STOPPING TIMES

ROBERT DOUGHERTY-BLISS AND CHARLES KENNEY

1. Stopped binary strings

Definition 1. A *binary string of length n* is a tuple $(x_1, x_2, \ldots, x_n)$ where each x_k is 0 or 1. Such a string is said to be *stopped* at index k if every index of the tuple in $(k/2, k]$ is zero. The *stopping time* of a binary string is the smallest k such that the string is stopped at index k, or ∞ if no such k exists. Note that, except 1 and ∞, every stopping time is even.

We are primarily interested in binary strings of length n which have stopping time n. Indeed, let $g(n)$ be the number of such strings. Our first goal is to show that $g(n)$ satifies the recurrence

$$\begin{aligned} g(1) = g(2) &= g(4) = 1 \\ g(4n) &= 2g(4n-2) \\ g(4n+2) &= 2g(4n) - g(2n). \end{aligned}$$

This will identify the sequence $a(n) = g(2n)$ as the *Narayana–Zidek–Capell numbers*, A2083 in the OEIS, which begin

$$1, 1, 1, 2, 3, 6, 11, 22, 42, 84, 165, 330, 654, \ldots$$

This recurrence was discovered thanks to the OEIS.

Our first definition discusses *finite* binary strings, but it is more convenient to work with *infinite* binary strings which are nonzero in only finitely many places. This is equivalent, and has the benefit that every such infinite string has a finite stopping time when considered as a sufficiently long, finite binary string.

Definition 2. Let

$$V = \bigoplus_{k=1}^{\infty} (\mathbb{Z}/2\mathbb{Z})$$

be the direct sum of infinitely many copies of $\mathbb{Z}/2\mathbb{Z}$. That is, the set of all infinite tuples $(x_1, x_2, x_3, \ldots)$ where each x_k is 0 or 1, and only finitely many are 1. For each positive integer k, let $\bigcirc_k$ be the set of elements of V which are zero beyond position k and, when the first k

entries are regarded as a finite binary string, they have stopping time k. That is, $\bigcirc_1 = \{0\} \subset V$,

$$\bigcirc_{2k} = \{v \in V : \forall j > k, v_j = 0, \text{ and } v \text{ has stopping time } 2k\}$$

and $\bigcirc_{2k+1} = \emptyset$ for every integer $k \geq 1$.

It is clear that $g(n) = |\bigcirc_n|$ for every positive integer n.

Theorem 1. *The sequence $g(n)$ satisfies the recurrence*

$$\begin{aligned} g(1) = g(2) = g(4) &= 1 \\ g(4n) &= 2g(4n-2) \\ g(4n+2) &= 2g(4n) - g(2n). \end{aligned}$$

Proof. Each nonzero element in V has a final nonzero entry. Let $e_0 \colon V \to V$ be a map which inserts a 0 in the position of this final entry, shifting the previous entry to the right. Similarly, let e_1 insert a 1 into this position. Symbolically,

$$e_0(b) = \begin{cases} 0 \text{ if } b = 0 \\ (b_1, ..., b_j, 0, 1, 0, 0, ...) \text{ if } b = (b_1, ..., b_j, 1, 0, 0, ...) \end{cases}$$

and

$$e_1(b) = \begin{cases} (1, 0, 0, \dots) \text{ if } b = 0 \\ (b_1, ..., b_j, 1, 1, 0, 0, ...) \text{ if } b = (b_1, ..., b_j, 1, 0, 0, ...) \end{cases}.$$

Note that e_0 and e_1 are injective, $e_0(V) \cap e_1(V) = \emptyset$, and $e_0(V) \cup e_1(V) = V$.

For any positive integer k,

$$\bigcirc_{2k+2} \subseteq e_0(\bigcirc_{2k}) \cup e_1(\bigcirc_{2k}).$$

Indeed, the final nonzero entry of every string in $\bigcirc_{2k+2}$ occurs at position $2k+1$. Removing the immediately preceding entry produces a string in $\bigcirc_{2k}$—e_0 and e_1 simply add the entry back.

In the other direction, for $k \geq 2$ we have

$$e_1(\bigcirc_{2k}) \subseteq \bigcirc_{2k+2},$$

because inserting a 1 into position k will never decrease the stopping time. For *odd* $k \geq 2$ we have

$$e_0(\bigcirc_{2k}) \subseteq \bigcirc_{2k+2},$$

because inserting a 0 at position k will only decrease the stopping time if the new stopping time *is* k; this is impossible if k is odd. This shows that

$$S_{2k+2} = e_0(S_{2k}) \cup e_1(S_{2k}),$$

and thus

$$g(2k+2) = 2g(2k)$$

for k odd.

To handle the even case, let

$$\bigcirc_{2k}^{1} = \{(b_1, ..., b_{2k}, 1, 0, 0, ...) \in V \text{ s.t. } (b_1, ..., b_{2k}, 0, 0, ...) \in \bigcirc_{2k}\}.$$

(Note that $|\bigcirc_{2k}^{1}| = |\bigcirc_{2k}| = g(2k)$.) Then

$$e_0(\bigcirc_{4k}) \subseteq \bigcirc_{4k+2} \cup \bigcirc_{2k}^{1}.$$

This implies

$$\bigcirc_{4k+2} \cup \bigcirc_{2k}^{1} = e_0(\bigcirc_{4k}) \cup e_1(\bigcirc_{4k}).$$

(That the left contains the right is clear given the preceding inclusion. For the other direction, the only problem is $\bigcirc_{2k}^{1} \subseteq e_0(\bigcirc_{4k}) \cup e_1(\bigcirc_{4k})$. This is easily handled by considering the element immediately preceding the final nonzero element.) Therefore

$$g(4k+2) + |\bigcirc_{2k}^{1}| = 2g(4k),$$

or

$$g(4k+2) = 2g(4k) - g(2k).$$

This establishes the recurrence. □

2. Stopped reals in the unit interval

Every element of the unit interval $[0, 1]$ can be regarded, via its binary expansion, as an infinite binary string, i.e., an element of $\mathbb{Z}/2Z \times \mathbb{Z}/2Z \times \cdots$. This differs from the set V of the previous section by allowing potentially infinitely many nonzero entries. For example, the real

$$(0.1010101010101\ldots)_2 = \frac{2}{3}$$

is perfectly well-defined, but not as a member of $\bigoplus_{k=1}^{\infty}(\mathbb{Z}/2\mathbb{Z})$. In this way, it makes since to discuss stopped *reals*, considering each real as an infinite binary string. (We shall interchangably refer to reals as both strings and numbers.) Our focus now becomes *topological* as we examine the *set* of stopped reals in the unit interval.

Definition 3. Let S_k be the set of reals in $[0, 1]$ which have stopping time k when regarded as an infinite binary string. Membership in this set is determined by examining only the first k bits of a binary expansion, so S_k is measureable. Let $S = \cup_{k \geq 1} S_k$ be the set of all stopped reals in $[0, 1]$, which is measurable since the S_k are.

Let us get a feel for what the sets S_k "look like." First, observe that every element of S_k consists of a binary string which has stopping time k followed by *any* binary string at all. Thus, each string with stopping time k determines an interval of length 2^{-k} included in S_k, and the disjoint union of these intervals is *all* of S_k. In this way, we can compute the following sets:

$$\begin{aligned}
S_1 &= [0, 1/2) \\
S_2 &= [1/2, 3/4) \\
S_3 &= \emptyset \\
S_4 &= [0, 1/2) \\
S_5 &= \emptyset \\
S_6 &= [13/16, 14/16) \cup [15/16, 1)
\end{aligned}$$

[Insert the stopping time plot somewhere around here.]

This simple description of S_k gives us a nice expression for the measure of S.

Theorem 2. *Let β_2 be the measure of S. Then*

$$\beta_2 = \sum_{k \geq 1} \frac{g(k)}{2^k} = \frac{1}{2} + \sum_{k \geq 1} \frac{g(2k)}{4^k} \approx 0.841657913173647.$$

Proof. The measure of S_k is $g(k)/2^k$, where $g(k)$ is the number of binary strings of length k with stopping time k. Since S is the pairwise disjoint union of the S_k, the "exact" result follows immediately. □

In the previous section we established that the sequence $a(n) = g(2n)$ is A2083 in the OEIS. It is known that $a(n) = O(2^n)$, so the above series converges very quickly. Using the recurrence of the previous section to generate the first hundred terms of $g(2k)$ it is easy to approximate the sum and provide rigorous lower bounds. To get more rigorous upper bounds, first note $g(2n)/2^n$ is monotonically decreasing. This implies that the error of using

$$\frac{1}{2} + \sum_{1 \leq k < n} \frac{g(2k)}{4^k}$$

as an approximation to the measure of S is

$$\sum_{k \geq n} \frac{g(2k)}{4^k} \leq \frac{g(2n)}{2^n} \sum_{k \geq n} \frac{1}{2^k} = \frac{2g(2n)}{4^n}.$$

For example, using the first 99 terms ($n = 100$) will give an error of no more than

$$\frac{22284668265087}{1584563250285286751870879006720} \approx 0.1406360286 \times 10^{-15}.$$

The constant β_2 seems new. We conjecture that it is irrational, and even transcendental. In fact, the only thing we know about it is that it is a value of the ordinary generating function of $g(n)$.

Definition 4. Let

$$G(z) = \sum_{k\geq 0} g(k)z^k = z + \sum_{k\geq 1} g(2k)z^{2k}$$

be the ordinary generating function of $g(n)$, and

$$A(z) = \sum_{k\geq 1} g(2k)z^k$$

the ordinary generating function of $g(2k)$. Note that $G(z) = z + A(z^2)$.

Lemma 1. *The generating function $A(z)$ is analytic on a disk of radius $1/2$ centered at the origin and satisfies*

$$A(z) = \frac{z(1 - A(z^2))}{1 - 2z}.$$

Proof. The equation is a routine computation using the recurrence for $g(2n)$ proved in the previous section. For convenience, let $a(n) = g(2n)$. Then:

$$\begin{aligned}
\sum_{k\geq 1} a(k)z^k &= \sum_{k\geq 1} a(2k)z^{2k} + \sum_{k\geq 0} a(2k+1)z^{2k+1} \\
&= \sum_{k\geq 1} 2a(2k-2)z^{2k} + a(1)z + \sum_{k\geq 1}(2a(2k) - a(k))z^{2k+1} \\
&= 2z\sum_{k\geq 0} a(2k+1)z^{2k+1} + 2z\sum_{k\geq 1} a(2k)z^{2k} - zA(z^2) + z \\
&= 2zA(z) + z(1 - A(z^2)).
\end{aligned}$$

Solving this for $A(z)$ yields the result. It is well-known that $a(n) = O(2^n)$, so $A(z)$ converges everywhere that $\sum_{k\geq 0}(2z)^k = (1-2z)^{-1}$ does, which is at least a disk of radius $1/2$. $\square$

It is clear that

$$\beta_2 = G(1/2) = \frac{1}{2} + A(1/4).$$

But again, this is very little information. We suspect that neither $A(z)$ nor $G(z)$ are algebraic.

Our choice of notation for β_2—including a subscript 2—is quite intentional. Everything that we have done with binary sequences generalizes almost directly to arbitrary bases. Take the definitions we gave before and replace any occurrence of "1" with "nonzero entry." Almost everything that we have said carries over, except that the recurrence for $g(n)$ is slightly different. We will now give these definitions in order to define an infinite family $\{\beta_n\}$ of constants.

Definition 5. For any integer $b \geq 2$, say that a finite "b-ary" tuple $(x_1, \dots, x_n)$ with $x_k \in \{0, 1, \dots, b-1\}$ is *stopped* at an index k provided that every index in $(k/2, k]$ is zero. The *stopping time* of such a string is the smallest k such that the string is stopped at k, or ∞ if no such k exists. The stopping time of an infinite tuple is the stopping time of its shortest stopped prefix, or ∞ if no such prefix exists.

Definition 6. For any integer $b \geq 2$, let $g_b(n)$ be the number of b-ary strings of length n with stopping time n.

The same argument used in Theorem 1 (defining, now, e_0, e_1, and so on up to e_{b-1}) will establish the following result.

Theorem 3.