FAREY SEQUENCES

ALEX T, KARTHIK B

1. Introduction

Definition 1.1. A Farey sequence F_n is the set of rational numbers $\frac{p}{q}$ with p and q coprime, and $0 \leq p \leq q \leq n$, ordered by size.

Example. $F_1 = \left\{\frac{0}{1}, \frac{1}{1}\right\}$
$F_2 = \left\{\frac{0}{1}, \frac{1}{2}, \frac{1}{1}\right\}$
$F_3 = \left\{\frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1}\right\}$
$F_4 = \left\{\frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1}\right\}$
$F_5 = \left\{\frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{1}{1}\right\}$

Farey sequences come from as far back as 1747. In the 1747 edition of "The Ladies Diary: or, the Woman's Almanac" there was the following question:

It is required to find (by a general theorem) the number of fractions of different values, each less than unity, so that the greatest denominator be less than 100?

This is equivalent to asking the size of the 99th Farey sequence. It turns out this problem is hard. It took the 18th century mathematicians 4 years to solve it.

This paper proves a method of generating the Farey sequence using the mediant, followed by a way to approximate any real number using fractions: Hurwitz's theorem.

Definition 1.2. The mediant of two fractions is given by $\frac{p}{q} \oplus \frac{r}{s} = \frac{p+r}{q+s}$.

Sometimes it will be convenient to have some kind of "weighted mediant" function. [Idi18]

Definition 1.3. The weighted mediant of two fractions is given by $\frac{p}{q}(a \oplus b)\frac{r}{s} = \frac{ap+br}{aq+bs}$. This means that the simple mediant can be expressed as $\frac{p}{q}(1 \oplus 1)\frac{r}{s}$.

2. Generating the Farey Sequence

Let us look at various methods of generating the Farey sequence. It is trivial to show that if you are allowed to take the mediant of any two terms from the previous sequence you will get the entire addition list to the next one - $\frac{p}{q} = \frac{p-1}{p}(1 \oplus 1)\frac{1}{q-p}$. It isn't trivial to show that taking the mediant of two adjacent terms provides the same.

Example. Let's first check by providing an example if this has a chance of working at all. Since $F_5 = F_4 \cup \{\frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}\}$, we just need to show that by taking the mediant of adjacent terms in F_4 you get those. $\frac{0}{1}(1 \oplus 1)\frac{1}{4} = \frac{1}{5}, \frac{1}{3}(1 \oplus 1)\frac{1}{2} = \frac{2}{5}, \frac{1}{2}(1 \oplus 1)\frac{2}{3} = \frac{3}{5}, \frac{3}{4}(1 \oplus 1)\frac{1}{1}$. So, this is at least true for F_5.

Since this holds to the first level of checking-trying an example-let's see if we can prove it. To prove this, we need some lemmata.

Date: December 10, 2018.

Lemma 2.1. *$\frac{p}{q} < \frac{p}{q}(a \oplus b)\frac{r}{s} < \frac{r}{s}$, where $a, b, q, s > 0, \frac{p}{q} < \frac{r}{s}$.*

To prove this lemma, the best method is just algebra.

Proof.

$$\frac{p}{q}(a \oplus b)\frac{r}{s} = \frac{ap+br}{aq+bs}.\frac{ap+br}{aq+bs} - \frac{p}{q} = \frac{apq+brq-aqp-bsp}{aq^2+bsq} = b\frac{rq-sp}{aq^2+bsq}.$$

This means we need $rq - sp > 0$. Since $\frac{p}{q} < \frac{r}{s}$, $ps < rq$, that is true. Thenceforth $\frac{p}{q} < \frac{p}{q}(a \oplus b)\frac{r}{s}$. Now,

$$\frac{r}{s} - \frac{ap+br}{aq+bs} = \frac{aqr+bsr-aps-brs}{aqs+bs^2} = a\frac{qr-ps}{aq+bs^2}.$$

Since $\frac{q}{r} > \frac{p}{s}$, $qr > ps$ and so $\frac{p}{q}(a \oplus b)\frac{r}{s} < \frac{r}{s}$. Q.E.D. ■

We also need the weighted mediant to be simplified.

Lemma 2.2. *If $\gcd(p, q) = \gcd(a, b) = \gcd(r, s) = 1$, and $rq - ps = 1$ then $\gcd(ap+br, aq+bs) = 1$.*

Proof. If

$$\gcd(q * (ap + br), p * (aq + bs)) = 1,$$

then obviously

$$\gcd(ap + br, aq + bs) = 1,$$

since we are just adding potential common divisors. We want

$$\gcd(apq + brq, apq + bps) = 1,$$

but we can be satisfied by at least simplifying it. Let's use the Euclidean algorithm.

$$apq + brq - paq - bps = b(rq - ps) = b.$$

So, we want

$$\gcd(b, aq + bs) = 1.$$

It doesn't matter which one we replace, after all. Using the Euclidean algorithm again, we need $\gcd(apq, b) = 1$. Since $\gcd(a, b) = 1$, we can get rid of the a.

$$\gcd(pq, b) = 1.$$

Henceforth,

$$\gcd(ap + br, aq + bs)|b.$$

Now, that means we can look at this mod b.

$$\gcd(ap + br, b) = \gcd(ap, b) = \gcd(p, b).$$

$$\gcd(aq + bs, b) = \gcd(aq, b) = \gcd(q, b).$$

Since $\gcd(p, q) = 1$, we know that they can't share any divisors, and so each half of the gcd have separate divisors of b and share none, and so $\gcd(ap + br, aq + bs) = 1$. Q.E.D. ■

One more interesting lemma we need is as follows.

Lemma 2.3. *For all fractions $\frac{p}{q} < \frac{x}{y} < \frac{r}{s}, \exists a, b \in \mathbb{Z}^+$ s.t. $\frac{p}{q}(a \oplus b)\frac{r}{s} = \frac{x}{y}$*

Proof. Let $a = ry - sx, b = qx - py$. They are both positive because of the relationship of the fractions. $\frac{p}{q}(a \oplus b)\frac{r}{s} = \frac{ap+br}{aq+bs} = \frac{rpy-sxp+qxr-pyr}{ryq-sxq+qxs-pys} = \frac{qxr-sxp}{ryq-pys} = \frac{x(rq-sp)}{y(rq-sp)} = \frac{x}{y}$. ■

Theorem 2.4. *You can generate F_n by taking all the mediants of adjacent terms of F_{n-1} with denominators n, and adding those on to F_{n-1}.*

To make it provable via induction, we need to add additional information.

Theorem 2.5. *You can generate F_n by taking all the mediants of adjacent terms of F_{n-1} with denominators n, and adding those on to F_{n-1}. Also, adjacent terms of the F_n, $\frac{p}{q}, \frac{r}{s}$ satisfy $rq - ps = 1$.*

Proof. We are going to prove this by induction. The base case is trivial. F_2 can be generated by taking the mediants of F_1, as the mediant of $\frac{0}{1}$ and $\frac{1}{1}$ is $\frac{1}{2}$, the one term added. Also, $1 * 1 - 0 * 2 = 1, 2 * 1 - 1 * 1 = 1$.

The inductive step is harder. Assume this is true for F_{n-1}. We need this to be true for F_n. First, due to 2.3, for all $\frac{x}{n}$ we need to add, we know that there is some a, b such that the terms surrounding it when taken the weighted mediant spit that out. Obviously there isn't already something in between, so what matters is the weighted mediant with the smallest denominator.

Since 2.2 says that it is impossible to simplify any mediant from F_{n-1} due to the second part of our assumption, the smallest denominator mediant is the $(1, 1)$ mediant, or the original. It follows that by taking the mediants, you will get the whole F_n.

We still need to check the second part, i.e. $rq - ps = 1$. Via the inductive hypothesis, we know that if we didn't add a mediant, then this remains true. If we did, then we have the new series of Farey neighbors $\frac{p}{q}, \frac{p+r}{q+s}, \frac{r}{s}$. We know from the inductive hypothesis $rq - ps = 1$, and so we look at $pq + rq - pq + sp = 1$, and $rq + sr - sp - sr = 1$. Q.E.D. ■

3. Hurwitz's Theorem

Surprisingly, Farey sequences can be used in approximating irrational numbers.
Hurwitz's Theorem [ADPW18] states that:

Definition 3.1. Given any irrational number ϵ, there exists infinitely many rational numbers h, k such that

$$| \epsilon - \frac{h}{k} | < \frac{1}{\sqrt{5}k^2}.$$

has infinitely many rational solutions $\frac{p}{q}$.

Proof. Proof by contradiction: We start off with two consecutive terms in Farey Sequence F_n $\frac{a}{b}, \frac{c}{d}$ with ϵ in between the two, and their mediant

$$\frac{a+c}{b+d} = \frac{e}{f}$$

Assumption: Assume that the statement of Hurwitz's theorem is false and that

$$\epsilon - \frac{a}{b} \geq \frac{1}{\sqrt{5}b^2}$$

$$\epsilon - \frac{c}{d} \geq \frac{1}{\sqrt{5}d^2}$$

$$\epsilon - \frac{e}{f} \geq \frac{1}{\sqrt{5}f^2}.$$

For the purposes of this proof we rewrite the second inequality into

$$\frac{c}{d} - \epsilon \geq \frac{1}{\sqrt{5}d^2}.$$

Adding the first and third equation to the second separately results in getting two equations

$$\frac{c}{d} - \frac{a}{b} \geq \frac{1}{\sqrt{5}}\left(\frac{1}{d^2} + \frac{1}{b^2}\right)$$
$$\frac{c}{d} - \frac{e}{f} \geq \frac{1}{\sqrt{5}}\left(\frac{1}{d^2} + \frac{1}{f^2}\right).$$

We know that $\frac{c}{d} - \frac{a}{b} = \frac{bc-ad}{bc} = \frac{1}{bc}$, and $\frac{c}{d} - \frac{e}{f} = \frac{cf-de}{df} = \frac{1}{df}$. So, we can simplify those two inequalities into

$$\frac{1}{bd} \geq \frac{1}{\sqrt{5}}\left(\frac{1}{d^2} + \frac{1}{b^2}\right)$$
$$\frac{1}{df} \geq \frac{1}{\sqrt{5}}\left(\frac{1}{d^2} + \frac{1}{f^2}\right).$$

Getting rid of the fractions gets

$$bd\sqrt{5} \geq d^2 + b^2, df\sqrt{5} \geq d^2 + f^2$$

Add these two together, and using the fact that $f = b + d$, we get

$$(b+f)d\sqrt{5} = (2b+d)d\sqrt{5} \geq 2d^2 + b^2 + f^2 = 3d^2 + 2b^2 + 2bd.$$

When we subtract off the left side to make it zero, we have a square left.

$$0 \geq 3d^2 + 2b^2 + 2bd - (2b+d)d\sqrt{5} = \frac{1}{2}((\sqrt{5}-1)d - 2b)^2.$$

Henceforth $\frac{1}{2}((\sqrt{5}-1)d - 2b)^2 = 0$, or $(\sqrt{5}-1)d - 2b = 0$. Adding $2b$ and dividing by d gets us $\sqrt{5} - 1 = \frac{2b}{d}$, which means that $\frac{2b}{d}$ is irrational, which is false, as b and d are integers. Contradiction. Q.E.D. ■

References

[ADPW18] Jonathan Ainsworth, Michael Dawson, John Pianta, and James. Warwick. The farey sequence. March 15, 2012. University of Edinburgh, www.maths.ed.ac.uk/ v1ranick/fareyproject.pdf. Accessed on 8 December, 2018.

[Idi18] Sum Idiot. Farey sequences. November 3, 2009. https://sumidiot.wordpress.com/2009/11/03/farey-sequenc es/. Accessed on 8 December, 2018.