Enumerating the rationals from left to right

S. P. GLASBY

2010 Mathematics subject classification: 11B57, 11B83, 11B75

There are three well-know sequences used to enumerate the rationals: the Stern-Brocot sequences $\mathcal{SB}_n$, the Calkin-Wilf sequences $\mathcal{CW}_n$, and the Farey sequences $\mathcal{F}_n$. The purpose of this note is to show that all three sequences can be constructed (left-to-right) using almost identical recurrence relations. The Stern-Brocot (S-B) and Calkin-Wilf (C-W) sequences give rise to complete binary trees related to the following rules:

These trees have many beautiful algebraic, combinatorial, computational, and geometric properties [2, 5, 4]. Well-written introductions to the S-B tree and Farey sequences can be found in [3], and to the C-W tree in [2]. We shall focus on *sequences* rather than *trees*.

Two fractions $\frac{a}{b} < \frac{c}{d}$ are called *adjacent* if $bc - ad = 1$. Adjacent fractions are necessarily reduced, i.e. $\gcd(a, b) = \gcd(c, d) = 1$. The *mediant* of $\frac{a}{b} < \frac{c}{d}$ is $\frac{a+c}{b+d}$. Simple algebra shows if $\frac{a}{b} < \frac{c}{d}$ are adjacent, then $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$ are pairwise adjacent (and hence reduced). The sequences $\mathcal{SB}_n$ are defined recursively: $\mathcal{SB}_0 = [\frac{0}{1}, \frac{1}{0}]$ represents 0 and ∞ as fractions, and $\mathcal{SB}_n$ is computed from $\mathcal{SB}_{n-1}$ by inserting mediants between consecutive fractions. Thus

$$\mathcal{SB}_1 = \left[\frac{0}{1}, \frac{\mathbf{1}}{\mathbf{1}}, \frac{1}{0}\right], \quad \mathcal{SB}_2 = \left[\frac{0}{1}, \frac{\mathbf{1}}{\mathbf{2}}, \frac{1}{1}, \frac{\mathbf{2}}{\mathbf{1}}, \frac{1}{0}\right], \quad \mathcal{SB}_3 = \left[\frac{0}{1}, \frac{\mathbf{1}}{\mathbf{3}}, \frac{1}{2}, \frac{\mathbf{2}}{\mathbf{3}}, \frac{1}{1}, \frac{\mathbf{3}}{\mathbf{2}}, \frac{2}{1}, \frac{\mathbf{3}}{\mathbf{1}}, \frac{1}{0}\right], \ \ldots.$$

A simple induction shows that $|\mathcal{SB}_n| = 2^n + 1$. Thus 2^{n-1} mediants are inserted into $\mathcal{SB}_{n-1}$ to form $\mathcal{SB}_n$. The C-W sequences are defined using the right rule above:

$$\mathcal{CW}_1 := \left[\frac{1}{1}\right], \mathcal{CW}_2 = \left[\frac{1}{2}, \frac{2}{1}\right], \mathcal{CW}_3 = \left[\frac{1}{3}, \frac{3}{2}, \frac{2}{3}, \frac{3}{1}\right], \mathcal{CW}_4 = \left[\frac{1}{4}, \frac{4}{3}, \frac{3}{5}, \frac{5}{2}, \frac{2}{5}, \frac{5}{3}, \frac{3}{4}, \frac{4}{1}\right], \ldots.$$

A simple induction shows that $|\mathcal{CW}_n| = 2^{n-1}$. Another simple induction (see [2, p360]) shows that the fractions in $\mathcal{CW}_n$ have the form

$$\mathcal{CW}_n = \left[\frac{b_{-1}}{b_0}, \frac{b_0}{b_1}, \ldots, \frac{b_{N-2}}{b_{N-1}}\right] \qquad \text{where } N = 2^{n-1},$$

and the denominator of a given fraction is the numerator of the succeeding fraction. Indeed, this property obtains even when the sequences $\mathcal{CW}_1, CW_2, CW_3, \ldots$ are concatenated to form $\mathcal{CW}_\infty := [\frac{1}{\mathbf{1}}, \overbrace{\frac{\mathbf{1}}{2}, \frac{2}{\mathbf{1}}}, \overbrace{\frac{\mathbf{1}}{3}, \frac{3}{2}, \frac{2}{3}, \frac{3}{\mathbf{1}}}, \overbrace{\frac{\mathbf{1}}{4}, \ldots, \frac{4}{\mathbf{1}}}, \ldots] = \left[\frac{c_0}{c_1}, \frac{c_1}{c_2}, \frac{c_2}{c_3}, \ldots\right].$

The Farey sequence of order n contains all the reduced fractions $\frac{p}{q}$ with $0 \leqslant p \leqslant q \leqslant n$, in their natural order. Thus

$$\mathcal{F}_1 = \left[\frac{0}{1}, \frac{1}{1}\right], \ \mathcal{F}_2 = \left[\frac{0}{1}, \frac{\mathbf{1}}{\mathbf{2}}, \frac{1}{1}\right], \ \mathcal{F}_3 = \left[\frac{0}{1}, \frac{\mathbf{1}}{\mathbf{3}}, \frac{1}{2}, \frac{\mathbf{2}}{\mathbf{3}}, \frac{1}{1}\right], \ \mathcal{F}_4 = \left[\frac{0}{1}, \frac{\mathbf{1}}{\mathbf{4}}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{\mathbf{3}}{\mathbf{4}}, \frac{1}{1}\right], \dots.$$

A standard way to compute $\mathcal{F}_n$ from $\mathcal{F}_{n-1}$ is to insert mediants between consecutive fractions of $\mathcal{F}_{n-1}$ only when it gives a denominator of size n (see [3, p118]). Thus $\mathcal{F}_n$ is a subsequence of $\mathcal{SB}_n$. It is easy to prove that $|\mathcal{F}_n| = 1 + \sum_{j=1}^{n} \varphi(j)$ where $\varphi(n)$ denotes the number of reduced fractions $\frac{a}{n}$ with $1 \leqslant a < n$. The mediant rule above implies that consecutive fractions in $\mathcal{SB}_n$ and $\mathcal{F}_n$ are adjacent (see also [3, p119]).

It is shown in [3] and [2] that $\mathcal{SB}_\infty$ and $\mathcal{CW}_\infty$ contain *every* (reduced) positive rational *precisely once*. Although $\mathcal{SB}_n, \mathcal{CW}_n, \mathcal{F}_n$ are defined "top-down" they can be computed from "left to right" via almost identical recurrence relations.

Theorem 1. *Write* $\mathcal{SB}_n = \left[\frac{a_{-1}}{b_{-1}}, \frac{a_0}{b_0}, \frac{a_1}{b_1}, \dots, \frac{a_{N-1}}{b_{N-1}}\right]$ *where* $N = 2^n$. *Then*

(1a) $\quad a_{-1} = 0,\ a_0 = 1, \qquad a_i = k_i a_{i-1} - a_{i-2} \qquad$ for $1 \leqslant i < N$,

(1b) $\quad b_{-1} = 1,\ b_0 = n, \qquad b_i = k_i b_{i-1} - b_{i-2} \qquad$ for $1 \leqslant i < N$,

where $k_i = 2\log_2 |i|_2 + 1$, *and* $|i|_2$ *denotes the largest power of* 2 *dividing* i.

Theorem 2. *Write* $\mathcal{CW}_\infty = \left[\frac{a_0}{a_1}, \frac{a_1}{a_2}, \dots, \frac{a_{i-1}}{a_i}, \dots\right]$ *and* $\mathcal{CW}_n = \left[\frac{b_{-1}}{b_0}, \frac{b_0}{b_1}, \frac{b_1}{b_2}, \dots, \frac{b_{N-2}}{b_{N-1}}\right]$ *where* $N = 2^{n-1}$. *Then the* a_i *and* b_i *can be computed via the recurrence relations*

(2a) $\quad a_{-1} = 0,\ a_0 = 1, \qquad a_i = k_i a_{i-1} - a_{i-2} \qquad$ for $1 \leqslant i < \infty$,

(2b) $\quad b_{-1} = 1,\ b_0 = n, \qquad b_i = k_i b_{i-1} - b_{i-2} \qquad$ for $1 \leqslant i < N$,

where $k_i = 2\log_2 |i|_2 + 1$. [Note that $\nu_2(i) := \log_2 |i|_2$ is the largest $\nu \in \mathbb{Z}$ satisfying $2^\nu \mid i$.]

Theorem 3. *Write the Farey sequence* $\mathcal{F}_n$ *of order* n *as* $\mathcal{F}_n = [\frac{A_{-1}}{B_{-1}}, \frac{A_0}{B_0}, \frac{A_1}{B_1}, \dots]$. *Then the numerators* A_i, *and the denominators* B_i *can be computed via the recurrence relations*

(3a) $\quad A_{-1} = 0,\ A_0 = 1, \qquad A_i = K_i A_{i-1} - A_{i-2} \qquad$ for $1 \leqslant i < N$,

(3b) $\quad B_{-1} = 1,\ B_0 = n, \qquad B_i = K_i B_{i-1} - B_{i-2} \qquad$ for $1 \leqslant i < N$,

where $K_i = \left\lfloor \frac{B_{i-2}+n}{B_{i-1}} \right\rfloor$, *and* $N = \sum_{j=1}^{n} \varphi(j)$.

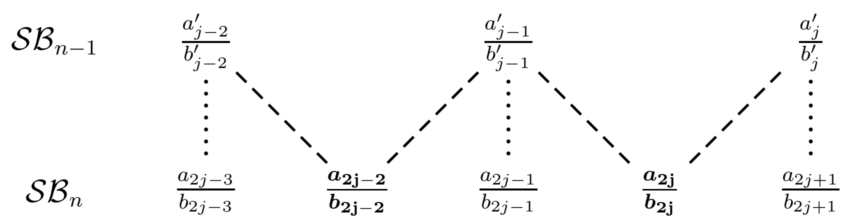
To illustrate Theorem 1, $\mathcal{SB}_4$ can be computed from left to right using the table

i	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
a_i	0	1	1	2	1	3	2	3	1	4	3	5	2	5	3	4	1
b_i	1	4	3	5	2	5	3	4	1	3	2	3	1	2	1	1	0
k_i			1	3	1	5	1	3	1	7	1	3	1	5	1	3	1

.

The numbers k_i are the same as the numbers k'_i generated by the recurrence $k'_1 = 1$, $k'_{2j+1} = 1$, $k'_{2j} = 2k'_j + 1$ for $j \geqslant 0$. (Proof by induction: $k_1 = k'_1$ and $k_{2j+1} = 1$, $k_{2j} = 2k_j + 1$ hold for $j \geqslant 1$ as $|2j+1|_2 = 1$ and $|2j|_2 = 2|j|_2$. Thus $k_i = k'_i$ for all $i \geqslant 1$.)

Proof (of Theorem 1). Our proof uses induction on n. It suffices to prove (1a) as the proof of (1b) is similar (just change the a's to b's). Clearly (1a) is true for $n = 0$ as $\mathcal{SB}_0 = [\frac{0}{1}, \frac{1}{0}]$. Assume $n > 0$ and (1a) is true for $\mathcal{SB}_{n-1}$. Let $\frac{a'_{-1}}{b'_{-1}}, \frac{a'_0}{b'_0}, \ldots, \frac{a'_{N/2-1}}{b'_{N/2-1}}$ be the fractions in $\mathcal{SB}_{n-1}$. The way mediants are inserted to create $\mathcal{SB}_n$ is shown below:



where dotted lines denote the repetition of a fraction, and dashed lines denote the formation of a mediant. The repetition of fractions means

$$a_{2j-1} = a'_{j-1} \quad \text{and} \quad b_{2j-1} = b'_{j-1} \quad \text{for } 0 \leqslant j < N/2, \tag{4}$$

and the formation of mediants means

$$a_{2j} = a'_{j-1} + a'_j \quad \text{and} \quad b_{2j} = b'_{j-1} + b'_j \quad \text{for } 0 \leqslant j < N/2. \tag{5}$$

We prove (1a) using induction on i. Certainly (1a) is true for $i = -1, 0$ as $\mathcal{SB}_n$ starts with $\frac{0}{1}, \frac{1}{n}$. Suppose now that $i \geqslant 1$, and consider the case when i is even and odd separately.
Case 1. $i = 2j$ is even and $j \geqslant 1$. The following shows that (1a) holds for even i:

$$\begin{aligned}
k_{2j}a_{2j-1} - a_{2j-2} &= (k_j + 2)a_{2j-1} - a_{2j-2} && \text{as } k_{2j} = k_j + 2,\\
&= (k_j + 2)a'_{j-1} - (a'_{j-2} + a'_{j-1}) && \text{by (4) and (5),}\\
&= k_j a'_{j-1} - a'_{j-2} + a'_{j-1} && \text{canceling } a'_{j-1},\\
&= a'_j + a'_{j-1} && \text{as } a'_j = k_j a'_{j-1} - a'_{j-2} \text{ by induction,}\\
&= a_{2j} && \text{by (5).}
\end{aligned}$$

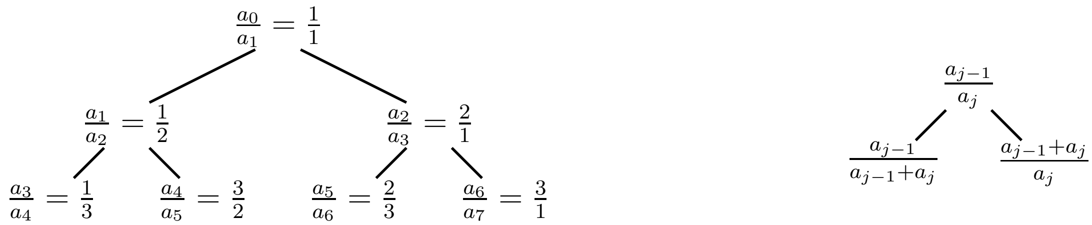
Case 2. $i = 2j + 1$ is odd and $j \geqslant 1$. Note that $k_{2j+1} = 1$, $a_{2j-1} = a'_{j-1}$, and $a_{2j+1} = a'_j$ by (4). These equations and (5) now imply

$$k_{2j+1}a_{2j} - a_{2j-1} = a_{2j} - a_{2j-1} = (a'_{j-1} + a'_j) - a'_{j-1} = a'_j = a_{2j+1}$$

as desired. This completes the inductions on i and n. □

A different (and very interesting) method for computing terms of $\mathcal{SB}_n$ is given in [1]. It uses continued fraction expansions and "normal additive factorizations". As the recurrence (1a) is independent of n, the numerators for $\mathcal{SB}_{n-1}$ reappear as the first $2^{n-1} + 1$ numerators for $\mathcal{SB}_n$. We now show that (half of) the denominators b_i in $\mathcal{SB}_n$ reappear (remarkably!) for $\mathcal{CW}_n$, and the numerators a_i also reappear in $\mathcal{CW}_\infty$. Accordingly, we shall use the *same notation* a_i, b_i in Theorem 2 as in Theorem 1.

Proof (of Theorem 2). The following diagram of the C-W tree (with rules)



shows that the numbers a_i must satisfy the recurrence relation:

$$a_0 = 1, \quad a_{2j-1} = a_{j-1} \quad \text{and} \quad a_{2j} = a_{j-1} + a_j \quad \text{for } j > 0. \tag{6}$$

The different recurrence relations (6) and (2a) determine the values $a_0, a_1, a_2, \ldots$. We must prove, therefore, that both recurrence relations generate the *same* numbers. For clarity, we write the numbers produced by (6) as a'_i. Thus

$$a'_{-1} = 0, \quad a'_0 = 1, \quad a'_{2j-1} \overset{(7.1)}{=} a'_{j-1} \quad \text{and} \quad a'_{2j} \overset{(7.2)}{=} a'_{j-1} + a'_j \quad \text{for } j \geqslant 0. \tag{7}$$

(Note that the definition $a'_{-1} := 0$ is consistent with $a'_{2j-1} = a'_{j-1}$ and $a'_{2j} = a'_{j-1} + a'_j$ when $j = 0$.) Our goal is to prove a_i defined by (2a) equals a'_i defined by (7) for $i \geqslant -1$.

We use induction on i. Certainly $a_i = a'_i$ holds for $i = -1, 0$. Assume $i \geqslant 1$ and $a_0 = a'_0$, $a_1 = a'_1, \ldots, a_{i-1} = a'_{i-1}$. Consider the cases when i is odd and even separately.

Case 1. $i = 2j - 1$ where $j \geqslant 1$. Then

$$\begin{aligned} a_{2j-1} &= k_{2j-1} a_{2j-2} - a_{2j-3} && \text{by (2a),} \\ &= a_{2j-2} - a_{2j-3} && \text{as } k_{2j-1} = 1, \\ &= a'_{2j-2} - a'_{2j-3} && \text{by induction on } i, \\ &= a'_{j-2} + a'_{j-1} - a'_{j-2} && \text{by (7.2) and (7.1),} \\ &= a'_{2j-1} && \text{by (7.1).} \end{aligned}$$

Case 2. $i = 2j$ where $j \geqslant 1$. Then

$$\begin{aligned} a_{2j} &= k_{2j} a'_{2j-1} - a_{2j-2} && \text{by (2a) and Case 1,} \\ &= (k_j + 2) a'_{2j-1} - a'_{2j-2} && \text{by } k_{2j} = k_j + 2 \text{ and induction,} \\ &= (k_j + 2) a'_{j-1} - (a'_{j-2} + a'_{j-1}) && \text{by (7.1) and (7.2),} \\ &= k_j a'_{j-1} - a'_{j-2} + a'_{j-1} && \text{canceling } a'_{j-1}, \\ &= a'_j + a'_{j-1} && \text{by induction on } i \text{ and (2a),} \\ &= a'_{2j} && \text{by (7.1).} \end{aligned}$$

This completes the inductive proof of (2a).

The proof of (2b) is now straightforward. As $\mathcal{CW}_n$ is a subsequence of $\mathcal{CW}_\infty$, there exists an m for which $\frac{a_{m-1}}{a_m}$ equals the first fraction $\frac{b_{-1}}{b_0} = \frac{1}{n}$ of $\mathcal{CW}_n$. Thus $a_{m-1} = b_{-1} = 1$ and $a_m = b_0 = n$. Since the recurrences (1a) and (1b) have the same form, it follows that $a_{m+i} = b_i$ for $1 \leqslant i < N$. Thus (1b) holds and (2b), which is the same, also holds. □

Theorem 3 is previously known (see Exercise 4.61 in [3, p150]). We include Theorem 3 and its proof both for comparison with Theorems 1 and 2, and for the reader's convenience.

Proof (of Theorem 3). Our proof uses induction on i. As the first two fractions of $\mathcal{F}_n$ are $\frac{0}{1}$ and $\frac{1}{n}$, the recurrences (3a,b) are correct for $i = -1, 0$. Suppose now that $i > 0$ and that (3a,b) are correct for subscripts less than i. Thus $\frac{A_{i-2}}{B_{i-2}}$ and $\frac{A_{i-1}}{B_{i-1}}$ are consecutive fractions of $\mathcal{F}_n$, and we wish to show that the next fraction is $\frac{A_i}{B_i}$ where $A_i = K_iA_{i-1} - A_{i-2}$ and $B_i = K_iB_{i-1} - B_{i-2}$. As consecutive Farey fraction are adjacent (i.e they satisfy $bc - ad = 1$), we know by induction that $A_{i-1}B_{i-2} - B_{i-1}A_{i-2} = 1$. However, the recurrences (3a,b) extend this property as

$$(8) \qquad \begin{aligned} A_iB_{i-1} - B_iA_{i-1} &= (K_iA_{i-1} - A_{i-2})B_{i-1} - (K_iB_{i-1} - B_{i-2})A_{i-1} \\ &= A_{i-1}B_{i-2} - B_{i-1}A_{i-2} = 1. \end{aligned}$$

Consider the inequalities $\frac{B_{i-2}+n}{B_{i-1}} - 1 < K_i \leqslant \frac{B_{i-2}+n}{B_{i-1}}$. Multiplying by B_{i-1} and subtracting B_{i-2} gives $n - B_{i-1} < B_i \leqslant n$. It follows from (8) and $0 < B_i \leqslant n$ that $\frac{A_{i-2}}{B_{i-2}} < \frac{A_{i-1}}{B_{i-1}} < \frac{A_i}{B_i}$. Suppose that $\frac{a}{b}$ is the next fraction in $\mathcal{F}_n$ after $\frac{A_{i-1}}{B_{i-1}}$. Then we know $\frac{A_{i-1}}{B_{i-1}} < \frac{a}{b} \leqslant \frac{A_i}{B_i}$, and we must show $\frac{a}{b} = \frac{A_i}{B_i}$. If not, then

$$(9) \qquad A_ib - aB_i \overset{(9.1)}{\geqslant} 1 \qquad \text{and} \qquad aB_{i-1} - bA_{i-1} \overset{(9.2)}{\geqslant} 1.$$

Multiplying (9.1) by B_{i-1}, and (9.2) by B_i, and then adding gives

$$n < B_{i-1} + B_i \leqslant (A_ib - aB_i)B_{i-1} + B_i(aB_{i-1} - bA_{i-1}) = (A_iB_{i-1} - B_iA_{i-1})b = b.$$

This is a contradiction since $\frac{a}{b} \in \mathcal{F}_n$ has $b \leqslant n$. Hence $\frac{a}{b} = \frac{A_i}{B_i}$. As both fractions are reduced (and $a, b, B_i > 0$), we conclude that $a = A_i$ and $b = B_i$, as desired. □

The On-Line Encyclopedia of Integer Sequences [6] has a wealth of useful information about the sequences $a_0, a_1, a_2, \ldots$ (A002487), and $k_1, k_2, k_3, \ldots$ (A037227), however, the connection in Theorem 2 between these sequences is new. Note that a_n counts the number of ways that n can be written as a sum of powers of 2, each power being used at most twice. For example, $a_4 = 3$ as $2^2 = 2 + 2 = 2 + 1 + 1$. Finally, we remark that each positive fraction $\frac{p}{q}$ can be associated with a string of L's and R's denoting its position in a binary tree [3, p119]. A simple induction (which we omit) shows that the S-B string of $\frac{p}{q}$ equals the *reverse* of the C-W string of $\frac{p}{q}$.

References

[1] B. Bates, M. Bunder and K. Tognetti, Locating terms in the Stern-Brocot tree, *European J. Combin.* **31** (2010), 1020–1033.

[2] N. Calkin and H.S. Wilf, Recounting the rationals, *Amer. Math. Monthly* **107** (2000), 360–363.

[3] R. L. Graham, D. E. Knuth and O. Patashnik, *Concrete Mathematics*, 2nd Ed., Addison-Wesley, 1994.

[4] M. Hockman, Continued fractions and the geometric decomposition of modular transformations, *Quaest. Math.* **29** (2006), 427–446.

[5] M. Niqui, Exact arithmetic on the Stern-Brocot tree, *J. Discrete Algorithms* **5** (2007), 356–379.

[6] On-Line Encyclopedia of Integer Sequences, published electronically at `http://oeis.org`, 2010.

Department of Mathematics, Central Washington University, WA 98926-7424, USA.
`http://www.cwu.edu/~glasbys/`