Sequences that satisfy $a(n-a(n)) = 0$

Nate Kube* Frank Ruskey†

October 13, 2005

Abstract

We explore the properties of some sequences for which $a(n - a(n)) = 0$. Under the natural restriction that $a(n) < n$ the number of such sequences is a Bell number. Adding other natural restrictions yields sequences counted by the Catalan numbers, the Narayana numbers, the triangle of triangular binomial coefficients, and the Schröder numbers.

1 Introduction, set partitions

We consider here sequences $a(1), a(2), \ldots, a(n)$ with the property that $a(j - a(j)) = 0$ for all $j = 1, 2, \ldots, n$. Naturally, we must have $1 \le j - a(j) \le n$ for $j = 1, 2, \ldots, n$. Let $\mathbf{F}(n)$ be the set of all such sequences, and let $\mathbf{F}(n, m)$ be the subset of those for which m of the $a(j)$ are zero.[1]

Theorem 1.1. *For all* $1 \le m \le n$,

$$|\mathbf{F}(n, m)| = \binom{n}{m} m^{n-m}.$$

Proof. There are $\binom{n}{m}$ ways to choose the m indices $J = \{j_1, j_2, \ldots, j_m\}$ for which $a(j) = 0$. Each of the $n - m$ other elements t can take on any value from the m-set $\{t - j : j \in J\}$. For such a value of t, we have $a(t - a(t)) = a(t - (t - j)) = a(j) = 0$. □

*Dept. of Computer Science, University of Victoria, research supported in part by a NSERC PGS-M. e-mail: nkube@cs.uvic.ca

†Dept. of Computer Science, University of Victoria, research supported in part by NSERC. WWW: www.cs.uvic.ca/~ruskey

[1] Throughout the paper we will use a bold-case letter, like $\mathbf{X}$ to denote a set of sequences, $\mathbf{X}(n)$ to denote the sequences in $\mathbf{X}$ of length n, and $\mathbf{X}(n, m)$ to denote the sequences in $\mathbf{X}(n)$ that contain exactly m zeroes.

The numbers occurring in Theorem 1.1 are not in the OEIS [6] but summing on m yields A000248, the number of labelled forests in which every tree is a star (isomorphic to K_{1s} for some s). To get a correspondence with our sequences, let the parent of node j be node $j - a(j)$, with roots regarded as self-parents.

Comtet calls these numbers the "idempotent numbers" [1] (see pg. 91,135). The number of idempotent functions on an n-set that have m fixed-points is $\binom{n}{m} m^{n-m}$.

If we take the subset of $\mathbf{F}(n)$ that is closed under the taking of prefixes (or, equivalently, that the $a(j)$ only take on non-negative values), then we get the additional constraint that $a(j) < j$ for $j = 1, 2, \ldots, n$. (Of course, the set of sequences that only satisfy the condition $a(j) < j$ has cardinality $n!$, but our condition is stronger.) Let $\mathbf{A}(n)$ denote the set of all such sequences and let $\mathbf{A}(n, m)$ denote the subset of $\mathbf{A}(n)$ consisting of sequences with exactly m zeroes. Below we list the elements of $\mathbf{A}(n)$ for $n = 1, 2, 3, 4$.

$$\begin{aligned}
\mathbf{A}(1) &= \{0\} \\
\mathbf{A}(2) &= \{00, 01\} \\
\mathbf{A}(3) &= \{000, 001, 002, 010, 012\} \\
\mathbf{A}(4) &= \{0000, 0001, 0002, 0003, 0010, 0012, 0013, 0020, \\
&\quad 0022, 0023, 0100, 0101, 0103, 0120, 0123\}
\end{aligned}$$

Note that 011 is missing from $\mathbf{A}(3)$ since then $a(3 - a(3)) = a(3 - 1) = a(2) = 1 \neq 0$. Using the notation of Comtet [1] and Knuth [3], we denote the n-th Bell number, A000110, by ϖ_n and Stirling numbers of the second kind, A008277, by $\left\{ {n \atop m} \right\}$.

Theorem 1.2. *For all* $0 < m \leq n$*,*

$$|\mathbf{A}(n)| = \varpi_n \quad \textit{and} \quad |\mathbf{A}(n, m)| = \left\{ {n \atop m} \right\}.$$

Proof. Let $j_1, j_2, \cdots, j_m$ be the positions for which $a(j) = 0$. Now define the i-th block of a partition to be the set

$$B_i = \{k : k - a(k) = j_i\}.$$

Note that j_i is the smallest element of B_i. It should be clear that this specifies a one-to-one correspondence. □

Example: The sequence that corresponds to the partition $\{1, 3, 4\}, \{2, 5, 8\}, \{6, 7\}$ is $(0, 0, 2, 3, 3, 0, 1, 6)$.

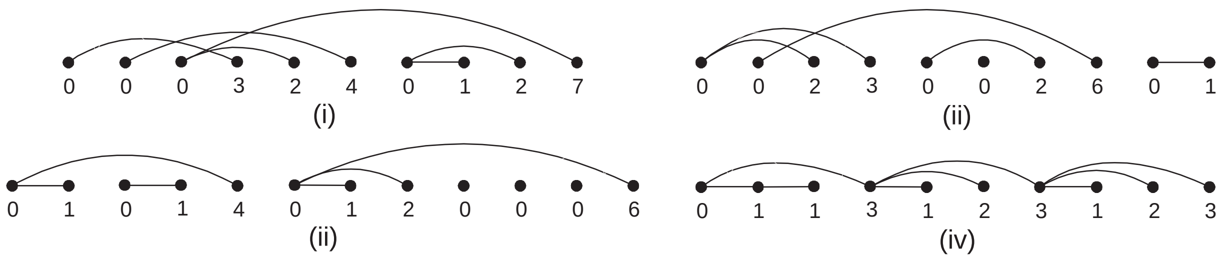

Figure 1: Linear difference diagram representation of an element from: (i) $\mathbf{A}(n)$, (ii) $\mathbf{B}(n)$, (iii) $\mathbf{C}(n)$, and (iv) $\mathbf{D}(n)$.

There is a natural pictorial representation of the sequences in $\mathbf{A}(n)$ as what we call a *linear difference diagram*, as shown in Figure 1(i) for the sequence $(0, 0, 0, 3, 2, 4, 0, 1, 2, 7)$. For each value $x \in \{1, 2, \ldots, n\}$, we draw an arc from x to $x - a(x)$, except if $a(x) = 0$. Our condition $a(n - a(n)) = 0$ then translates into the property that each connected component of the underlying graph is a star.

Define the set $\mathbf{B}(n)$ to be the subset of $\mathbf{A}(n)$ that satisfy the constraint that if $a(j) \neq 0$, then $a(j - 1) < a(j)$. We have that $\mathbf{B}(n) = \mathbf{A}(n)$ for $n = 1, 2, 3$ and $\mathbf{B}(4) = \mathbf{A}(4) \setminus \{0022\}$. Let $\mathbf{B}(n, m)$ denote the subset of $\mathbf{B}(n)$ consisting of sequences with exactly m zeroes. The numbers $|\mathbf{B}(n, m)|$ appear in OEIS [6] as sequence A098568 but no combinatorial interpretation is assigned to them. Summing on m gives the sequence A098569.

In terms of set partitions, the set $\mathbf{B}$ corresponds to those in which every element j such that j is not smallest in its block is in a block whose smallest element is no greater than the smallest element of the block containing $j - 1$.

Example The sequence $(0, 0, 2, 3, 0, 0, 2, 6, 0, 1)$, depicted in Figure 1(ii), is in $\mathbf{B}(n)$.

Theorem 1.3. *For all* $1 \leq m \leq n$,

$$|\mathbf{B}(n, m)| = \binom{n - 1 + \binom{m}{2}}{n - m}. \tag{1}$$

Proof. Denote $B(n, m) = |\mathbf{B}(n, m)|$. Classify the sequences in $\mathbf{B}(n, m)$ according to the index k of the rightmost zero. The sequences that occur in the first $k - 1$ positions are exactly those in $\mathbf{B}(k - 1, m - 1)$. The values that can go into positions $k + 1$ to n must be increasing and can be thought of as a selection with repetition of size $n - k - 1$ from the set of positions of the 0's, call them $1 = j_1 < j_2 < \cdots < j_m = k$. Arrange the selection as a nonincreasing sequence $l_{k+1} \geq l_{k+2} \geq \cdots \geq l_n$. Now, if $l_s = j_t$, then set $a(s) = s - j_t$. Note that $a(s - a(s)) = a(j_t) = 0$. Furthermore, $a(s) < a(s + 1)$ since $a(s) = s - j_t < s + 1 - j_{t'} = a(s + 1)$ where $t' \geq t$. This classification implies that the following recurrence relation holds, with the initial condition that $B(n,1) = 1$.

$$B(n,m) = \sum_{k=m}^{n} B(k-1, m-1)\binom{n-k+m-1}{n-k}$$

We will now show that the expression in (1) satisfies this recurrence relation. The following identity is well-known (see Gould [2], equation (3.2)).

$$\binom{x+y+t+1}{t} = \sum_{j=0}^{t} \binom{j+x}{j}\binom{t-j+y}{t-j} \tag{2}$$

We wish to show that

$$\binom{n-m-1+\binom{m+1}{2}}{n-m} = \sum_{k=m}^{n} \binom{k-m-1+\binom{m}{2}}{k-m}\binom{n-k+m-1}{n-k}.$$

But this is the same as (2) with $j = k-m$, $t = n-m$, $x = \binom{m}{2} - 1$, and $y = m-1$. □

2 Catalan and Schröder correspondences

Note that the linear difference diagram of Figures 1 (i) and 1 (ii) have crossing arcs. How many such sequences have no crossing arcs?

Define the set $\mathbf{C}(n)$ to be the set of sequences $a(1), a(2), \ldots, a(n)$ for which (a) $0 \leq a(j) < j$ and (b) there is no subsequence such that $i - a(i) < j - a(j) \leq i < j$. Note that $\mathbf{C}(n) = \mathbf{B}(n)$ for $n = 1, 2, 3, 4$ and $\mathbf{C}(5) = \mathbf{B}(5) \setminus \{00203\}$, since $3 - a(3) < 5 - a(5) \leq 3 < 5$. Let $\mathbf{C}(n,m)$ denote the subset of $\mathbf{C}(n)$ consisting of sequences with exactly m zeroes. The numbers $|\mathbf{C}(n,m)|$ appear in OEIS [6] as sequence A001263, the Catalan triangle. Summing on m gives the Catalan numbers A000108.

Lemma 2.1. *For all $n \geq 1$, $\mathbf{C}(n) \subseteq \mathbf{A}(n)$.*

Proof. Suppose that there is some value j for which $a(j - a(j)) > 0$, and let $i = j - a(j)$. We will show that

$$i - a(i) < j - a(j) = i < j,$$

which will prove the lemma. First note that $a(j) > 0$ since otherwise $a(j - a(j)) = a(j) = 0$. Thus $i < j$. Finally, $i - a(i) = j - a(j) - a(j - a(j)) < j - a(j)$ by our assumption that $a(j - a(j)) > 0$. □

Lemma 2.2. *For all $n \geq 1$, $\mathbf{C}(n) \subseteq \mathbf{B}(n)$*

Proof. If there is some sequence $a(1), a(2), \ldots, a(n)$ that is not in $\mathbf{B}(n)$, then there is some value of j such that $a(j-1) \geq a(j)$ and $a(j) > 0$. Setting $i = j - 1$ we would then have $i - a(i) < j - a(j) \leq i < j$ and so the sequence is not in $\mathbf{C}(n)$ either. □

A Dyck path on $2n$ steps is a lattice path in the coordinate plane (x, y) from $(0, 0)$ to $(2n, 0)$ with steps $(1, 1)$ (*Up*) and $(1, 1)$ (*Down*), never falling below the x-axis. Figure 2 shows a typical Dyck path of length 24.

The numbers shown below for $|\mathbf{C}(n, m)|$ are called the Narayana numbers [4]. They count the number of Dyck paths on $2n$ steps with m peaks.

Let C_n denote the n-th Catalan number, $C_n = \frac{1}{n+1}\binom{2n}{n}$. The correspondence used in the proof below is mentioned in [5], problem 6.19(f^4).

Theorem 2.3.

$$|\mathbf{C}(n)| = C_n \textit{ and } |\mathbf{C}(n, m)| = \frac{1}{m}\binom{n-1}{m-1}\binom{n}{m-1}.$$

Proof. Considering *Up* steps as left parentheses and *Down* steps as right parentheses, a Dyck path of length $2n$ corresponds to a well-formed parentheses string of equal length. Furthermore, a peak corresponds to a `()` pair.

Let $\mathbf{S}_{2n}$ denote the set of well-formed parenthesis strings of length $2n$. Define the function f from $\mathbf{S}_{2n}$ to $\mathbf{C}(n)$ as $f(s) = (a(1), a(2), \ldots, a(n))$ where $a(j)$ is the number of right parentheses that are properly enclosed by the j-th parentheses pair where the parentheses pairs are numbered by the order in which their right parentheses occur. The left parenthesis that matches the j-th right parenthesis is referred to as *j's match*. For example, consider $s =$ `((())(()))(((()))()()())`. Subscripting the right parentheses and overlining matching parentheses pairs we obtain

$$((()_1)_2(()_3)_4)_5(((()_6)_7)_8()_9()_{10}()_{11})_{12}$$

and thus $f(s) = (0, 1, 0, 1, 4, 0, 1, 2, 0, 0, 0, 6)$ (represented as a linear difference diagram in Figure 1(iii)).

We now need to explain why $f(\mathbf{S}_{2n}) \subseteq \mathbf{C}(n)$. Let $s \in \mathbf{S}_{2n}$ and consider $f(s) = (a(1), a(2), \ldots, a(n))$. By definition $a(i) < i$. Furthermore the sequence satisfies the non-crossing property. If it did not, for some $i < j$ we would have that both $j - a(j) \leq i$ and $i - a(i) < j - a(j)$. Notice that $j - a(j)$ is the position of the leftmost right parenthesis to the right of j's match. Now, $j - a(j) \leq i < j$ implies that the i-th right parenthesis must lie between j and its match. As well, $i - a(i) < j - a(j)$ implies that the leftmost right parenthesis to the right of i's match is left of the leftmost right parenthesis to the right of j's match. This implies that i's match is left of j's match which contradicts the fact that s is well-formed. Hence $f(s) \in \mathbf{C}(n)$ thus $f(\mathbf{S}_{2n}) \subseteq \mathbf{C}(n)$.

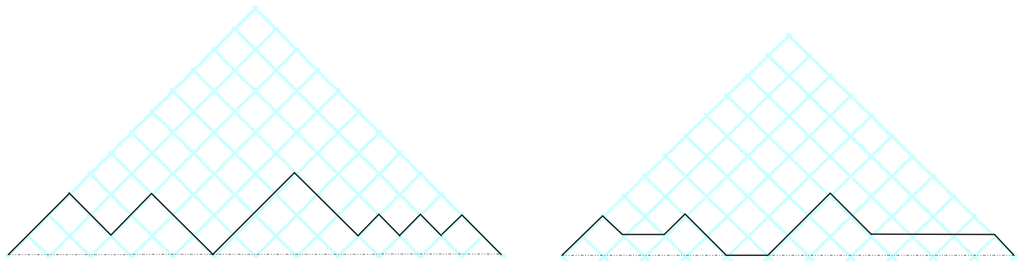

Figure 2: The sequence $(0, 1, 0, 1, 4, 0, 1, 2, 0, 0, 0, 6)$ represented as: (i) a Dyck path of length 24, (ii) a Schröder path of length 11.

We now show that f is indeed a bijection. Let $b = (b(1), b(2), \ldots, b(n)) \in \mathbf{C}(n)$. We show by induction that it is possible to construct exactly one $s \in \mathbf{S}_{2n}$ such that $f(s) = (a(1), a(2), \ldots, a(n)) = b$. Let $k = 1$ and s be the well-formed parenthesis string `()`. Then $f(s) = (a(1)) = (0) = b(1)$ and s is the only such string. Assume $f(s) = (a(1), a(2), \ldots, a(n)) = (b(1), b(2), \ldots, b(n))$ for some $n \geq 1$ where s is the only such string. Consider $b(n+1)$. If $b(n+1) = 0$ then appending `()` to s, in which there is only one way, results in $f(s) = (a(1), a(2), \ldots, a(n), a(n+1)) = (b(1), b(2), \ldots, b(n), b(n+1))$ and s is the only such string. Similarly, if $b(n+1) = n$, then enclosing s within a right and a left parenthesis produces the desired result.

Suppose that $0 < b(n+1) < n$. Consider the substring s' consisting of all elements of s to the right of the $\{n+1-b(n+1)-1\}$-th right parenthesis. If s' is not well-formed then there exists a right parenthesis i with $n+1-b(n+1) \leq i < n+1$ whose match is to the left of the $\{n+1-b(n+1)-1\}$-th right parenthesis. This implies that $i - b(i) \leq n+1-b(n+1)-1$. However then, $i - b(i) < n+1-b(n+1) \leq i < n+1$ which contradicts the fact that b is in $\mathbf{C}(n)$. Therefore s' is well-formed and simply enclosing it in a left and right parentheses pair within s produces the desired result.

Furthermore, note that a zero in an element of $\mathbf{C}(n)$ corresponds to a `()` in an element of $\mathbf{S}_{2n}$ which corresponds to a peak in a Dyck path of length $2n$. Thus $|\mathbf{C}(n, m)|$ is the number of Dyck paths of length $2n$ with m peaks. $\square$

The reader may wonder what happens if we were allowed to have $i - a(i) < j - a(j) = i < j$ but not $i - a(i) < j - a(j) < i < j$. Call the resulting set $\mathbf{D}(n)$. Such sequences need no longer satisfy $a(j - a(j)) = 0$ so strictly speaking are outside the scope of this paper, but the question is interesting nonetheless. They are counted by Schröder numbers.

Let D_n denote, A006318, the n-th (large) Schröder number, $D_n = \langle z^n \rangle \frac{1-z-\sqrt{1-6z+z^2}}{2z}$ (see [4], pg. 178).

Example The sequence $(0, 1, 1, 3, 1, 2, 3, 1, 2, 3)$, depicted as a linear difference diagram in Figure 1(iv), is in $\mathbf{D}(n)$ but not $\mathbf{C}(n)$.

A Schröder path is a lattice path in the coordinate plane (x, y) from $(0, 0)$ to $(n, 0)$ with steps $(1, 1)$ (*Up*), $(1, -1)$ (*Down*) and $(1, 0)$ (*Straight*) never falling below the x-axis. The length of a Schröder path is the number of *Up* and *Straight* steps in the path. Figure 2shows a typical Schröder path of length 11.

The numbers $\binom{2n-m-1}{m-1}C_{n-m}$ count the number of Schröder paths from $(0, 0)$ to $(n-1, n-1)$ containing $m-1$ *Straight* steps (A060693).

Theorem 2.4.

$$|\mathbf{D}(n)| = D_{n-1} \text{ and } |\mathbf{D}(n, m)| = \binom{2n-m-1}{m-1}C_{n-m}$$

Proof. Let $\mathbf{P}_n$ denote the set of Schröder paths of length n. Define the function g from $\mathbf{P}_{n-1}$ to $\mathbf{D}(n)$ as $g(p) = (0, a(1), a(2), \ldots, a(n-1))$ where $a(j)$ is 0 if the j-th counted step is *Straight* or is the number of counted steps (starting with itself) between it and its corresponding *Down* step. For example, let p be the Schröder path shown in Figure 2. Then $g(p) = (0, a(1), a(2), \ldots, a(11)) = (0, 1, 0, 1, 4, 0, 1, 2, 0, 0, 0, 6)$. Notice that $a(i) = 0$ exactly when the i-th counted step in p is *Straight*.

We now need to explain why $g(\mathbf{P}_{n-1}) \subseteq \mathbf{D}(n)$. Let $p \in \mathbf{P}_{n-1}$. Consider $g(p) = (0, a(1), a(2), \ldots, a(n-1)) = (b(1), b(2), \ldots, b(n))$. Since p is a Schröder path $a(i) \leq i$. Since $b(i+i) = a(i)$ we have that $b(i+1) < i+1$. Now, suppose that there exists an $1 < i < j \leq n$ such that $i - b(i) < j - b(j) < i < j$. Then there exists some $1 \leq x < y < n$ such that $x - a(x) < y - a(y) < x < y$. Since both $a(y)$ and $a(x)$ must be non-zero to satisfy this inequality, we have that they both count the number of countable steps (beginning with themselves) between them and their respective matches. Now, x lies between y and its match. Furthermore, $y - a(y)$ is the position of the first countable step to the left of y's match. Since $x - a(x) < y - a(y)$, the first countable step to the left of x's match is to the left of the first countable step to the left of y's match which implies that x's match is to the left of y's match. This means that between y and its match there is one more *Up* step then *Down* step thus y and its match are not on the same level contradicting the fact that this is indeed y's match. Hence $g(p) \in \mathbf{D}(n)$ thus $g(\mathbf{P}_{n-1}) \subseteq \mathbf{D}(n)$.

We now show that g is a bijection. Let $b = (b(1), b(2), \ldots, b(n)) \in \mathbf{D}(n)$. We show by induction that it is possible to construct exactly one $p \in \mathbf{P}_{n-1}$ such that $g(p) = (0, a(1), a(2), \ldots, a(n-1)) = b$. Let $k = 2$. If $b(2) = a(1) = 0$ let p' be the Schröder path of length 1 consisting of 1 *Straight* step. Then $g(p') = (0, 0) = (b(1), b(2))$ and there was only one such p'. Otherwise let p' be the Schröder path of length 1 consisting of one *Up* step and its match. Then $g(p') = (0, 1) = (b(1), b(2))$ and there was only one such p'.

Assume $g(p') = (0, a(1), a(2), \ldots, a(j-1)) = (b(1), b(2), \ldots, b(j))$ for some $j \geq 1$ and p' is the only such path. Consider $b(j+1)$. If $b(j+1) = 0$ then appending a *Straight* step to p', in which there is only one way, results in $g(p') = (0, a(1), a(2), \ldots, a(j)) = (b(1), b(2), \ldots, b(j), b(j+1))$ and p' is the only such string. Similarly, if $b(j+1) = j$ appending an *Up* step to the end p' and placing its match at the front produces the desired result.

Suppose that $0 < b(j+1) < j$. Consider the path p'' consisting of all elements of p' to the right of the $j-a(j)$-th countable step. If p'' is not a Schröder path then there exists some *Up* step at position $j-a(j) < i < j$ whose match is to the left of the $j-a(j)$-th countable step. However, this implies that the first countable step to the left of i's match is to the left of the first countable step to the left of j's match. This implies that $i-a(i) < j-a(j) < i < j$ and hence $i+1-b(i+1) < j+1-b(j+1) < i+1 < j+1$ contradicting the fact that b is in $\mathbf{D}(n)$. Therefore p'' is a Schröder path. Now, within p', simply appending an *Up* step to the end of p'' and placing its match at the front of p'' produces the desired result.

Furthermore, since a zero in an element of $\mathbf{D}(n)$ in any position other than the first corresponds to a *Straight* step in a Schröder path of length $n-1$, $|\mathbf{D}(n,m)|$ = the number of Schröder paths of length $n-1$ with $m-1$ zeros. □

References

[1] Louis Comtet, *Advanced Combinatorics*, D. Reidel Publishing Company, 1974.

[2] H.W. Gould, *Combinatorial Identities*, Morgantown Printing and Binding, revised edition, 1972.

[3] Donald E. Knuth, *The Art of Computer Programming*, pre-fascicle 3B (a draft of sections 7.2.1.4-5: Generating all Partitions), Addison-Wesley.

[4] Richard P. Stanley, *Enumerative Combinatorics, Volume 2*, Cambridge University Press, 1999.

[5] Richard P. Stanley, *Catalan Addendum*, version of 17 February 2005, available online at `http://www-math.mit.edu/~rstan/ec/catadd.pdf`.

[6] Neil J. Sloane, *The Online Encyclopedia of Integer Sequences (OEIS)*, `http://www.research.att.com/~njas/sequences/`, March 2005.

2000 Mathematics Subject Classification: Primary 05A15, Secondary 68R15.

Keywords: Bell number, Catalan number, Schröder number, bijection.

Concerned with sequences: A000248, A000110, A008277, A098568, A098569, A001263, A000108, A006318, A060693.