REGULARITY PROPERTIES OF THE STERN ENUMERATION OF THE RATIONALS

BRUCE REZNICK

Abstract. The Stern sequence $s(n)$ is defined by $s(0) = 0, s(1) = 1, s(2n) = s(n),$ $s(2n + 1) = s(n) + s(n + 1)$. Stern showed in 1858 that $gcd(s(n), s(n + 1)) = 1$, and that for every pair of relatively prime positive integers $(a, b)$ there exists a unique $n \geq 1$ with $s(n) = a$ and $s(n + 1) = b$. We show that in a strong sense, the average value of $\frac{s(n)}{s(n+1)}$ is $\frac{3}{2}$, and that for $d \geq 2$, $(s(n), s(n + 1))$ is uniformly distributed among all feasible pairs of congruence classes modulo $d$. More precise results are presented for $d = 2$ and $3$.

1. Introduction and History

In 1858, M. A. Stern [18] defined the diatomic array, an unjustly neglected mathematical construction. It is a Pascal triangle with memory: each row is created by inserting the sums of pairs of consecutive elements into the previous row.

\[
\begin{array}{cccccc}
   & a & b \\
 a & a+b & b \\
 a & 2a+b & a+b & a+2b & b \\
 a & 3a+b & 2a+b & 3a+2b & a+b & 2a+3b & a+2b & a+3b & b \\
 & & & & & & & & \\
\end{array}
\]

(1.1)

When $(a, b) = (0, 1)$, it is easy to see that each row of the diatomic array repeats as the first half of the next row down. The resulting infinite Stern sequence can also be defined recursively by:

\[
s(0) = 0, \quad s(1) = 1, \quad s(2n) = s(n), \quad s(2n + 1) = s(n) + s(n + 1).
\]

(1.2)

Taking $(a, b) = (1, 1)$ in (1.1), we obtain blocks of $(s(n))$ for $2^r \leq n \leq 2^{r+1}$. Although $s(2^r) = 1$ is repeated at the ends, each pair $(s(n), s(n + 1))$ appears below exactly

\begin{itemize}
  \item [Date:] October 18, 2006.
  \item [1991 Mathematics Subject Classification.] Primary: 05A15, 11B37, 11B57, 11B75.
\end{itemize}
once as a consecutive pair in a row:

\[
\begin{array}{ccc}
(r = 0) & 1 & 1 \\
(r = 1) & 1 & 2 & 1 \\
(r = 2) & 1 & 3 & 2 & 3 & 1 \\
(r = 3) & 1 & 4 & 3 & 5 & 2 & 5 & 3 & 4 & 1 \\
\end{array}
\]

(1.3)

Mirror symmetry (or an easy induction) implies that for \(0 \leq k \leq 2^r\), we have

\[
s(2^r + k) = s(2^{r+1} - k).
\]

(1.4)

In his original paper, Stern proved that for all \(n\),

\[
gcd(s(n), s(n + 1)) = 1;
\]

moreover, for every pair of positive relatively prime integers \((a, b)\), there is a unique \(n\) so that \(s(n) = a\) and \(s(n + 1) = b\). Stern's discovery predates Cantor's proof of the countability of \(\mathbb{Q}\) by fifteen years. This property of the Stern sequence has been recently made explicit and discussed in [4]. Another enumeration of the positive rationals involves the Stern-Brocot array, which also predates Cantor; see [8], pp. 116–123, 305–306. This was used by Minkowski in defining his \(z\)-function; see [14]. The Stern sequence and Stern-Brocot array make brief appearances in Dickson's History, see [6], pp. 156, 426. Apparently, de Rham [5] was the first to consider the sequence \((s(n))\) per se, attributing the term "Stern sequence" to Bachmann [2], p. 143, who had only considered the array. The Stern sequence has recently arisen as well in the discussion of 2-regular sequences [1] and the Tower of Hanoi graph [10]. Some other Stern identities and a large bibliography relating to the Stern sequence are given in [19]. A further discussion of the Stern sequence will be found in [16].

Let

\[
t(n) = \frac{s(n)}{s(n + 1)}.
\]

(1.6)

Here are blocks of \((t(n))\), for \(2^r \leq n < 2^{r+1}\) for small \(r\):

\[
\begin{array}{ccc}
(r = 0) & \frac{1}{1} \\
(r = 1) & \frac{1}{2} & \frac{2}{1} \\
(r = 2) & \frac{1}{3} & \frac{2}{3} & \frac{2}{1} \\
(r = 3) & \frac{1}{4} & \frac{1}{3} & \frac{3}{2} & \frac{3}{2} & \frac{3}{1} \\
\end{array}
\]

(1.7)
In Section 3, we shall show that

\[ \sum_{n=0}^{N-1} t(n) = \frac{3}{2} N + O(\log^2 N), \]

so the “average” element in the Stern enumeration of \( \mathbb{Q}_+ \) is \( \frac{3}{2} \).

For a fixed integer \( d \geq 2 \), let

\[ S_d(n) := (s(n) \mod d, s(n + 1) \mod d) \]

and let

\[ S_d = \{ (i \mod d, j \mod d) : \gcd(i, j, d) = 1 \}. \]

It follows from (1.5) that \( S_d(n) \in S_d \) for all \( n \). In Section 4, we shall show that for each \( d \), the sequence \( (S_d(n)) \) is uniformly distributed on \( S_d \), so the “probability” that \( s(n) \equiv i \mod d \) can be explicitly computed. More precisely, let

\[ T(N; d, i) = |\{ n : 0 \leq n < N & s(n) \equiv i \mod d \}|. \]

Then there exists \( \tau_d < 1 \) so that

\[ T(N; d, i) = r_{d,i} N + O(N^{\tau_d}), \]

where

\[ r_{d,i} = \frac{1}{d} \prod_{p | d \cap p \nmid d} \frac{p}{p+1} \prod_{p^k | d} \frac{p^2}{p^2 - 1}. \]

In particular, the probability that \( s(n) \) is a multiple of \( d \) is \( I(d)^{-1} \), where

\[ I(d) = d \prod_{p \mid d} \frac{p+1}{p} \in \mathbb{N}. \]

In Section 5, we present more specific information for the cases \( d = 2 \) and 3. It is an easy induction that \( s(n) \) is even if and only if \( n \) is a multiple of 3, so that \( \tau_2 = 0 \). We show that \( \tau_3 = \frac{1}{2} \) and give an explicit formula for \( T(2^r; 3, 0) \), as well as a recursive description of those \( n \) for which \( 3 \mid s(n) \). We also prove that, for all \( N \geq 1 \), \( T(N; 3, 1) - T(N; 3, 2) \in \{0, 1, 2, 3\} \).

It will be proved in [16] that

\[ T(2^r; 4, 0) = T(2^r; 5, 0), \quad T(2^r; 6, 0) = T(2^r; 9, 0) = T(2^r; 11, 0); \]

we conjecture that \( T(2^r; 22, 0) = T(2^r; 27, 0) \). (The latter is true for \( r \leq 19 \).) These exhaust the possibilities for \( T(2^r; N_1, 0) = T(2^r; N_2, 0) \) with \( N_i \leq 128 \). Note that \( I(4) = I(5) = 6, I(6) = I(8) = I(9) = I(11) = 12 \) and \( I(22) = I(27) = 36 \). However, \( T(2^r; 8, 0) \neq T(2^r; 6, 0) \), so there is more than just asymptotics at work.
2. Basic facts about the Stern sequence

We formalize the definition of the diatomic array. Define \( Z(r, k) = Z(r, k; a, b) \) recursively for \( r \geq 0 \) and \( 0 \leq k \leq 2^r \) by:

\[
Z(0, 0) = a, \quad Z(0, 1) = b;
\]

\[
Z(r + 1, 2k) = Z(r, k), \quad Z(r + 1, 2k + 1) = Z(r, k) + Z(r, k + 1).
\]

The following lemma follows from (1.2), (2.1) and a simple induction.

**Lemma 2.1.** For \( 0 \leq k \leq 2^r \), we have

\[
(2.2) \quad Z(r, k; 0, 1) = s(k).
\]

Lemma 2.1 leads directly to a general formula for the diatomic array.

**Theorem 2.2.** For \( 0 \leq k \leq 2^r \), we have

\[
(2.3) \quad Z(r, k; a, b) = s(2^r - k)a + s(k)b.
\]

**Proof.** Clearly, \( Z(r, k; a, b) \) is linear in \( (a, b) \) and it also satisfies a mirror symmetry

\[
(2.4) \quad Z(r, k; a, b) = Z(r, 2^r - k; b, a)
\]

for \( 0 \leq k \leq 2^r \), c.f. (1.4). Thus,

\[
(2.5) \quad Z(r, k; a, b) = aZ(r, k; 1, 0) + bZ(r, k; 0, 1) = aZ(r, 2^r - k; 0, 1) + bZ(r, k; 0, 1).
\]

The result then follows from Lemma 2.1. \( \square \)

The diatomic array contains a self-similarity: any two consecutive entries in any row determine the corresponding portion of the succeeding rows. More precisely, we have a relation whose simple inductive proof is omitted, and which immediately leads to the iterated generalization of (1.2).

**Lemma 2.3.** If \( 0 \leq k \leq 2^r \) and \( 0 \leq k_0 \leq 2^{r_0} - 1 \), then

\[
(2.6) \quad Z(r + r_0, 2^r k_0 + k; a, b) = Z(r, k; Z(r_0, k_0; a, b), Z(r_0, k_0 + 1; a, b)).
\]

**Corollary 2.4.** If \( n \geq 0 \) and \( 0 \leq k \leq 2^r \), then

\[
(2.7) \quad s(2^n + k) = s(2^r - k)s(n) + s(k)s(n + 1).
\]

**Proof.** Take \( (a, b, k_0, r_0) = (0, 1, n, \lceil \log_2(n + 1) \rceil) \) in Lemma 2.3, so that \( k_0 < 2^{r_0} \), and then apply Theorem 2.2. \( \square \)

We turn now to \( t(n) \). Clearly, \( t(2n) < 1 \leq t(2n + 1) \) for all \( n \); after a little algebra, (1.2) implies

\[
(2.8) \quad t(2n) = \frac{1}{1 + \frac{1}{t(n)}}, \quad t(2n + 1) = 1 + t(n).
\]

The mirror symmetry (1.4) yields two other formulas which are evident in (1.7):

\[
(2.9) \quad t(2^r + k)t(2^{r+1} - k - 1) = 1,
\]

The mirror symmetry (1.4) yields two other formulas which are evident in (1.7):
for $0 \leq k \leq 2^r - 1$, which follows from

$$t(2^{r+1} - k - 1) = \frac{s(2^{r+1} - k - 1)}{s(2^{r+1} - k)} = \frac{s(2^r + k + 1)}{s(2^r + k)} = \frac{1}{t(2^r + k)};$$

and

$$t(2^r + 2\ell) + t(2^{r+1} - 2\ell - 2) = 1,$$

for $r \geq 1$ and $0 \leq 2\ell \leq 2^r - 2$, which follows from

$$\frac{s(2^r + 2\ell)}{s(2^r + 2\ell + 1)} + \frac{s(2^{r+1} - 2\ell - 2)}{s(2^{r+1} - 2\ell - 1)} = \frac{s(2^r + 2\ell)}{s(2^r + 2\ell + 1)} + \frac{s(2^r + 2\ell + 2)}{s(2^r + 2\ell + 1)},$$

since $s(2m) + s(2m + 2) = s(2m + 1)$.

Although we will not use it directly here, we mention a simple closed formula for $t(n)$, and hence for $s(n)$. Stern had already proved that if $2^r \leq n < 2^{r+1}$, then the sum of the denominators in the continued fraction representation of $t(n)$ is $r+1$; this is clear from (2.8). Lehmer [11] gave an exact formulation, of which the following is a variation. Suppose $n$ is odd and $[n]_2$, the binary representation of $n$, consists of a block of $a_1$ 1’s, followed by $a_2$ 0’s, $a_3$ 1’s, etc, ending with $a_{2v}$ 0’s and $a_{2v+1}$ 1’s, with $a_j \geq 1$. (That is, $n = 2^{a_1+a_2+\cdots+a_{2v+1}} - 2^{a_{2v+2}+\cdots+a_{2v+1}} \pm \cdots \pm 2^{a_{2v+1} - 1}$.) Then

$$t(n) = \frac{s(n)}{s(n + 1)} = \frac{p}{q} = a_{2v+1} + \frac{1}{a_{2v} + \frac{1}{\cdots + \frac{1}{a_1}}}.$$

Conversely, if $\frac{p}{q} > 1$ and (2.13) gives its presentation as a simple continued fraction with an odd number of denominators, then the unique $n$ with $t(n) = \frac{p}{q}$ has the binary representation described above. (If $n$ is even or $\frac{p}{q} < 1$, apply (2.9) first.)

The Stern-Brocot array is named after the clockmaker Achille Brocot, who used it [3] in 1861 as the basis of a gear table; see also [9]. This array caught the attention of several French number theorists, and is discussed in [12]. It is formed by applying the diatomic rule to numerators and denominators simultaneously:

$$(r = 0) \begin{array}{ccc} 0 & 1 & 1 \\ 1 & 0 & 0 \end{array}$$

$$(r = 1) \begin{array}{ccc} 0 & 1 & 1 \\ 1 & 1 & 0 \end{array}$$

$$(r = 2) \begin{array}{ccc} 1 & 2 & 1 \\ 2 & 1 & 0 \end{array}$$

$$(r = 3) \begin{array}{ccc} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 2 & 1 & 1 \\ 3 & 2 & 1 \\ 1 & 0 & 0 \end{array}$$


This array is not quite the same as (1.7). If $\frac{a}{b}$ and $\frac{c}{d}$ are consecutive in the $r$-th row, then they repeat in the $(r + 1)$-st row, separated by $\frac{a+c}{b+d}$. It is easy to see that the elements of the $r$-th row are $\frac{s(k)}{s(2^{r+1} - k)}$, $0 \leq k \leq 2^r$. It is also easy to show that the
elements of each row are increasing, and moreover, that they share a property with the Farey sequence.

**Lemma 2.5.** For $0 \leq k \leq 2^r - 2$,

$$
\frac{s(k+1)}{s(2^r - k - 1)} - \frac{s(k)}{s(2^r - k)} = \frac{1}{s(2^r - k)s(2^r - k - 1)}.
$$

That is,

$$
(2.16) \quad s(k+1)s(2^r - k) - s(k)s(2^r - k - 1) = 1.
$$

This lemma has a simple proof by induction, which can be found in [12], p.467 and [8], p.117.

The “new” entries in the $(r+1)$-st row of (2.14) are a permutation of the $r$-th row of (1.7). The easiest way to express the connection (see [16]) for rationals $\frac{p}{q} > 1$ is that if $0 < k < 2^r$ is odd, then

$$
(2.17) \quad \frac{p}{q} = \frac{s(2^r + k)}{s(2^r - k)} = \frac{s(2^r + k)}{s(2^r + k + 1)},
$$

where $\frac{p}{q}$ denotes the integer so that $[n]_2$ and $[\overline{n}]_2$ are the reverse of each other. If $\frac{p}{q} < 1$, then apply mirror symmetry to the instance of (2.17) which holds for $\frac{x}{y}$.

The Müntzian $?\text{-functions}$ can be defined using the first half of the rows of (2.14). For odd $\ell$, $0 \leq \ell \leq 2^r$,

$$
(2.18) \quad ?\left(\frac{s(\ell)}{s(2^r+1 - \ell)}\right) = \frac{\ell}{2^r}.
$$

This gives a strictly increasing map from $\mathbb{Q} \cap [0,1]$ to the dyadic rationals in $[0,1]$, which extends to a continuous strictly increasing map from $[0,1]$ to itself, taking quadratic irrationals to non-dyadic rationals.

Finally, suppose $N$ is a positive integer, written as

$$
(2.19) \quad N = 2^{r_1} + 2^{r_2} + \cdots + 2^{r_v}, \quad r_1 > r_2 > \cdots > r_v.
$$

We shall define

$$
(2.20) \quad N_0 = 0; \quad N_j = 2^{r_1} + \cdots + 2^{r_j} \text{ for } j = 1, \ldots, v.
$$

Further, for $1 \leq j \leq v$, let $M_j = 2^{-r_j} N_{j+1}$, so that

$$
(2.21) \quad N_j = N_{j-1} + 2^{r_j} = 2^{r_j}(M_j + 1) = 2^{r_j-1} M_{j-1}.
$$

and, for $a < b \in \mathbb{Z}$, let

$$
(2.22) \quad [a,b) := \{ k \in \mathbb{Z} : a \leq k < b \}.
$$

Our proofs will rely on the observation that

$$
(2.23) \quad [0, N) = \bigcup_{j=0}^{v-1} [N_j, N_{j+1}) = \bigcup_{j=1}^{v} [2^{r_j} M_j, 2^{r_j}(M_j + 1)),
$$

\nu
where the above unions are disjoint, so that, formally,

$$
\sum_{n=0}^{N-1} = \sum_{j=0}^{v-1} \sum_{n=N_j}^{N_j+1-1} = \sum_{j=1}^{v} \sum_{n=2jM_j}^{2j(M_j+1)-1}.
$$

### 3. The Stern-Average Rational

We begin by looking at the sum of $t(n)$ along the rows of (1.7). Let

$$
A(r) = \sum_{n=2^r}^{2^{r+1}-1} t(n) \quad \text{and} \quad \tilde{A}(r) = \sum_{n=0}^{2^r-1} t(n) = \sum_{i=0}^{r-1} A(i).
$$

#### Lemma 3.1

For $r \geq 0$,

$$
A(r) = \frac{3}{2} \cdot 2^r - \frac{1}{2} \quad \text{and} \quad \tilde{A}(r) = \frac{3}{2} \cdot 2^r - \frac{r+3}{2}.
$$

**Proof.** First note that $A(0) = t(1) = \frac{1}{2} - \frac{3}{2}$. Now observe that for $r \geq 0$,

$$
A(r+1) = \sum_{j=0}^{2^{r+1}-1} t(2^{r+1} + j) = \sum_{k=0}^{2^r-1} t(2^{r+1} + 2k) + \sum_{k=0}^{2^r-1} t(2^{r+1} + 2k + 1).
$$

Using (2.11) and (2.8), we can simplify this summation:

$$
\sum_{k=0}^{2^r-1} t(2^{r+1} + 2k) = \frac{1}{2} \left( \sum_{k=0}^{2^r-1} t(2^{r+1} + 2k) + t(2^{r+2} - 2k - 2) \right) = 2^{r-1},
$$

and

$$
\sum_{k=0}^{2^r-1} t(2^{r+1} + 2k + 1) = \sum_{k=0}^{2^r-1} \left( 1 + t(2^r + k) \right) = 2^r + A(r).
$$

Thus, $A(r+1) = 2^{r-1} + 2^r + A(r)$, and the formula for $A(r)$ is established by induction. This also immediately implies the formula for $\tilde{A}(r)$.

#### Lemma 3.2

If $m$ is even, then

$$
\tilde{A}(r) \leq \sum_{k=0}^{2^r-1} t(2^r m + k) < A(r).
$$

**Proof.** For fixed $(k, r)$, let

$$
\Phi_{k,r}(x) = \frac{s(2^r - k)x + s(k)}{s(2^r - (k + 1))x + s(k + 1)}.
$$

Then it follows from (2.16) that

$$
\Phi'_{k,r}(x) = \frac{s(k + 1)s(2^r - k) - s(k)s(2^r - k - 1)}{(s(2^r - (k + 1))x + s(k + 1))^2} > 0.
$$
Using (2.7), we see that
\[
t(2^r m + k) = \frac{s(2^r m + k)}{s(2^r m + k + 1)} = \frac{s(2^r - k)s(m) + s(k)s(m + 1)}{s(2^r - k - 1)s(m) + s(k + 1)s(m + 1)}
\]
(3.9)
\[
= \Phi_{k,r} \left( \frac{s(m)}{s(m + 1)} \right) = \Phi_{k,r}(t(m)).
\]
Since \( m \) is even, \( 0 \leq t(m) < 1 \); monotonicity then implies that
\[
t(k) = \Phi_{k,r}(0) \leq t(2^r m + k) < \Phi_{r,k}(1) = t(2^r + k).
\]
(3.10)
Summing (3.10) on \( k \) from 0 to \( 2^r - 1 \) gives (3.6).

We use these estimates to establish (1.8).

**Theorem 3.3.** If \( 2^r \leq N < 2^{r+1} \), then
\[
\frac{3N}{2} - \frac{r^2 + 7r + 6}{4} \leq \sum_{n=0}^{N-1} t(n) < \frac{3N}{2} - \frac{1}{2}.
\]
(3.11)

**Proof.** Recalling (2.24), we apply Lemma 3.2 for each \( j \), with \( r = r_j \) and \( m = M_j \), so that
\[
\frac{3}{2} \cdot 2^{r_j} - \frac{r_j + 3}{2} \leq \sum_{n=N_{j-1}}^{N_j-1} t(n) < \frac{3}{2} \cdot 2^{r_j} - \frac{1}{2}.
\]
(3.12)
After summing on \( j \), we find that
\[
\frac{3N}{2} - \frac{r_1 + \cdots + r_v + 3v}{2} \leq \sum_{n=0}^{N-1} t(n) \leq \frac{3N}{2} - \frac{v}{2}.
\]
(3.13)
To obtain (3.11), note that \( \sum r_j + 3v \leq \frac{r(r+1)}{2} + 3r + 3 = \frac{r^2 + 7r + 6}{2} \).

**Corollary 3.4.**
\[
\sum_{n=0}^{N-1} t(n) = \frac{3N}{2} + \mathcal{O} \left( \log^2 N \right).
\]
(3.14)
Since \( t(2^r - 1) = \frac{r}{2} \), the true error term is at least \( \mathcal{O}(\log N) \). Numerical computations using Mathematica suggest that \( \log^2 N \) can be replaced by \( \log N \log \log N \). It also seems that, at least for small fixed positive integers \( t \),
\[
\alpha_t := \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \frac{s(n)}{s(n + t)}
\]
exists. We have seen that \( \alpha_1 = \frac{3}{2} \), and if they exist, the evidence suggests that \( \alpha_2 \approx 1.262 \), \( \alpha_3 \approx 1.643 \) and \( \alpha_4 \approx 1.161 \). We are unable to present an explanation for these specific numerical values.
4. Stern Pairs, mod d

We fix $d \geq 2$ with prime factorization $d = \prod p_i^{\ell_i}$, $e_i \geq 1$, and recall the definitions of $S_d$ and $S_d(n)$ from (1.9) and (1.10). Let

$$N_d = |S_d|,$$

and for $0 \leq i < d$, let

$$N_d(i) = \{|j \mod d : (i \mod d, j \mod d) \in S_d\}|.$$

We now give two lemmas whose proofs rely on the Chinese Remainder Theorem.

**Lemma 4.1.** The map $S_d : \mathbb{N} \rightarrow S_d$ is surjective.

**Proof.** Suppose $\alpha = (i, j) \in S_d$ with $0 \leq i, j \leq d - 1$. We shall show that there exists $w \in \mathbb{N}$ so that $\gcd(i, j + wd) = 1$. Consequently, there exists $n$ with $s(n) = i$ and $s(n + 1) = j + wd$, so that $S_d(n) = \alpha$.

Write $i = \prod \ell q_i^{f_i}$, $f_i \geq 1$, with $q_i$ prime. If $q_i \mid j$, then $q_i \nmid d$. There exists $w \geq 0$ so that $w \equiv d^{-1} (\mod q_i^{f_i})$ if $q_i \mid j$ and $w \equiv 0 (\mod q_i^{f_i})$ if $q_i \nmid j$. Then $j + wd \neq 0 (\mod q_i^{f_i})$ for all $\ell$, so no prime dividing $i$ divides $j + wd$, as desired. \qed

**Lemma 4.2.** For $0 \leq i \leq d - 1$,

$$N_d = d^2 \prod \frac{p_i^2 - 1}{p_i^2} \quad \text{and} \quad N_d(i) = d \prod_{p_i \mid i} \frac{p_i - 1}{p_i}.$$

**Proof.** To compute $N_d$, we use the Chinese Remainder Theorem by counting the choices for $(i \mod p_i^{f_i}, j \mod p_i^{f_i})$ for each $\ell$. Missing are those $(i, j)$ in which $p_i$ divides both $i$ and $j$, and so the total number of classes is $(p_i^{f_i} - p_i^{f_i - 1})^2$ for each $\ell$.

Now fix $i$. If $p_i \mid i$, then $(i, j) \in S_d$ if and only if $p_i \nmid j$; if $p_i \nmid i$, then there is no restriction on $j$. Thus, there are either $p_i^{f_i} - p_i^{f_i - 1}$ or $p_i^{f_i}$ choices for $j$, respectively. \qed

Suppose $\alpha = (i, j) \in S_d$; let $L(\alpha) := (i, i + j)$ and $R(\alpha) = (i + j, j)$, where $i + j$ is reduced mod $d$ if necessary. Then $L(\alpha), R(\alpha) \in S_d$ and the following lemma is immediate.

**Lemma 4.3.** For all $n$, we have $S_d(2n) = L(S_d(n))$ and $S_d(2n + 1) = R(S_d(n))$.

We now define the directed graph $G_d$ as follows. The vertices of $G_d$ are the elements of $S_d$. The edges of $G_d$ consist of $(\alpha, L(\alpha))$ and $(\alpha, R(\alpha))$ where $\alpha \in S_d$. Iterating, we see that $L^k(\alpha) = (i, i + k)$ and $R^k(\alpha) = (i + k, j)$, so that $L^d = R^d = id$, and $L^{-1} = L^{d-1}$ and $R^{-1} = R^{d-1}$. Thus, if $(\alpha, \beta)$ is an edge of $G_d$, then there is a walk of length $d - 1$ from $\beta$ to $\alpha$.

Each vertex of $G_d$ has out-degree two; since $(L^{-1}(\alpha), \alpha)$ and $(R^{-1}(\alpha), \alpha)$ are edges, each vertex has in-degree two as well. Let $M_d = [m_{\alpha, \beta}] = [m_{\alpha, \beta}]$ denote the adjacency matrix for $G_d$: $M_d$ is the $N_d \times N_d$ 0-1 matrix so that $m_{\alpha L(\alpha)} = m_{\alpha R(\alpha)} = 1$, with other entries equal to 0. For a positive integer $r$, write

$$M_d^r = [m_{\alpha, \beta}^r];$$
then $m^{(r)}_{\alpha\beta}$ is the number of walks of length $r$ from $\alpha$ to $\beta$. Finally, for $\gamma \in S_d$, and integers $U_1 < U_2$, let

(4.5) \[ B(\gamma; U_1, U_2) = \{|m : U_1 \leq m < U_2 \& S_d(m) = \gamma\} \]

The following is essentially equivalent to Lemma 2.3.

**Lemma 4.4.** Suppose $\alpha = S_d(m)$, $\beta \in S_d$ and $r \geq 1$. Then $B(\beta; 2^r m, 2^r (m + 1)) = m^{(r)}_{\alpha\beta}$ is equal to the number of walks of length $r$ in $G_d$ from $\alpha$ to $\beta$.

**Proof.** The walks of length 1 starting from $\alpha$ are $(\alpha, L(\alpha))$ and $(\alpha, R(\alpha))$; that is, $(S_d(n), S_d(2n))$ and $(S_d(n), S_d(2n + 1))$. The rest is an easy induction. \(\square\)

**Lemma 4.5.** For sufficiently large $N$, $m^{(N)}_{\alpha\beta} > 0$ for all $\alpha, \beta$.

**Proof.** Let $\alpha_0 = (0, 1) = S_d(0)$. Note that $L(\alpha_0) = \alpha_0$, hence if there is a walk of length $w$ from $\alpha_0$ to $\gamma$, then there are such walks of every length $\geq w$. By Lemma 4.1, for each $\alpha \in S_d$, there exists $n_\alpha$ so that $S_d(n_\alpha) = \alpha$. Choose $r$ sufficiently large that $n_\alpha < 2^r$ for all $\alpha$. Then by Lemma 4.4, for every $\gamma$, there is a walk of length $r$ from $\alpha_0$ to $\gamma$, and so there is a walk of length $(d - 1)r$ from $\gamma$ to $\alpha_0$. Thus, for any $\alpha, \beta \in S_d$, there is at least one walk of length $dr$ from $\alpha$ to $\beta$ via $\alpha_0$. \(\square\)

We need a version of Perron-Frobenius. Observe that $A_d = \frac{1}{2} M_d$ is doubly stochastic and the entries of $A_d^N = 2^{-N} M_d^N$ are positive for sufficiently large $N$. Thus $A_d$ is irreducible (see [13], Ch.1), so it has a simple eigenvalue of 1, and all its other eigenvalues are inside the unit disk. It follows that $M_d$ has a simple eigenvalue of 2. Let

(4.6) \[ f_d(T) = T^k + c_{k-1} T^{k-1} + \cdots + c_0 \]

be the minimal polynomial of $M_d$. Let $\rho_d < 2$ be the maximum modulus of any non-2 root of $f_d$, and let $1 + \sigma_d$ be the maximum multiplicity of any such maximal root. Then for $r \geq 0$ and all $(\alpha, \beta)$,

(4.7) \[ m^{r+k}_{\alpha\beta} + c_{k-1} m^{r+k-1}_{\alpha\beta} + \cdots + c_0 m^r_{\alpha\beta} = 0. \]

It follows from the standard theory of linear recurrences that for some constants $c_{\alpha\beta}$,

(4.8) \[ m^r_{\alpha\beta} = c_{\alpha\beta} 2^r + (r^{\sigma_d} \rho_d^r) \quad \text{as } r \to \infty. \]

In particular, $\lim_{r \to \infty} A^r_d = A_{d0} := [c_{\alpha\beta}]$, and since $A^{r+1}_d = A_d A^r_d$, it follows that each column of $A_{d0}$ is an eigenvector of $A_d$, corresponding to $\lambda = 1$. Such eigenvectors are constant vectors and since $A_{d0}$ is doubly stochastic, we may conclude that for all $(\alpha, \beta)$, $c_{\alpha\beta} = \frac{1}{N_d}$. Then there exists $c_d > 0$ so that for $r \geq 0$ and all $(\alpha, \beta)$,

(4.9) \[ \left| m^r_{\alpha\beta} - \frac{2^r}{N_d} \right| < c_d r^{\sigma_d} \rho_d^r. \]
Computations show that for for small values of $d$ at least, $\rho_d = \frac{1}{2}$ and $\sigma_d = 0$. In any event, by choosing $2 > \bar{\rho}_d > \rho_d$ if $\sigma_d > 0$, we can replace $\tau_d^{(2)}\rho_d^{(2)}$ by $\bar{\rho}_d^{(2)}$ in the upper bound. Putting this together, we have proved the following theorem.

**Theorem 4.6.** There exist constants $c_d$ and $\bar{\rho}_d < 2$ so that if $m \in \mathbb{N}$ and $\alpha \in S_d$, then for all $r \geq 0$,

$$B(\alpha; 2^r m, 2^r (m + 1)) - \frac{2^r}{N_d} < c_d \bar{\rho}_d^{(2)}.$$  

We now use this result on blocks of length $2^r$ to get our main theorem.

**Theorem 4.7.** For fixed $d \geq 2$, there exists $\tau_d < 1$ so that, for all $\alpha \in S_d$,

$$B(\alpha; 0, N) = \frac{N}{N_d} + O(N^{\tau_d}).$$

**Proof.** By (2.25), we have

$$B(\alpha; 0, N) = \sum_{j=0}^{v-1} B(\alpha; N_j, N_{j+1}) = \sum_{j=1}^{v} B(\alpha; 2^j M_j, 2^j (M_j + 1)).$$

It follows that

$$B(\alpha; 0, N) - \frac{N}{N_d} \leq c_d \bar{\rho}_d^{(1)} + \cdots + \bar{\rho}_d^{(v)}.$$  

If $\bar{\rho}_d \leq 1$, the upper bound is $O(\tau_1) = O(\log N) = O(N^{\epsilon})$ for any $\epsilon > 0$. If $1 \leq \bar{\rho}_d < 2$, the upper bound is $O(\bar{\rho}_d^{(1)}) = O(N^{\tau_d})$ for $\tau_d = \frac{\log \bar{\rho}_d}{\log 2}$, since $N \leq 2^{\tau_1 + 1}$. \hfill \Box

Using the notation (1.11), we have

$$T(N; d, i) = \sum_{\alpha = (i, j) \in S_d} B(\alpha; 0, N),$$

and the following is an immediate consequence of Lemma 4.2 and Theorem 4.7.

**Corollary 4.8.** Suppose $d \geq 2$. Then

$$T(N; d, i) = r_{d,i} N + O(N^{\tau_d}),$$

where, recalling that $d = \prod p_i^{\epsilon_i}$,

$$r_{d,i} = \frac{1}{d} \cdot \prod_{p_i | i} \frac{p_i}{p_i - 1} \cdot \frac{p_i^{\epsilon_i}}{p_i^{\epsilon_i} - 1}.$$  

For example, if $p$ is prime, then $f(p, 0) = \frac{1}{p+1}$ and $f(p, i) = \frac{p}{p^{i+1}}$ when $p \nmid i$.

In some sense, the model here is a Markov Chain, if we imagine going from $m$ to $2m$ or $2m + 1$ with equal probability, so that the $B(\beta; 2^r m, 2^r (m + 1))$’s represent the distribution of destinations after $r$ steps. Ken Stolarsky has pointed out that [17] is a somewhat different application of the limiting theory of Markov Chains in a number theoretic setting.
5. Small values of \(d\)

It is immediate to see (and to prove) that \(2 \mid s(n)\) if and only if \(3 \mid n\), thus \(S_2(n)\) cycles among \{(0, 1), (1, 1), (1, 0)\} and \(\tau_2 = 0\). This generalizes to a family of partition sequences. Suppose \(d \geq 2\) is fixed, and let \(b(d; n)\) denote the number of ways that \(n\) can be written in the form

\[
n = \sum_{i \geq 0} \epsilon_i 2^i, \quad \epsilon_i \in \{0, \ldots, d - 1\},
\]

so that \(b(2; n) = 1\). It is shown in [15] that

\[
\sum_{n=0}^{\infty} s(n)X^n = X \prod_{j=0}^{\infty} \left(1 + X^{2^j} + X^{2^{j+1}}\right).
\]

A standard partition argument shows that

\[
\sum_{n=0}^{\infty} b(d; n)X^n = \prod_{j=0}^{\infty} \frac{1 - X^{d \cdot 2^j}}{1 - X^{2^{j+1}}}. 
\]

Thus, \(s(n) = b(3; n - 1)\). An examination of the product in (5.3) modulo 2 shows that \(b(d; n)\) is odd if and only if \(n \equiv 0, 1 \pmod{d}\) (see [15], Theorems 5.2 and 2.14.)

Suppose now that \(d = 3\). Write the 8 elements of \(S_3\) in lexicographic order:

\[
(0, 1), (0, 2), (1, 0), (1, 1), (1, 2), (2, 0), (2, 1), (2, 2).
\]

Then in the notation of the last section,

\[
M_3 = \begin{pmatrix}
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0
\end{pmatrix}.
\]

The minimal polynomial of \(M_3\) is

\[
f_3(T) = T^5 - 2T^4 + T^3 - 4T^2 + 4T = T(T - 1)(T - 2)(T - \mu)(T - \tilde{\mu}),
\]

where

\[
\mu = \frac{-1 + \sqrt{7}i}{2}, \quad \tilde{\mu} = \frac{-1 - \sqrt{7}i}{2}.
\]

Since the roots of \(f_3\) are distinct, we see that for each \((\alpha, \beta) \in S_3\), for \(r \geq 1\), there exist constants \(v_{\alpha\beta r}\) so that

\[
m_{\alpha\beta}^{(r)} = v_{\alpha\beta 1} + v_{\alpha\beta 2}\mu^r + v_{\alpha\beta 3}\tilde{\mu}^r + \frac{1}{8} \cdot 2^r = \frac{1}{8} \cdot 2^r + \mathcal{O}(2^{r/2}).
\]
(As it happens, there are only eight distinct sequences \( m_{a,b}^{(r)} \).) Corollary 4.8 then implies that
\[
T(N; 3, 0) = \frac{N}{4} + O(\sqrt{N}),
\]
\[
T(N; 3, 1) = \frac{3N}{8} + O(\sqrt{N}), \quad T(N; 3, 2) = \frac{3N}{8} + O(\sqrt{N}).
\]
Since \( T(N; 3, 0) + T(N; 3, 1) + T(N; 3, 2) = N \), we gain complete information from studying \( T(N; 3, 0) \) and
\[
\Delta(N) = \Delta_3(N) := T(N; 3, 1) - T(N; 3, 2).
\]
(That is, \( \Delta_3(N + 1) - \Delta_3(N) \) equals 0, 1, −1 when \( s(N) \equiv 0, 1, 2 \) mod 3, respectively.)

To study \( T(N; 3, 0) \), we first define the set \( A_3 \subset \mathbb{N} \) recursively by:
\[
0, 5, 7 \in A_3, \quad 0 < n \in A_3 \implies 2n, 8n \pm 5, 8n \pm 7 \in A_3.
\]
Thus,
\[
A_3 = \{0, 5, 7, 10, 14, 20, 28, 33, 35, 40, 45, 47, 49, 51, 56, 61, 63, \ldots \}.
\]

**Theorem 5.1.** If \( n \geq 0 \), then \( 3 \mid s(n) \) if and only if \( n \in A_3 \).

**Proof.** It follows recursively from (1.2) or directly from (2.7) that
\[
s(2n) = s(n), \quad s(8n \pm 5) = 2s(n) + 3s(n \pm 1), \quad s(8n \pm 7) = s(n) + 3s(n \pm 1).
\]
Thus, 3 divides \( s(n) \) if and only if 3 divides \( s(2n), s(8n \pm 5) \) or \( s(8n \pm 7) \). Since every \( n > 1 \) can be written uniquely as \( 2n', 8n' \pm 5 \) or \( 8n' \pm 7 \) with \( 0 \leq n' < n \), the description of \( A_3 \) is complete. \( \square \)

In the late 1970’s, E. W. Dijkstra [7](pp. 215–6, 230–232) studied the Stern sequence under the name “fusc”, and gave a different description of \( A_3 \) (p. 232):

Inspired by a recent exercise of Don Knuth I tried to characterize the arguments \( n \) such that \( 3 \mid fusc(n) \). With braces used to denote zero or more instances of the enclosed, the vertical bar as the BNF ‘or’, and the question mark ‘?’ to denote either a 0 or a 1, the syntactical representation for such an argument (in binary) is \( \{0\}1\{?0\}1\{0\}?1\{0\}?1\{0\} \).

I derived this by considering – as a direct derivation of my program – the finite state automaton that computes \( fusc(N) \) mod 3.

Let
\[
a_r = |\{ n \in A_3 : 2^r \leq n < 2^{r+1} \}| = T(2^{r+1}; 3, 0) - T(2^r; 3, 0).
\]
It follows from (5.12) that
\[
a_0 = a_1 = 0, \quad a_2 = a_3 = a_4 = 2, \quad a_5 = 10.
\]

**Lemma 5.2.** For \( r \geq 3 \), \( (a_r) \) satisfies the recurrence
\[
a_r = a_{r-1} + 4a_{r-3}.
\]
Proof. This is evidently true for \( r = 3, 4, 5 \). If \( 2^r \leq n < 2^{r+1} \) and \( n = 2n' \), then
\[ 2^{r-1} \leq n' < 2^r, \]
so the even elements of \( A_3 \) counted in \( a_r \) come from elements of \( A_3 \) counted in \( a_{r-1} \). If \( 2^r \leq n < 2^{r+1} \) and \( n = 8n' \pm 5 \) or \( n = 8n' \pm 7 \), then \( 2^{r-3} < n' < 2^{r-2} \) and \( n' \in A_3 \). Thus the odd elements of \( A_3 \) counted in \( a_r \) come (in fours) from elements of \( A_3 \) counted in \( a_{r-3} \).

The characteristic polynomial of the recurrence (5.16) is \( T^3 - T^2 - 4 \) (necessarily a factor of \( f_3(T) \)), and has roots \( T = 2, \mu \) and \( \bar{\mu} \). The details of the following routine computation are omitted.

**Theorem 5.3.** For \( r \geq 0 \), we have the exact formula

\[
(5.17) \quad a_r = \frac{1}{4} \cdot 2^r + \left( \frac{-7 + 5\sqrt{7}i}{56} \right) \mu^r + \left( \frac{-7 - 5\sqrt{7}i}{56} \right) \bar{\mu}^r.
\]

Keeping in mind that \( s(0) = 0 \) is not counted in any \( a_r \), we find after a further computation that the error estimate \( \mathcal{O}(\sqrt{N}) \) is best possible for \( T(N;3,0) \):

**Corollary 5.4.**

\[
(5.18) \quad T(2^r;3,0) = \frac{1}{4} \cdot 2^r + \left( \frac{7 - \sqrt{7}i}{56} \right) \mu^r + \left( \frac{7 + \sqrt{7}i}{56} \right) \bar{\mu}^r + \frac{1}{2}.
\]

To study \( \Delta(N) \), we first need a somewhat surprising lemma.

**Lemma 5.5.** For all \( N \), \( \Delta(2N) = \Delta(4N) \).

*Proof.* The simplest proof is by induction, and the assertion is trivial for \( N = 0 \). There are eight possible “short” diatomic arrays modulo 3:

\[
\begin{array}{ccccccccc}
& s(N) & s(2N) & s(2N+1) & s(4N) & s(4N+1) & s(4N+2) & s(4N+3) & s(4N+4) \\
0 & 1 & 0 & 2 & 1 & 0 & 1 & 1 & 1 \\
0 & 1 & 1 & 2 & 1 & 0 & 2 & 1 & 2 \\
1 & 2 & 0 & 2 & 1 & 1 & 0 & 1 & 2 \\
1 & 0 & 2 & 2 & 0 & 2 & 1 & 2 & 1 \\
1 & 1 & 0 & 2 & 2 & 2 & 0 & 1 & 0 \\
1 & 1 & 1 & 2 & 1 & 2 & 0 & 1 & 1 \\
1 & 0 & 2 & 0 & 2 & 1 & 1 & 0 & 2 \\
1 & 1 & 0 & 2 & 2 & 0 & 1 & 1 & 0 \\
\end{array}
\]

By counting the elements in the rows mod 3 in each case, we see that \( \Delta(2N+2) - \Delta(2N) = \Delta(4N+4) - \Delta(4N) \) is equal to: \( 1, -1, 2, 0, 1, -2, -1, 0 \), respectively. \( \square \)
Theorem 5.6. For all \( n \), \( \Delta(n) \in \{0, 1, 2, 3\} \). More specifically,

\[
\begin{align*}
S_3(m) &= (0, 1) \implies \Delta(2m) = 0, \Delta(2m + 1) = 0; \\
S_3(m) &= (0, 2) \implies \Delta(2m) = 3, \Delta(2m + 1) = 3; \\
S_3(m) &= (1, \ast) \implies \Delta(2m) = 1, \Delta(2m + 1) = 2; \\
S_3(m) &= (2, \ast) \implies \Delta(2m) = 2, \Delta(2m + 1) = 1.
\end{align*}
\]

(5.20)

Proof. To prove the theorem, we first observe that it is correct for \( m \leq 4 \). We now assume it is true for \( m \leq 2^r \) and prove it for \( 2^r \leq m < 2^{r+1} \). There are sixteen cases: \( m \) can be even or odd and there are eight choices for \( S_3(m) \). As a representative example, suppose \( S_3(m) = (2, 1) \). We shall consider the cases \( m = 2t \) and \( m = 2t + 1 \) separately. The proofs for the other seven choices of \( S_3(m) \) are very similar and are omitted.

Suppose first that \( m = 2t < 2^{r+1} \). Then \( S_3(m) = S_3(2t) = (2, 1) \), hence \( S_3(t) = (2, 2) \). We have \( \Delta(2m) = 2 \) by hypothesis, and hence \( \Delta(4m) = 2 \) by Lemma 5.5. The eighth array in (5.19) shows that \( s(4t) \equiv 2 \mod 3 \), so that \( \Delta(4m+1) = \Delta(4m) - 1 = 1 \), as asserted in (5.20).

If, on the other hand, \( m = 2t + 1 < 2^{r+1} \) and \( S_3(m) = S_3(2t + 1) = (2, 1) \), then \( S_3(t) = (1, 1) \). We now have \( \Delta(2t) = 1 \) and \( \Delta(2t+1) = 2 \) by hypothesis and \( \Delta(4t) = 1 \) by Lemma 5.5. The fourth array in (5.19) shows that \( (s(4t), s(4t + 1), s(4t + 2)) \equiv (1, 0, 2) \mod 3 \). Thus, it follows that \( \Delta(2m) = \Delta(4t + 2) = \Delta(4t + 1 + 0) = 2 \) and \( \Delta(2m + 1) = \Delta(4t + 3) = \Delta(4t + 2) - 1 = 1 \), again as desired. \( \square \)

Since \( S_3(m) \) is uniformly distributed on \( S_3 \), (5.20) shows that \( \Delta(n) \) takes the values \( (0, 1, 2, 3) \) with limiting probability \( \left( \frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8} \right) \).

We conclude with a few words about the results announced at the end of the first section, but not proved here. For each \( (d, i) \), \( T(2^r; d, i) \) will satisfy a recurrence whose characteristic equation is a factor of the minimal polynomial of \( S_d \). It happens that \( T(2^r; 4, 0) = T(2^r; 5, 0) \) for small values of \( r \) and both satisfy the recurrence with characteristic polynomial \( T^4 - 2T^3 + T^2 - 4 \) (roots: 2, -1, -\( \tau \), -\( \bar{\tau} \)) so that equality holds for all \( r \). The same applies to \( T(2^r; 6, 0) = T(2^r; 9, 0) = T(2^r; 11, 0) \), with a more complicated recurrence. Results similar to Lemma 5.5 and Theorem 5.6 hold for \( d = 4 \), with a similar proof; Antonios Hondroulis has shown that this is also true for \( d = 6 \). No result has been found yet for \( d = 5 \), although a Mathematica check for \( N \leq 2^{19} \) shows that \( -5 \leq T(N; 5, 1) - T(N; 5, 4) \leq 11 \). These topics will be discussed in greater detail in [16].

References


Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801
E-mail address: reznick@math.uiuc.edu