J \le 64$ and near each power of 2 up to 2^{16}, satisfy Proposition 8.3. Table 7 gives the extreme and mean values of $N(J)/J$ on the dyadic ranges. The integers $J \le 65536$ with $N(J) = J$, that is with every state successful, are $1, 2, 4, 12, 16, 17356, 17360, 17376, 17632$.

Table 7: The minimum, maximum and mean of $N(J)/J$ for $2^m \le J < 2^{m+1}$, $0 \le m \le 15$, with the integers J where the minimum and maximum are attained.

m	min $N(J)/J$	at J	max $N(J)/J$	at J	mean $N(J)/J$
0	1.0000	1	1.0000	1	1.0000
1	0.6667	3	1.0000	2	0.8333
2	0.5714	7	1.0000	4	0.7595
3	0.4444	9	1.0000	12	0.7565
4	0.5161	31	1.0000	16	0.7090
5	0.4364	55	0.9412	34	0.6160
6	0.4507	71	0.9355	124	0.7170
7	0.5561	187	0.9846	130	0.8218
8	0.5727	447	0.9927	274	0.7646
9	0.2976	699	0.7489	876	0.5527
10	0.3962	2019	0.9245	1298	0.7189
11	0.2440	2787	0.7865	3260	0.5317
12	0.1846	5027	0.8713	8098	0.5086
13	0.5205	9899	0.9999	15362	0.8212
14	0.8215	26105	1.0000	17632	0.9672
15	0.7311	39643	0.9937	51332	0.8992

By Proposition 8.1(2), the sets S_J give the members of $\mathcal{A}$ with candidate $K \le 65537$, hence $A(x)$ for $x \le 65551 = \beta(65537) - 3$. The value $A(20000) = 14836$ obtained in this way agrees with Section 10.2. Table 8 gives, for $h \le 16$, the proportion of members among the candidates $K \in (2^{h-1}, 2^h]$ and the mean of $N(K-1)/(K-1)$ over the same K; for $7 \le h \le 16$ the two numbers differ by at most 0.0171. Table 9 gives $A(x)$ for $x = 25000, 30000, \ldots, 65000$. For $x = 1000, 2000, \ldots, 65000$, the ratio $A(x)/x$ takes values between 0.5132, at $x = 6000$, and 0.8654, at $x = 62000$.

Table 8: Members among the candidates $K \in (2^{h-1}, 2^h]$, and the mean of $N(J)/J$ over $J = K - 1$ for the same K.

h	candidates	members	proportion	mean N/J	h	candidates	members	proportion	mean N/J
1	1	1	1.0000	1.0000	9	256	192	0.7500	0.7646
2	2	2	1.0000	0.8333	10	512	277	0.5410	0.5527
3	4	2	0.5000	0.7595	11	1024	744	0.7266	0.7189
4	8	8	1.0000	0.7565	12	2048	1054	0.5146	0.5317
5	16	10	0.6250	0.7090	13	4096	2087	0.5095	0.5086
6	32	16	0.5000	0.6160	14	8192	6716	0.8198	0.8212
7	64	45	0.7031	0.7170	15	16384	15878	0.9691	0.9672
8	128	103	0.8047	0.8218	16	32768	29503	0.9004	0.8992

Table 9: The counting function $A(x)$ for $25000 \le x \le 65000$, obtained from the sets S_J.

x	$A(x)$	$A(x)/x$	x	$A(x)$	$A(x)/x$
25000	19692	0.7877	50000	42335	0.8467
30000	24433	0.8144	55000	47180	0.8578
35000	29021	0.8292	60000	51819	0.8637
40000	33160	0.8290	65000	56153	0.8639
45000	37737	0.8386			

10.5. The forest of states

The script checks Theorem 8.4 as follows. The number $F(J)$ of failing states obtained from the rows equals the number given by part (1) for every $J \le 65536$, and the largest value of $F(J)$ in this range is 3. The sum $\sum_{J \le L} F(J)$ equals $\sum_{\sigma=2}^{L} \lfloor (L+1-\sigma)/2^{\sigma-1} \rfloor$ for $L = 10, 100, 1000, 10000, 65536$, with the values $6, 94, 991, 9991, 65519$. Every state (J', w') with $J' \le 600$ has exactly $v_2(J'+1+w')$ preimages under Φ, counted by applying Φ to all states of levels up to 613, and the recursion of part (3) for $N(J)$ holds for $J \le 1200$.

The script labels every state by the root of its tree and counts, at the levels $J = 2^m$, $5 \le m \le 16$, the number of distinct trees among the J states of the level and the shares of the largest trees (Table 10). At each of these levels the largest tree is the tree of the root $(1, 0)$, and its share equals $N(J)/(2J)$, as stated by Proposition 8.5. At the level $J = 65536$, the ten largest trees are those of the roots $(1, 0)$, $(47321, 0)$, $(51081, 0)$, $(47257, 0)$ and $(9631, 0)$, with 29090, 16376, 6838, 2810 and 1962 states, and those of the failing states of the levels 61364 ($s = 5$), 61361 ($s = 2$), 52582 ($s = 3$), 52581 ($s = 2$) and 56327 ($s = 8$), with 1614, 1614, 1578, 1578 and 386 states.

Table 10: The trees of the forest at the levels $J = 2^m$: the number of distinct trees among the J states of the level, the proportion $N(J)/J$ of successful states, and the shares of the largest tree, which is the tree of $(1, 0)$ at every level, and of the 5 and 10 largest trees.

J	trees	$N(J)/J$	tree of $(1, 0)$	5 largest	10 largest
32	8	0.6250	0.3125	0.8750	1.0000
64	12	0.5625	0.2812	0.7188	0.9688
128	9	0.9375	0.4688	0.9531	1.0000
256	15	0.8438	0.4219	0.8047	0.9531
512	17	0.7422	0.3711	0.7188	0.9102
1024	20	0.5469	0.2734	0.7422	0.9355
2048	21	0.5293	0.2646	0.7783	0.9521
4096	25	0.4668	0.2334	0.6982	0.8931
8192	22	0.8335	0.4167	0.7751	0.9609
16384	17	0.9868	0.4934	0.9744	0.9945
32768	19	0.9535	0.4767	0.9462	0.9908
65536	27	0.8878	0.4439	0.8709	0.9742

10.6. The bottom-up recursion

The script computes S_J for $J \le 131072$ with Theorem 8.6, in 27.5 seconds. The sets obtained coincide with those of Section 10.4 for every $J \le 3000$; the values of $N(J)$ coincide for $J \le 65536$, and so does the membership of every initial state with $K \le 65537$. For every $J \le 131072$, S_J is symmetric, S_J has period $J/2$ when J is even, $N(2k+1) = N(2k)$, and $4 \mid N(J)$ when $J \ge 4$. For the block $h = 17$, the proportion of members among the 65536 candidates is $56692/65536 = 0.8651$, and the mean of $N(K-1)/(K-1)$ over the block is 0.8652. On the range $2^{16} \le J < 2^{17}$, $N(J)/J$ takes values between 0.6683, at $J = 70695$, and 0.9996, at $J = 95306$, with mean 0.8652. Table 11 gives $A(x)$ for $70000 \le x \le 130000$; the value $A(65000) = 56153$ coincides with Table 9. For $x = 1000, 2000, \ldots, 131000$, the ratio $A(x)/x$ takes values between 0.5132, at $x = 6000$, and 0.8750, at $x = 109000$.

Table 11: The counting function $A(x)$ for $70000 \le x \le 130000$, obtained from the recursion of Theorem 8.6.

x	$A(x)$	$A(x)/x$	x	$A(x)$	$A(x)/x$
70000	60188	0.8598	105000	91604	0.8724
75000	63912	0.8522	110000	96233	0.8748
80000	68038	0.8505	115000	100429	0.8733
85000	72221	0.8497	120000	104238	0.8687
90000	76886	0.8543	125000	108240	0.8659
95000	81814	0.8612	130000	112452	0.8650
100000	86767	0.8677			

10.7. The families of Theorem 8.7

The formula $c(2^{2^r}-1) = 2^{2^r+r-1}$ holds for $0 \le r \le 12$. For $h \le 300$, Theorem 8.7(2) gives 282 pairs (i,h) and Theorem 8.7(3) gives 222 pairs (i,h); for all of them w is an even positive integer and $2^{h-1}+1 < K < 2^h$. Every exponent of parts (1) and (2) at most $131088 = \beta(131073) - 3$ is in $\mathcal{A}$, every exponent of part (3) in this range is not, according to the sets S_J of Section 10.6, and the big-integer greedy algorithm agrees for $a \le 4000$. Table 12 lists the exponents of parts (2) and (3) with $h \le 17$.

Table 12: The exponents $a = K + h - 3$ of Theorem 8.7(2) (members) and (3) (non-members) with $h \le 17$, with $\sigma = 2^{i-1}+1$.

Members, part (2)					Non-members, part (3)				
i	σ	h	K	a	i	σ	h	K	a
1	2	4	10	11	2	3	5	19	21
2	3	6	35	38	2	3	8	147	152
1	2	7	74	78	3	5	8	133	138
3	5	9	261	267	2	3	11	1171	1179
1	2	10	586	593	3	5	13	4229	4239
2	3	10	547	554	4	9	13	4105	4115
1	2	13	4682	4692	2	3	14	9363	9374
2	3	14	8739	8750	2	3	17	74899	74913
4	9	14	8201	8212					
3	5	15	16645	16657					
1	2	16	37450	37463					

Table 13 gives the number of exponents up to x produced by all the families of Theorems 3.3, 5.1, 6.1 and 8.7, with $H = \lfloor \log_2 x \rfloor - 1$, together with $(3+C_1)H$, $(\frac{5}{4}+\frac{E}{2}+C_2)H$ and, for $x \le 100000$, the exact values of $A(x)$ and $B(x)$. For every x of the table, the member and non-member exponents are at most x and disjoint, and for $x \le 100000$ every member exponent is in $\mathcal{A}$ and every non-member exponent is not.

Table 13: Exponents up to x produced by all the explicit families, compared with $(3+C_1)H$ and $(\frac{5}{4}+\frac{E}{2}+C_2)H$ of Theorem 8.8, and with $A(x)$ and $B(x)$.

x	H	members	$(3+C_1)H$	$A(x)$	non-members	$(\frac{5}{4}+\frac{E}{2}+C_2)H$	$B(x)$
10^3	8	22	31.72	636	17	22.54	365
10^4	12	37	47.59	5964	25	33.81	4037
10^5	15	49	59.48	86767	35	42.27	13234
10^6	18	61	71.38	–	42	50.72	–
10^9	28	100	111.04	–	69	78.90	–
10^{12}	38	139	150.69	–	97	107.08	–

10.8. Odd primes

For $p = 3, 5, 7, 11$ the script checks Lemma 9.2(3) for $n \le 50000$ and Lemma 9.2(1) for $pn \le 200000$ against the definition of c_p, and Proposition 9.1 for $n \le 30$ by computing the product $(1 - X)\cdots(1 - X^n)$ in $\mathbb{F}_p[X]$ and dividing it by $1 - X$ as long as possible. A brute-force search over $n \le 200000$ decides the membership of the exponents $a \le 13, 8, 7, 5$ for $p = 3, 5, 7, 11$; the members found are $0, 3, 4, 6, 7, 8, 9, 10, 13$ for $p = 3$ (with $c_3(n) = 3^a$ for $n = 1, 13, 27, 215, 513, 1457, 3645, 9841, 196829$), $0, 5, 6, 8$ for $p = 5$, $0, 7$ for $p = 7$ and 0 for $p = 11$. In these ranges the recursion of Theorem 9.4 and the greedy algorithm applied to the big integers p^{a+1} and t_i agree with the search, and they agree with each other for $a \le 1200$. The exponents $a \le 20000$ without candidate are exactly those of Theorem 9.3(3): $1, 5, 15, 43, 125, 369, 1099, 3287, 9849$ for $p = 3$, $1, 7, 33, 159, 785, 3911, 19537$ for $p = 5$, $1, 9, 59, 403, 2805, 19613$ for $p = 7$ and $1, 13, 135, 1467, 16109$ for $p = 11$. The three parts of Theorem 9.5 are confirmed; part (3) gives the exponents $2, 22, 186, 1646, 14770, 132870$ for $p = 3$, $3, 80, 1957$ for $p = 5$, $4, 202, 9808$ for $p = 7$ and $6, 734, 88582$ for $p = 11$. Table 14 gives $A_p(x)$ for $x \le 3000$ and the outcomes of the recursion for $a \le 3000$.

Table 14: The counting functions $A_p(x)$ for $p = 3, 5, 7, 11$, and the outcomes of the recursion of Theorem 9.4 for $0 \le a \le 3000$: no candidate, immediate failure, a digit $d > p - 1$, a remainder $w' > (p - 1)i$, and success.

	$p = 3$	$p = 5$	$p = 7$	$p = 11$
$A_p(500)$	381	239	212	215
$A_p(1000)$	733	397	556	397
$A_p(1500)$	1105	616	896	577
$A_p(2000)$	1442	882	1173	804
$A_p(2500)$	1893	1252	1485	1100
$A_p(3000)$	2317	1676	1812	1334
$A_p(3000)/3000$	0.7723	0.5587	0.6040	0.4447
no candidate	7	5	5	4
immediate failure	4	7	6	10
digit $d > p - 1$	343	197	222	125
remainder $w' > (p - 1)i$	330	1116	956	1528
success	2317	1676	1812	1334

For $p = 3$ the 4 immediate failures up to 3000 are the exponents $2, 22, 186, 1646$ of Theorem 9.5(3), as Theorem 9.6 states. Grouping the candidates by g, the proportion of members is $0.6667, 0.7778, 0.7037, 0.8395, 0.6626, 0.7586$ for $g = 2, \ldots, 7$ when $p = 3$; $0.4000, 0.6400, 0.5200, 0.3952$ for $g = 2, \ldots, 5$ when $p = 5$; $0.2857, 0.6735, 0.4052, 0.6318$ for $g = 2, \ldots, 5$ when $p = 7$; and $0.1818, 0.6198, 0.3681$ for $g = 2, 3, 4$ when $p = 11$.

10.9. Immediate failures and the bottom-up recursion for odd primes

For $a \le 100000$, the script determines the immediate failures from the first step of Theorem 9.4 and compares them with the exponents given by the triples (g, e, j) of Theorem 9.6. For $p = 2$ there is none. For $p = 3, 5, 7, 11$ the two sets coincide; they are listed in Table 15. For $p = 3$ they are the exponents of Theorem 9.5(3), and for every p and g their number is at most $(p - 1)(p - 2)$; the largest number for a single g is $1, 3, 4, 6$ for $p = 3, 5, 7, 11$.

Table 15: The immediate failures $a \le 100000$ (Theorem 9.6).

p	immediate failures $a \le 100000$
2	none
3	2, 22, 186, 1646, 14770
5	2, 3, 41, 54, 80, 1306, 1957, 24420, 32558, 48834
7	2, 4, 20, 82, 102, 202, 3272, 4906, 6540, 9808
11	2, 3, 4, 6, 20, 185, 246, 368, 490, 734, 3224, 19688, 25312, 29530, 44293, 59056, 88582

The script computes the sets $S_K^{(p)}$ for $K \le 10000$ with Theorem 9.7, in less than 3 seconds for each prime. For every $K \le 10000$ and $p = 3, 5, 7, 11$, $S_K^{(p)}$ is symmetric under $w \mapsto (p-1)K - w$, $N_p(K+1) = N_p(K)$ when $p \nmid K$, and $N_p(K+1) = p\#\{w \in S_K^{(p)} : p \mid w\}$ when $p \mid K$; the members obtained from these sets coincide with those of Theorem 9.4 for $a \le 1500$. Table 16 gives $A_p(x)$ for $x \le 10000$. Table 17 groups the candidates K by $g = g(K)$ and compares the proportion of members with the mean of $N_p(K)/((p-1)K+1)$, the proportion of successful states, over the same K. For every complete group with at least 2000 candidates, namely $g = 8, 9$ for $p = 3$, $g = 6$ for $p = 5$ and $g = 5$ for $p = 7$, the two numbers differ by at most 0.0110.

Table 16: The counting functions $A_p(x)$ for $x \le 10000$, obtained from Theorem 9.7.

x	$A_3(x)$	$A_5(x)$	$A_7(x)$	$A_{11}(x)$	$A_3(x)/x$	$A_5(x)/x$	$A_7(x)/x$
2000	1442	882	1173	804	0.7210	0.4410	0.5865
4000	3082	2413	2348	1718	0.7705	0.6032	0.5870
6000	4664	3611	3406	2677	0.7773	0.6018	0.5677
8000	5684	4866	4764	3576	0.7105	0.6082	0.5955
10000	6518	6097	6244	4433	0.6518	0.6097	0.6244

Table 17: Candidates K grouped by $g = g(K)$: number of candidates, proportion of members, and mean of $N_p(K)/((p-1)K+1)$ over the group (complete groups only).

p	g	candidates	members	mean density	p	g	candidates	members	mean density
3	5	81	0.8395	0.7932	5	4	125	0.5200	0.5304
3	6	243	0.6626	0.6559	5	5	625	0.3952	0.4173
3	7	729	0.7586	0.7565	5	6	3125	0.6474	0.6513
3	8	2187	0.7796	0.7686	7	4	343	0.4052	0.4243
3	9	6561	0.6033	0.6005	7	5	2401	0.6318	0.6319
					11	4	1331	0.3681	0.3926

For $x = 1000, 2000, \ldots, 10000$, the ratio $A_p(x)/x$ takes values between 0.6518 and 0.7773 for $p = 3$, between 0.3970 and 0.6132 for $p = 5$, between 0.5560 and 0.6244 for $p = 7$, and between 0.3970 and 0.4486 for $p = 11$.

11. Conclusion

The identity $c(n) = n + \frac{1}{2}\sum_{i \in \mathcal{B}(n)} i2^i$ turns the question of Nicolas into the question of which powers of 2 are sums of distinct terms of the superincreasing sequence $(j2^j)_{j \ge 1}$ (Theorem 2.2). For every exponent there is at most one possible leading term, and the exponents without one are exactly $2^r + r - 2$ (Theorem 3.3); the remaining exponents are decided by a recursion on pairs of integers bounded by $a + 3$ (Theorem 4.1, Corollary 4.2). The closed formula $c(2^{K-2} + 2^{K-s-2}) = 2^{K+h-3}$ and the failure criterion of Theorem 4.1 give infinite explicit families of members and of non-members, which contain the families of Nicolas as the cases $s = 1$ and $s = 2$ (Theorems 5.1 and 6.1) and are exactly the exponents decided by at most one iteration of the recursion (Theorem 6.2), and they give the bounds $A(x) \ge (2 + o(1)) \log_2 x$ and $B(x) \ge (2.0533\ldots + o(1)) \log_2 x$ (Theorem 7.1). The experiments of Section 10 confirm these statements by three independent methods, reproduce the table of Nicolas, whose values are the numbers $A(x) - 1$ of exponents $1 \le a \le x$ in $\mathcal{A}$, and give $A(x)$ for $x \le 65000$, with $A(x)/x$ between 0.5132 and 0.8654 for $x = 1000, 2000, \ldots, 65000$. The successful states satisfy the symmetry $w \mapsto J - 1 - w$, $N(2k+1) = N(2k)$ and $4 \mid N(J)$ (Propositions 8.2 and 8.3), and for $7 \le h \le 16$ the proportion of members among the candidates $K \in (2^{h-1}, 2^h]$ differs from the mean of $N(K-1)/(K-1)$ by at most 0.0171. The states form a forest whose roots are the states $(J, 0)$ and the failing states, with $v_2(J' + 1 + w')$ children at each state (J', w') (Theorem 8.4); exactly half of the successful states of each level lie in the tree of $(1, 0)$ (Proposition 8.5), and at the levels $2^m \le 65536$ the ten largest trees contain between 89.31% and 100% of the states of the level. The sets S_J also satisfy a recursion from the smallest elements of T

upwards, with period $J/2$ for even J (Theorem 8.6); with this recursion $A(x)$ is computed for $x \le 131000$, and $A(x)/x$ lies between 0.5132 and 0.8750 for $x = 1000, 2000, \dots, 131000$. The symmetry gives further explicit families and the bounds $A(x) \ge (3.9655\ldots + o(1)) \log_2 x$ and $B(x) \ge (2.8178\ldots + o(1)) \log_2 x$ (Theorems 8.7 and 8.8). For every prime p, the function $c_p(n) = \sum_{k \le n} p^{v_p(k)}$, which counts the factor $1 - X$ in $(1 - X) \cdots (1 - X^n)$ over $\mathbb{F}_p[X]$, admits the same reduction, a unique candidate and a recursion, and for odd p there are infinitely many members and non-members (Theorems 9.3, 9.4 and 9.5); the immediate failures, which do not occur for $p = 2$, are described completely, and the bottom-up recursion and the symmetry extend to every prime (Theorems 9.6 and 9.7). The asymptotic behaviour of $A(x)$ asked for in [1] is not determined by these results.

Declaration of interests

The author declares no competing interests.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of generative AI in scientific writing

The author used Claude Opus 5.5 to assist in deriving the proofs, to write the experiment script `powers_of_two_c.py`, and in preparing this manuscript; the author ran the experiments. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Data availability

The experiments of Section 10 are reproduced by running the script `powers_of_two_c.py` with the options `-xmax 20000 -jmax2 131072 -pxmax 3000 -primes 3,5,7,11 -kmax 10000`, using Python 3.12.3 and NumPy 2.4.6.

References

[1] J.-L. Nicolas, Some open questions, The Ramanujan Journal 9 (2005) 251–264. doi:10.1007/s11139-005-0836-2.
[2] N. J. A. Sloane, S. Plouffe, The Encyclopedia of Integer Sequences, Academic Press, San Diego, 1995.
[3] H.-K. Hwang, S. Janson, T.-H. Tsai, Identities and periodic oscillations of divide-and-conquer recurrences splitting at half, Advances in Applied Mathematics 155 (2024) 102653, full version: arXiv:2210.10968. doi:10.1016/j.aam.2023.102653.
[4] R. de la Bretèche, C. Pomerance, G. Tenenbaum, Products of ratios of consecutive integers, The Ramanujan Journal 9 (2005) 131–138. doi:10.1007/s11139-005-0831-7.