# 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 $\infty$ if no such $k$ exists. Note that, except 1 and $\infty$, every stopping time is even.

We should first note that this definition and our later results generalize immediately to bases other than 2. The only change in the definition is that the elements of the strings may be anything in $\{0, 1, \ldots, b-1\}$ in base $b$. In the later arguments, "1" should be replaced with "any nonzero number." For simplicity we will sometimes discuss only binary, but will quote the arbitrary base results when they are needed.

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 $(x_1, x_2, x_3, \dots)$ 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. (1)* Each nonzero element in $V$ has a final nonzero entry (a one). Let $e_0 \colon V \to V$ be a map which inserts a 0 in the position of this final entry, shifting the previous final entry to the right one space. Similarly, let $e_1$ insert a 1. 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 $k+1$. 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* $k \geq 3$ we have

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

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

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

and thus

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

for $k$ odd.

To handle the even case, let

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

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

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

This implies

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

(That the left contains the right is clear given the preceding inclusion. For the other direction, note $\octagon^1_{2k} \subseteq e_0(\octagon_{4k})$, since if $(b_1, ..., b_{2k}, 0, 0, ...) \in \octagon_{2k}$ then $(b_1, ..., b_{2k-1}, 1, 0, 0, ...)$ has stopping time precisely $4k$.) Therefore

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

or

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

This establishes the recurrence. □

*Proof. (2)*

**Definition 3.** A *composition* of $n$ is a vector in $(\mathbb{Z}_{>0})^m$, for some $m \in \{1, 2, ..., n\}$, whose components add to $n$.

The $n$th Narayana-Zidek-Capell number, $a(n)$, is the cardinality of the set

$$R_n = \left\{ \text{Compositions } p \text{ of } n \text{ satisfying } p_1 = 1 \text{ and } p_k \leq \sum_{j=1}^{k-1} p_j \right\}.$$

For $j \in \mathbb{Z}_{>0}$, let $f(j) = (0, 0, ..., 0, 1)$, with $j-1$ zeroes preceding the one. Then for $p \in (\mathbb{Z}_{>0})^m$, let

$$F(p) = (f(p_1), f(p_2), ..., f(p_m), 0, 0, ...),$$

concatenated in the natural way, with a tail of zeroes appended at the end. Then $F : R_n \xrightarrow{\sim} \bigcirc_{2n}$ is a bijection, so $|\bigcirc_{2n}| = g(2n) = a(n)$. Indeed, if $p \in R_n$, then since $\sum_{j=1}^m p_j = n$, the final nonzero entry in $F(p)$ is in spot $n$. Thus $F(p)$ is stopped at time $2n$, and is not stopped at time $t$ for any $t \in [n, 2n-1]$. And if the zeroes from $f(p_r)$ in $F(p)$ caused $F(p)$ to be stopped at some time $t \leq n-1$,

$$F(p) = (f(p_1), ..., f(p_{r-1}), 0, 0, ..., \underset{\substack{\uparrow \\ \text{index } t}}{0}, ...)$$

then $p_r > \lceil t/2 \rceil \geq \sum_{j=1}^{r-1} p_j$. But this would contradict the fact that $p \in R_n$. And, if $b \in \bigcirc_{2n}$ has 1s at indices $a_1 = 1, a_2, ..., a_\ell = n$, then

$$\begin{aligned} F^{-1} &: \bigcirc_{2n} \to R_n \\ F^{-1} &: b \mapsto (a_1, a_2 - a_1, a_3 - a_2, ..., a_\ell - a_{\ell-1}) \end{aligned}$$

provides an inverse to $F$, since the index $a_r$ of the $r$th one in b is at most twice the index $a_{r-1}$ of the $(r-1)$st one. $\square$

A slightly different recurrence holds for arbitrary bases. Let $g_b(n)$ be the number of $b$-ary strings of length $n$ with stopping time $n$. Then

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

## 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 4.** 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:

$$\begin{aligned} S_1 &= [0, 1/2) \\ S_2 &= [1/2, 3/4) \\ S_4 &= [3/4, 13/16) \\ S_6 &= [7/8, 57/64) \\ S_8 = [13/16, 209/256) &\cup [15/16, 241/256). \end{aligned}$$

[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 $\beta_2$ 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 $g(k)/2^{2k}$, 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. $\square$

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

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

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

$$\frac{22284668265087}{1584563250285286751870879006720} \approx 1.406360286 \times 10^{-16}.$$

The constant $\beta_2$ seems new. Accordingly, we conjecture that it is irrational, and even transcendental. But we really know nothing about $\beta_2$. The only information we have is that it is a value of the ordinary generating function of $g(n)$.

**Definition 5.** 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 $G(z) = z + A(z^2)$.

**Proposition 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:

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

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.

Expanding reals in $[0,1]$ with different bases yields different "stopped sets" and different measures. The same arguments apply, except now the measure of the $S_k$ will be $g_b(k)/b^k$. Thus, the measure of the set of $b$-ary stopped reals is

$$\beta_b = \frac{1}{b} + \sum_{k\geq 1} \frac{g_b(2k)}{b^{2k}}.$$

These constants go as follows:

$$\begin{aligned}
\beta_2 &= 0.84165791317364708989\\
\beta_3 &= 0.62119074589923243760\\
\beta_4 &= 0.48141057151328202149\\
\beta_5 &= 0.39071175514852239829\\
\beta_6 &= 0.32805820924751380527.
\end{aligned}$$

These seem to be decreasing, and indeed they are, down to 0.

**Theorem 3.** *For $b \geq 2$,*

$$\frac{1}{b} \leq \beta_b \leq \frac{2}{b}.$$

*Proof.* The definition of $g_b$ shows that it is positive, and the recurrence established in the previous section shows that $g_b(2k)/b^k$ is monotonically decreasing. In particular,

$$g_b(2k)/b^k \leq g_b(2)/b = \frac{b-1}{b}.$$

for $k \geq 1$. Therefore

$$\beta_b \leq \frac{1}{b} + \frac{b-1}{b} \sum_{k\geq 1} \frac{1}{b^k} = \frac{2}{b}.$$

The lower bound is from the definition of $\beta_b$. □

## 3. Stopped integers

**Definition 6.** Say that a positive integer $n$ is *maximally stopped* provided that its binary expansion has stopping time equal to its length.

The maximally stopped integers begin

$$2, 12, 56, 208, 240, 864, 928, 992, 3392, 3520, 3648, \dots,$$

and do not yet appear in the OEIS.

??? [I have no idea what to say about these. But look—an integer sequence!]