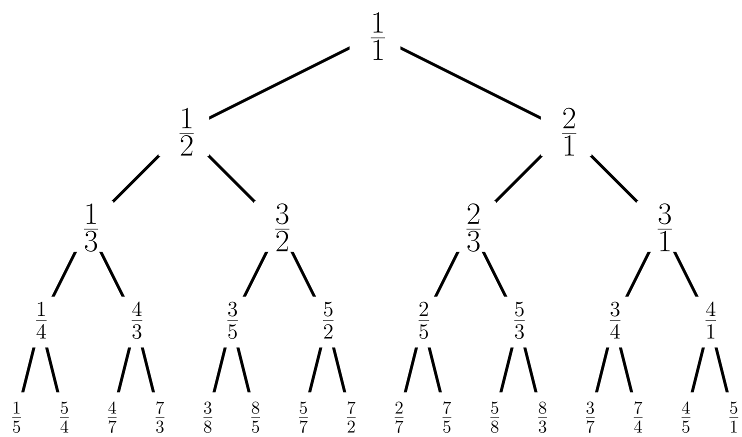
An arborist's guide to the rationals

KATHERINE E. STANGE

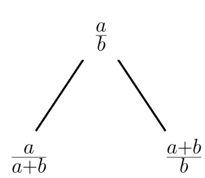
Abstract. There are two well-known ways to enumerate the positive rational numbers in an infinite binary tree: the Farey/Stern-Brocot tree and the Calkin-Wilf tree. In this brief note, we describe these two trees as 'transpose shadows' of a tree of matrices (a result due to Backhouse and Ferreira) via a new proof using yet another famous tree of rationals: the topograph of Conway and Fung.

1. Four Trees

In 2000, Calkin and Wilf studied an explicit enumeration of the positive rationals which naturally arranges itself into an infinite tree [5], the first few levels of which are shown here:



The generation rule is that a parent $\frac{a}{b}$ has the following left and right children:



Date: March 18, 2014, Draft #1.

2010 *Mathematics Subject Classification.* Primary: 11B57, Secondary: 11E99.

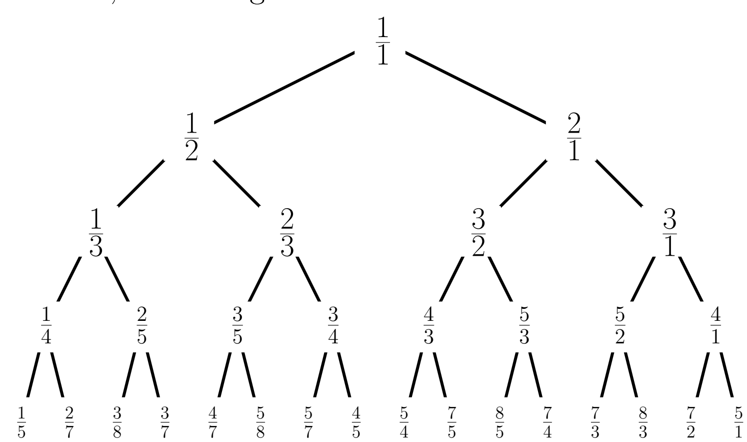
Key words and phrases. special linear group, topograph, Möbius transformation, tree, enumeration.

The author's research has been supported by NSA Grant Number H98230-14-1-0106
.

Every positive rational number appears in this tree exactly once. Calkin and Wilf consider the integer sequence $b(n)$ which enumerates the representations of n as a sum of powers of 2, where each power is allowed to appear at most twice. The function $b(n)/b(n+1)$ reads off the entries in the tree left to right, top row downwards:

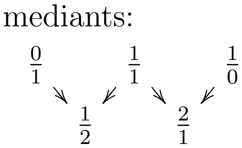
$$\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{3}{2}, \frac{2}{3}, \frac{3}{1}, \frac{1}{4}, \frac{4}{3}, \frac{3}{5}, \dots$$

This tree is reminiscent of the more famous Farey tree, also known as the Stern-Brocot tree, which begins as follows:



The latter name is in honour of two independent descriptions of related ideas in the mid 1800's by Stern [15] and Brocot [4]. Brocot was a french clockmaker who created an array of fractions for the purpose of designing clockwork gears[1]. Stern studied an array of integers, which can be used to generate *both* the Stern-Brocot tree and the Calkin-Wilf tree (it has been quite reasonably suggested the latter tree be called the *Eisenstein-Stern tree*, but this name is not prevalent [2]). The name 'Farey tree' comes from its relationship to Farey sequences[2] (which were themselves likely invented by Charles Haros [8]). For more on the muddy historical waters, and the trees themselves, see [2, 11, 12].

The *mediant* of rationals $\frac{a}{b}$ and $\frac{c}{d}$ (in lowest form) is $\frac{a+c}{b+d}$. The root of the tree is $\frac{1}{1}$, which forms the first row in the tree. Bracket this row by $\frac{0}{1}$ and $\frac{1}{0}$, and then take the list of mediants:



[1]Brocot wrote a book and a paper by the same title, *Calcul des rouages par approximation, nouvelle méthode*; it is the book which contains the array.

[2]Given a bound D, the associated Farey sequence is the sequence of rationals in $[0, 1]$ which, in lowest form, have denominator less than or equal to D. The name has gradually become associated to a wider variety of structures generated by means of the mediant operation we will describe in a moment.

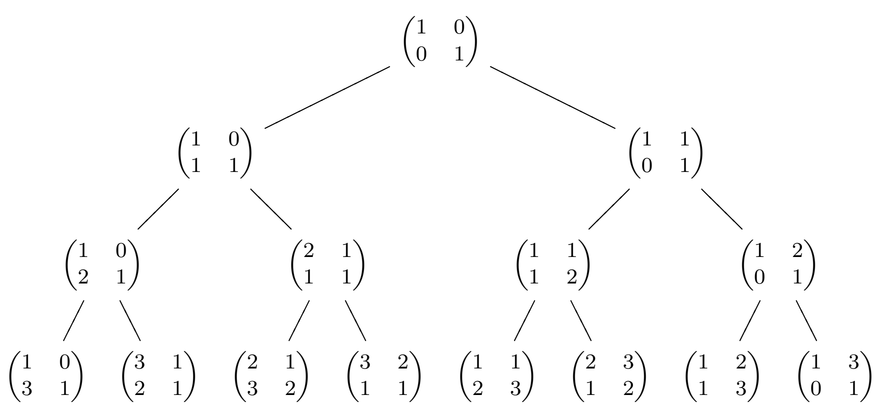
These mediants form the second row of the tree. In general, beginning with the full list of fractions appearing in rows 1 through n, listed in order of size, ones brackets as above, resulting in what has been called a *Brocot sequence* or a *Farey-like sequence*:

$$\frac{0}{1}, \frac{1}{n}, \ldots, \frac{n}{1}, \frac{1}{0}.$$

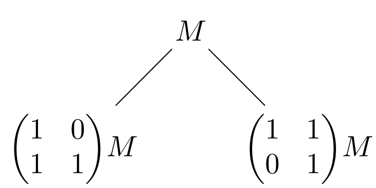
The mediants of the Brocot sequence form the $(n+1)$-st row. Continue ad infinitum and the tree will, just as the Calkin-Wilf tree does, contain exactly one instance of each positive rational number.

It should not be surprising that these two trees share a common genesis. Stern's array gives rise to the *Stern sequence* $s(n)$. The $b(n)$ of the Calkin-Wilf tree is exactly $s(n+1)$ [14], while the fractions of the Farey tree are of the form $s(n)/s(2^r - n)$ (see [12] for history). Some algebraic connections between these two trees are described in [3, 7].

It is the purpose of our arboreal tour to explore a connection between these two famous rational-enumerating trees via a single tree of *matrices*.



To obtain this tree, place the 2×2 identity matrix at the root, and apply the following generation rule:



The matrix tree is a visualization of the folk theorem that the monoid $\mathrm{SL}_2(\mathbb{Z}^{\geq 0})$ is freely generated by the two elements

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix};$$

every element of $\mathrm{SL}_2(\mathbb{Z}^{\geq 0})$ appears in the tree exactly once. See Section 3.

The relationship between these three trees is originally due to Backhouse and Ferreira, and deserves to be better known.

Theorem (Backhouse, Ferreira [1, 2])**.** *To recover the Calkin-Wilf tree, one replaces, in the matrix tree above,*

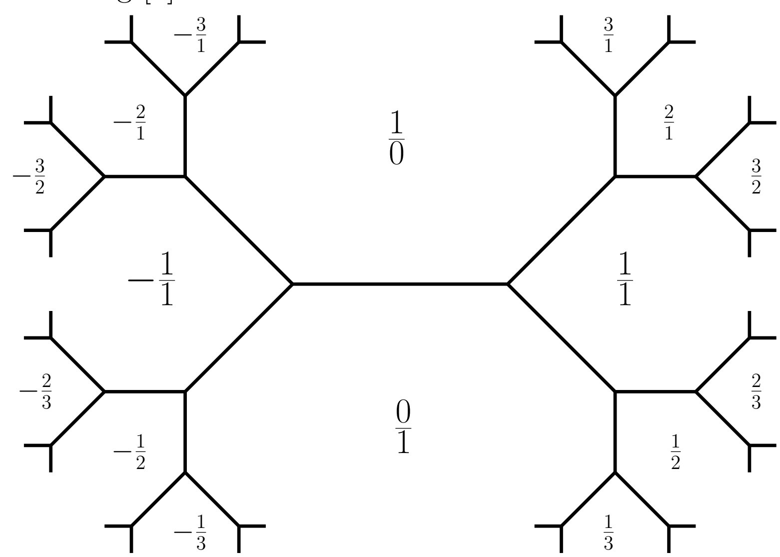
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \textit{with} \quad \frac{a+b}{c+d}.$$

To recover the Farey tree, one replaces

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \textit{with} \quad \frac{d+b}{c+a}.$$

The relation to the Calkin-Wilf tree is immediate by comparing the generation rules of the two trees, and this relationship was exploited in [10, 13].

The relationship to the Farey tree is also not too difficult to verify directly, but it is our purpose to provide a new proof which arises by turning to yet another beautiful tree that enumerates the rationals: the topograph of Conway and Fung [6].



This time, it is the regions that are labelled by the rational numbers. Surrounding each vertex are three fractions which are, in some order (and with appropriate use of signs), a pair of fractions together with their mediant.

Acknowledgements. The author would like to thank João Ferreira for gently bringing to her attention that the result was previously known. She would also like to thank João Ferreira, Jeffrey Lagarias and Jennifer Lansing for help in correcting the historical background contained in this article. Thank you also to Jesse Levine and Amy Feaver for comments on a draft.

2. The topograph

Write $\infty = \frac{1}{0}$, and $\mathbb{Q}^\infty = \mathbb{Q} \cup \{\infty\}$.

Definition. Two points $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}^\infty$ (given in lowest terms) are called $\mathbb{Z}$-*distinct* if $ad - bc = \pm 1$.

The definition is symmetric and doesn't depend on the convention for minus signs in one's definition of 'lowest terms.'

Definition. The *topograph* is the graph whose set of vertices is all triples of pairwise $\mathbb{Z}$-distinct points, with the stipulation that two such triples are connected by an edge whenever they have a pair of elements in common.

We can identify an edge with the pair of $\mathbb{Z}$-distinct elements shared by its vertices[3]. Note that every pair appears exactly once in the topograph since the pair $\frac{a}{b}, \frac{c}{d}$ is part of only two triples:

$$\left\{\frac{a}{b}, \frac{c}{d}, \frac{a+c}{b+d}\right\}, \quad \text{and} \quad \left\{\frac{a}{b}, \frac{c}{d}, \frac{a-c}{b-d}\right\}.$$

The graph can be made planar in such a way that the boundary of each region (an infinite tree of valence 2, i.e. a line), consists of all pairs and triples containing a fixed element. In this way, each region can be labelled with a point of $\mathbb{Q}^\infty$ [6].

3. Möbius transformations

The automorphisms of $\mathbb{Q}^\infty$ are the Möbius transformations,

$$z \mapsto \frac{az+b}{cz+d}, \quad a, b, c, d \in \mathbb{Q}, \quad ad - bc \neq 0,$$

forming a group under composition. This is isomorphic to the matrix group

$$\mathrm{PGL}_2(\mathbb{Q}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Q}, ad - bc \in \mathbb{Q}^* \right\} / \left\{ kI_{2\times 2} : k \in \mathbb{Q}^* \right\},$$

by the map

$$\left(z \mapsto \frac{az+b}{cz+d} \right) \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The subset of matrices having representatives with non-negative integer entries and determinant 1 is closed under multiplication but not inverses, forming the monoid

$$\mathrm{SL}_2(\mathbb{Z}^{\geq 0}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Z}^{\geq 0}, ad - bc = 1 \right\}.$$

The monoid $\mathrm{SL}_2(\mathbb{Z}^{\geq 0})$ is closed under transposition,

$$\gamma \mapsto \gamma^T, \quad \frac{az+b}{cz+d} \mapsto \frac{az+c}{bz+d}.$$

With this notation, the Theorem can now be phrased as follows: replacing the transformation γ with $\gamma(1)$ gives the Calkin-Wilf tree, while replacing it with $1/\gamma^T(1)$ gives the Farey tree[4].

[3] Conway and Fung consider $\mathbb{Q}^\infty$ as $\mathbb{P}^1(\mathbb{Q})$, so that points become primitive vectors of $\mathbb{Z}^2$; each edge corresponds to a basis of $\mathbb{Z}^2$, and vertices give triples called *superbases*.

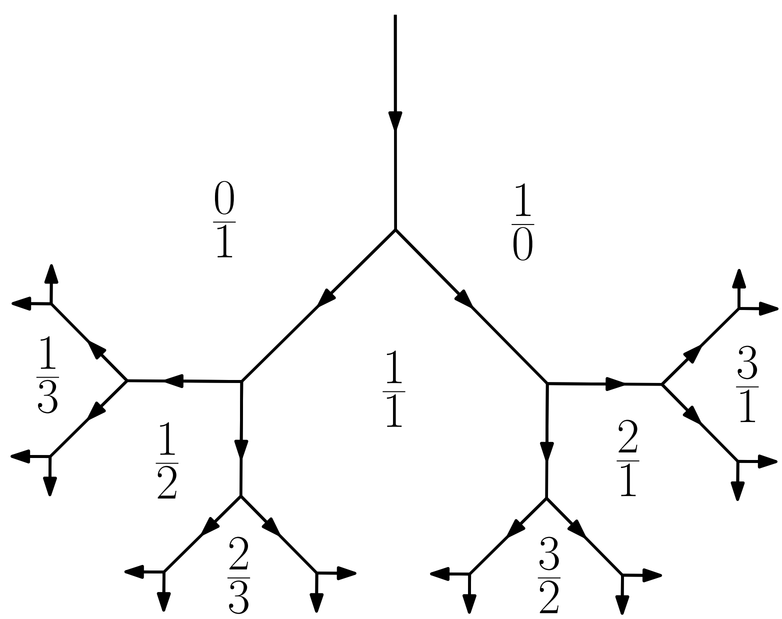
[4] The reciprocal is immaterial, since it would disappear if we wrote the Farey tree right-to-left instead of left-to-right, i.e. reflected in its vertical midline.

4. The Proof

The proof proceeds by labelling the topograph two ways: first, to create the Farey tree, and second, to create the matrix tree. Comparing the two labellings generates the rule given in the Theorem.

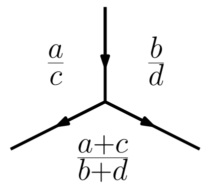
Definition. A *flow* of the topograph (or a portion of it) is an assignment of direction to every edge in such a way that in-degree is exactly one at each vertex.

Once one edge is assigned a direction, the portion of the topograph that is forward of that edge (according to the edge direction), has all its directions determined uniquely by the condition of flow, and forms a rooted binary tree directed away from the root. Choosing the edge $\{0, \infty\}$, and directing it toward the vertex $\{0, \infty, 1\}$, we obtain the following.



Each vertex has one incoming edge; with respect to this direction, there's a left, right and forward region. If we label a vertex with the region bounded by the two outgoing edges (i.e. moving the region labels up to the 'peaks' of their respective regions), we obtain the Farey tree [6]. In particular, all regions are labelled with positive rationals.

By contrast, to such a vertex we may also associate a Möbius transformation $\gamma(z)$ by specifying its values at $\frac{1}{0}$, $\frac{0}{1}$ and $\frac{1}{1}$ are exactly the labels of the regions to the left, right and below the vertex. In other words, if these labels are, respectively, $\frac{a}{c}$, $\frac{b}{d}$, and $\frac{a+c}{b+d}$, then we obtain the transformation $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.



$$\mapsto \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

This gives a tree of matrices. In fact, this is *not* the matrix tree in the introduction, but by applying the map

$$\gamma(z) \mapsto (1/\gamma(z))^T, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} c & a \\ d & b \end{pmatrix},$$

at each vertex, we obtain the matrix tree of the introduction. The verification of this is straightforward: the roots agree and the $\mathbb{Z}$-distinctness condition at each vertex translates into the matrix tree generation rule. This is sufficient to complete the proof.

Afterthoughts

There is another rational-enumerating tree, the Bird Tree [9], which is a levelwise permutation of the Farey tree, but sadly we won't explore it in this note. It is also interesting to observe that if we define a flow of the topograph which directs the boundary of the region $\frac{1}{0}$ from negative to positive regions, we obtain the extended Farey tree of [11].

References

[1] Roland Backhouse and João F. Ferreira. Recounting the rationals: twice! In *Mathematics of program construction*, volume 5133 of *Lecture Notes in Comput. Sci.*, pages 79–91. Springer, Berlin, 2008.

[2] Roland Backhouse and João F. Ferreira. On Euclid's algorithm and elementary number theory. *Sci. Comput. Programming*, 76(3):160–180, 2011.

[3] Bruce Bates, Martin Bunder, and Keith Tognetti. Linking the Calkin-Wilf and Stern-Brocot trees. *European J. Combin.*, 31(7):1637–1661, 2010.

[4] Achille Brocot. *Calcul des rouages par approximation, nouvelle méthode*. A. Brocot, 1862.

[5] Neil Calkin and Herbert S. Wilf. Recounting the rationals. *Amer. Math. Monthly*, 107(4):360–363, 2000.

[6] John H. Conway. *The sensual (quadratic) form*, volume 26 of *Carus Mathematical Monographs*. Mathematical Association of America, Washington, DC, 1997. With the assistance of Francis Y. C. Fung.

[7] S. P. Glasby. Enumerating the rationals from left to right. *Amer. Math. Monthly*, 118(9):830–835, 2011.

[8] Scott B. Guthery. *A motif of mathematics*. Docent Press, Boston, MA, 2011. History and application of the mediant and the Farey sequence.

[9] Ralf Hinze. Functional pearls: the bird tree. *J. Funct. Programming*, 19(5):491–508, 2009.

[10] Robert A. Kucharczyk. Enumerating trees. *http://arxiv.org/abs/1201.1851*, 2012.

[11] J. C. Lagarias and C. P. Tresser. A walk along the branches of the extended Farey tree. *IBM J. Res. Develop.*, 39(3):283–294, 1995.

[12] Jennifer Lansing. *On the Stern sequence and a related sequence*. PhD thesis, University of Illinois at Urbana-Champaign, 2014.
[13] Melvyn B. Nathanson. A tree of linear fractional transformations. *http://arxiv.org/abs/1401.0012*, 2014.
[14] Bruce Reznick. Some binary partition functions. In *Analytic number theory (Allerton Park, IL, 1989)*, volume 85 of *Progr. Math.*, pages 451–477. Birkhäuser Boston, Boston, MA, 1990.
[15] M. A. Stern. Über eine zahlentheoretische funktion. *J. Reine Angew. Math.*, 55:193–220, 1858.

Department of Mathematics, University of Colorado, Campux Box 395, Boulder, Colorado 80309-0395
E-mail address: kstange@math.colorado.edu