Math 4527 (Number Theory 2)

Lecture #3 of 38 $\sim$ January 25, 2021

Farey Sequences

- The Frobenius Coin Problem (part 2)
- The Farey Sequences
- Rational Approximation

This material represents §6.2.1 from the course notes.

More Frobenius Coin Problem, I

The problem of describing the largest integer that cannot be written as a nonnegative linear combination of two integers (also called the Frobenius coin problem) was first solved by Sylvester:

Theorem (Sylvester)

If a and b are relatively prime integers, then there are exactly $\frac{1}{2}(a-1)(b-1)$ integers that cannot be written in the form $ax + by$ with $x, y \geq 0$, and the largest such integer is $ab - a - b$.

I proved this theorem last time. We'll do a few quick applications today.

More Frobenius Coin Problem, II

Example: There are postage stamps worth 5 cents and stamps worth 13 cents. What is the largest non-attainable amount of postage, and how many non-attainable amounts are there?

More Frobenius Coin Problem, II

Example: There are postage stamps worth 5 cents and stamps worth 13 cents. What is the largest non-attainable amount of postage, and how many non-attainable amounts are there?

- By Sylvester's theorem with $a = 5$ and $b = 13$, the largest non-representable integer is $5 \cdot 13 - 13 - 5 = 47$.
- In total, there are $\frac{1}{2} \cdot 4 \cdot 12 = 24$ unattainable totals.

More Frobenius Coin Problem, III

We could of course generalize this problem, to ask: for given integers $a_1, a_2, \dots, a_k$, what is the largest integer n that cannot be written as a nonnegative integer linear combination of the a_i?

- It turns out that there is no known general formula when $k > 2$ (though the result is fairly effectively computable for $k = 3$).
- For a fixed number of denominations k, there does exist a polynomial-time algorithm (polynomial in $\log a_k$, specifically) for computing this maximum integer n, but it is not appreciably faster than merely attempting to list the possibilities!
- For a variable number of denominations k, it is known that computing n is NP-hard.

For small values, we can use Sylvester's theorem and some case analysis to solve the more general problem.

More Frobenius Coin Problem, IV

Example: Find the largest amount of postage that cannot be given using some combination of 6-cent, 11-cent, and 14-cent stamps.

More Frobenius Coin Problem, IV

Example: Find the largest amount of postage that cannot be given using some combination of 6-cent, 11-cent, and 14-cent stamps.

- By Sylvester's theorem with $a = 3$ and $b = 7$, we know that $3 \cdot 7 - 3 - 7 = 11$ is the largest amount that cannot be made with 3 and 7.
- Thus, 6 and 14 can make any even total that exceeds 22.
- Adding one 11-cent stamp if needed, we can make any total greater than or equal to 34. We can also clearly make $33 = 3 \cdot 11$, $31 = 6 + 11 + 14$, $29 = 3 \cdot 6 + 11$.
- But 27 is not possible (we would need one 11-cent stamp, but we cannot make 16 from 6 and 14). So the largest amount we cannot make is 27 cents.

Rational Approximation, I

We will now study some problems related to rational approximation of real numbers by rational numbers.

- Although it may not seem so clear at the moment, we will in fact be able to use some of these results to solve Diophantine equations.
- When describing real numbers, for convenience we often want to give a nearby rational number that is a good approximation.

Rational Approximation, II

Indeed, this idea is implicitly embedded in the notion of the decimal expansion of a real number.

- For example, writing $e = 2.7182818284590\ldots$ formally means that e is the limit of the sequence 2, 2.7, 2.71, 2.718, 2.7182, $\ldots$.
- Thus, truncating this sequence after some finite number of steps will provide a good approximation of e.
- More specifically, in the case of the decimal expansion to n digits, the approximation is accurate to within an error of 10^{-n}.
- Decimal numbers are all well and good, but we can often get better approximations using arbitrary rational numbers, rather than just ones whose denominators are powers of 10.

Farey Sequences, I

If we are seeking to approximate a real number α, one thing we might first look at is the set of rational numbers of small denominator.

- Since we want to understand distances between nearby numbers, we should arrange the rationals in increasing order, and then identify where our real number α lands between the nearest pair.

This resulting sequence of rationals is known as a Farey sequence:

Definition

The Farey sequence of level n is the set of rational numbers between 0 and 1 whose denominators (in lowest terms) are $\leq n$, arranged in increasing order.

Farey Sequences, II

Here are the first few Farey sequences:

- Level 1: $\frac{0}{1}, \frac{1}{1}.$
- Level 2: $\frac{0}{1}, \frac{1}{2}, \frac{1}{1}.$
- Level 3: $\frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1}.$
- Level 4: $\frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1}.$
- Level 5: $\frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{1}{1}.$

Farey Sequences, III

To get the $(n+1)$st Farey sequence from the nth one, we just need to insert the fractions with denominator $n+1$ properly:

- Level 6:
  $$\frac{0}{1}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{1}{1}.$$
- Level 7:
  $$\frac{0}{1}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{2}{7}, \frac{1}{3}, \frac{2}{5}, \frac{3}{7}, \frac{1}{2}, \frac{4}{7}, \frac{3}{5}, \frac{2}{3}, \frac{5}{7}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{1}{1}.$$
- Level 8:

$$\frac{0}{1}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{2}{7}, \frac{1}{3}, \frac{3}{8}, \frac{2}{5}, \frac{3}{7}, \frac{1}{2}, \frac{4}{7}, \frac{3}{5}, \frac{5}{8}, \frac{2}{3}, \frac{5}{7}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{7}{8}, \frac{1}{1}.$$

Audience participation time: find some patterns.

Farey Sequences, IV

A few fairly obvious patterns:

- Here's 4 sets of newly-inserted terms:
  $\frac{1}{2}, \frac{3}{5}, \frac{2}{3} \qquad \frac{0}{1}, \frac{1}{7}, \frac{1}{6} \qquad \frac{2}{5}, \frac{3}{7}, \frac{1}{2} \qquad \frac{1}{3}, \frac{3}{8}, \frac{2}{5}.$
- It appears that the first term that appears between two consecutive terms a/b and c/d is $(a+c)/(b+d)$. This quantity is called the <u>mediant</u> (or the "baseball average").
- Here's 3 pairs of differences between consecutive terms:
  $\frac{2}{5} - \frac{1}{3} = \frac{1}{15} \qquad \frac{2}{7} - \frac{1}{5} = \frac{1}{35} \qquad \frac{6}{7} - \frac{5}{6} = \frac{1}{42}.$
- Note that the difference between the terms a/b and c/d is always $1/(bd)$. Equivalently, the value $bc - ad$ for consecutive terms a/b and c/d is always equal to 1.

Farey Sequences, V

Proposition (Properties of Farey Sequences)

Let n be a positive integer.

1. *If a/b and c/d are consecutive terms in the Farey sequence of level n, then $bc - ad = 1$.*
2. *If a/b, e/f, and c/d are three consecutive terms in a Farey sequence, then $e/f = (a+c)/(b+d)$.*
3. *If $0 \le a/b, c/d \le 1$ with $bc - ad = 1$, then a/b and c/d are consecutive in the Farey sequence of level $\max(b, d)$. The first term that appears between them in any later sequence is $(a+c)/(b+d)$ in the Farey sequence of level $b+d$.*
4. *a/b and c/d are consecutive terms in the Farey sequence of level n if and only if $bc - ad = 1$ and $b + d > n$.*
5. *If a/b and e/f are consecutive terms in the Farey sequence of level n, the term immediately following e/f is c/d, where $c = \left\lfloor \frac{n+b}{f} \right\rfloor e - a$ and $d = \left\lfloor \frac{n+b}{f} \right\rfloor f - b$.*

Farey Sequences, VI

In order to prove the first result, we will need the following surprisingly useful fact from elementary geometry:

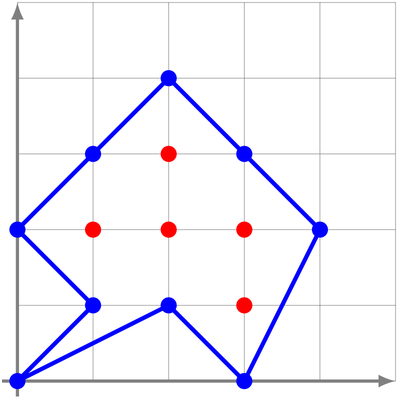
Theorem (Pick's Theorem)

If R is a polygon in the plane whose vertices are all lattice points, then the area of R is given by the formula $A = I + \frac{1}{2}B - 1$, where I is the number of lattice points in the interior of R and B is the number of lattice points on the boundary of R.

A boundary point is a point on one of the sides of R, while an interior point is a point not on one of the sides of R.

Farey Sequences, VII

Pick's Theorem is easiest to see with an example: this polygon has 9 boundary points and 5 interior points, and by drawing triangles around it, one can verify its area is $\frac{17}{2} = 5 + \frac{9}{2} - 1$:

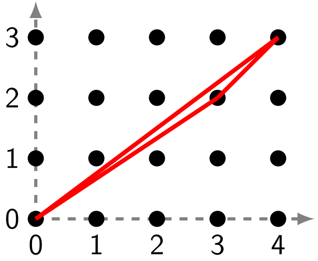


Farey Sequences, VIII

1. If a/b and c/d are consecutive terms in the Farey sequence of level n, then $bc - ad = 1$.

Proof:

- Suppose a/b and c/d are consecutive terms in the Farey sequence of level n.
- In the plane, draw the triangle whose vertices are (0,0), (b, a), and (d, c):



- By Pick's Theorem, the area of this lattice-point triangle is $\frac{1}{2}B + I - 1$.
- We claim $B = 3$ and $I = 0$ (as is clear in the picture).

Farey Sequences, VIII

1. If a/b and c/d are consecutive terms in the Farey sequence of level n, then $bc - ad = 1$.

<u>Proof</u> (continued):

- To show $B = 3$ and $I = 0$, first suppose there were a lattice point (x, y) in the interior, where (necessarily) $y \leq \max(b, d)$. Then the slope of the line joining $(0, 0)$ to (x, y) would be between a/b and c/d: but then y/x would be between a/b and c/d in the Farey sequence, impossible.
- Now consider a non-vertex boundary point. It cannot lie on the side joining (0,0) and (b, a), since a and b are relatively prime. Similarly, it cannot lie on the side joining (0,0) and (d, c). If it were on the side joining (b, a) and (d, c), then by the same argument given above, there would be a term between a/b and c/d in the Farey sequence.

Farey Sequences, IX

1. If a/b and c/d are consecutive terms in the Farey sequence of level n, then $bc - ad = 1$.

<u>Proof</u> (finally):

- Thus, $B = 3$ and $I = 0$, so the triangle has area $\frac{1}{2}$.
- By basic geometry we can see that the area of the triangle with vertices $(0, 0)$, (b, a), and (d, c) is $\frac{1}{2}|bc - ad|$.
- Some approaches for this: either enclose this triangle with larger right triangles, or note that the area of the triangle is half of the magnitude of the cross product $\langle b, a, 0\rangle \times \langle d, c, 0\rangle = \langle 0, 0, bc - ad\rangle$.
- So, since $bc > ad$, setting $\frac{1}{2}|bc - ad| = \frac{1}{2}$ immediately gives $bc - ad = 1$.

Farey Sequences, X

2. If a/b, e/f, and c/d are three consecutive terms in a Farey sequence, then $e/f = (a+c)/(b+d)$.

Proof:

- By (1), since a/b and e/f are consecutive we have $be - af = 1$, and by (2) since e/f and c/d are consecutive we have $cf - de = 1$.
- This is a system of two linear equations in the two variables e and f, so solving it yields $e = (a+c)/(bc-ad)$ and $f = (b+d)/(bc-ad)$.
- Thus, $\dfrac{e}{f} = \dfrac{a+c}{b+d}$, as claimed.
- One can check directly that $\dfrac{a+c}{b+d}$ appears between $\dfrac{a}{b}$ and $\dfrac{c}{d}$ in the Farey sequence of level $b+d$, since $\dfrac{a}{b} < \dfrac{a+c}{b+d} < \dfrac{c}{d}$.

Farey Sequences, XI

3. If $0 \leq a/b,\ c/d \leq 1$ with $bc - ad = 1$, then a/b and c/d are consecutive in the Farey sequence of level $\max(b, d)$. The first term that appears between them in any later sequence is $(a+c)/(b+d)$ in the Farey sequence of level $b+d$.

Proof:

- First suppose $\dfrac{e}{f}$ is the term immediately following $\dfrac{a}{b}$ in the Farey sequence of level $\max(b, d)$. Then $be - af = 1$ by (1).
- Subtracting $bc - ad = 1$ yields $b(c-e) - a(d-f) = 0$, so $b(c-e) = a(d-f)$. Since a and b are relatively prime, we conclude that b divides $d - f$. Since $f \leq \max(b, d) < b + d$, the only possibility is that $f = d$, and then $e = c$.
- Alternatively, we could have observed that both (e, f) and (c, d) are solutions to the linear Diophantine equation $bx - ay = 1$, and used the structure of the solutions to win.

Farey Sequences, XII

3. If $0 \leq a/b,\ c/d \leq 1$ with $bc - ad = 1$, then a/b and c/d are consecutive in the Farey sequence of level $\max(b, d)$. The first term that appears between them in any later sequence is $(a+c)/(b+d)$ in the Farey sequence of level $b + d$.

Proof (continued):

- For the second statement, we just showed that a/b and c/d are consecutive in the Farey sequence of level $\max(b, d)$.
- Now increase the level of the sequence in increments of 1.
- If e/f is the first term to appear between a/b and c/d, then by (2), it would necessarily be the case that $e = (a+c)/(bc-ad) = a+c$ and $f = (b+d)/(bc-ad) = b+d$.
- So in fact, this is the first term that appears between them.

Farey Sequences, XIII

4. a/b and c/d are consecutive terms in the Farey sequence of level n if and only if $bc - ad = 1$ and $b + d > n$.

<u>Proof</u>:

- We must have $bc - ad = 1$ by (1).
- Also, if $b + d \leq n$, then $\dfrac{a+c}{b+d}$ is a term between a/b and c/d as noted in (2).
- Thus, we must have $bc - ad = 1$ and $b + d > n$.
- But if both conditions hold, then (3) immediately implies that there are no terms between a/b and c/d in the Farey sequence of level n, so they are in fact consecutive.

Farey Sequences, XIV

5. If a/b and e/f are consecutive terms in the Farey sequence of level n, the term immediately following e/f is c/d, where $c = \left\lfloor \frac{n+b}{f} \right\rfloor e - a$ and $d = \left\lfloor \frac{n+b}{f} \right\rfloor f - b$.

Proof:

- By the mediant property (2), we know that $\frac{e}{f} = \frac{a+c}{b+d}$.
- Thus, there must exist some integer k such that $a + c = ke$ and $b + d = kf$, so that $c = ke - a$ and $d = kf - b$.
- Since the closest term to e/f will have k as large as possible, and since $d \leq n$, the largest possible value of k is $\left\lfloor \frac{n+b}{f} \right\rfloor$.

Farey Sequences, XV

Using the above results, we can construct the portion of any Farey sequence around any desired rational number, without needing to compute all of the terms in the sequence.

- To find the next term after a/b we can solve $bx - ay = 1$, and then we can use use the two-term recursion to extend the sequence.
- We can also fill in the terms in between any two given terms by taking mediants, or (if desired) by using the above procedure to generate the terms after a/b.

Farey Sequences, XVI

Example: Verify that $7/19$ and $10/27$ are consecutive in the Farey sequence of level 35, and find the next term that appears between them in a higher sequence.

Farey Sequences, XVI

Example: Verify that $7/19$ and $10/27$ are consecutive in the Farey sequence of level 35, and find the next term that appears between them in a higher sequence.

- By the above results, since $10 \cdot 19 - 7 \cdot 27 = 190 - 189 = 1$, the terms are consecutive in the Farey sequence of level 27.
- The next term will be the mediant $17/46$, and so they are still consecutive in the Farey sequence of level 35.

Farey Sequences, XVII

<u>Example</u>: Verify that 10/17 and 13/22 are consecutive in the Farey sequence of level 25, and find the next three terms after them.

- Recurrence: $c = \left\lfloor \dfrac{n+b}{f} \right\rfloor e - a$ and $d = \left\lfloor \dfrac{n+b}{f} \right\rfloor f - b$.

Farey Sequences, XVII

<u>Example</u>: Verify that $10/17$ and $13/22$ are consecutive in the Farey sequence of level 25, and find the next three terms after them.

- Recurrence: $c = \left\lfloor \dfrac{n+b}{f} \right\rfloor e - a$ and $d = \left\lfloor \dfrac{n+b}{f} \right\rfloor f - b$.
- By the above results, since $13 \cdot 17 - 10 \cdot 22 = 221 - 220 = 1$, the terms are consecutive in the Farey sequence of level 22.
- The next term will be the mediant $23/39$, and so they are still consecutive in the Farey sequence of level 25.
- For the next terms we use the recursion: after a/b, e/f, the next term is $c = \left\lfloor \dfrac{n+b}{f} \right\rfloor e - a$ and $d = \left\lfloor \dfrac{n+b}{f} \right\rfloor f - b$.
- Plugging in yields the next terms $3/5, 14/23, 11/18$.

Farey Sequences, XVIII

Example: Find the next term after $11/202$ in the Farey sequence of level 500.

Farey Sequences, XVIII

<u>Example</u>: Find the next term after $11/202$ in the Farey sequence of level 500.

- By the above results, if $11/202$ and c/d are consecutive terms, then $202c - 11d = 1$.
- We can solve this linear Diophantine equation using the method from last class: reducing mod 11 yields $4c \equiv 1 \pmod{11}$ so $c = 3$ is a solution, yielding $d = 55$.
- So the solutions are $(c, d) = (3 + 11k, 55 + 202k)$ for $k \in \mathbb{Z}$.
- The larger the value of k is, the smaller the value of $\dfrac{c}{d} - \dfrac{11}{202} = \dfrac{1}{202d}$ will be.
- The largest possible value for k with $55 + 202k \leq 500$ is $k = 2$, so the next term is $25/457$.

Farey Sequences, XIX

Example: Find all terms between $7/33$ and $14/65$ in the Farey sequence of level 100.

Farey Sequences, XIX

Example: Find all terms between $7/33$ and $14/65$ in the Farey sequence of level 100.

- These terms are not consecutive anywhere because $14 \cdot 33 - 7 \cdot 65 = 7$, not 1.
- We start by finding terms between them: the mediant of these two terms is $21/98 = 3/14$.
- Now $7/33$ and $3/14$ are consecutive in the Farey sequence of level 33, since $33 \cdot 3 - 7 \cdot 14 = 1$.
- Also, $3/14$ and $14/65$ are consecutive in the Farey sequence of level 65, since $14 \cdot 14 - 3 \cdot 65 = 1$.
- At this point, we can find all of the remaining terms by computing mediants (we can stop when the sum of two consecutive denominators exceeds 100). We get $\frac{7}{33}, \frac{17}{80}, \frac{10}{47}, \frac{13}{61}, \frac{16}{75}, \frac{19}{89}, \frac{3}{14}, \frac{20}{93}, \frac{17}{79}, \frac{14}{65}$.

More Farey Sequences, III

We can use the Farey sequences to do rational approximation.

Proposition (Rational Approximation via Farey)

Let n be a positive integer and α be a real number. Then the following hold:

1. *There exists a rational number p/q such that $0 < q \leq n$ and* $\left|\alpha - \dfrac{p}{q}\right| \leq \dfrac{1}{q(n+1)}$.
2. *If α is irrational, then there are infinitely many distinct rational numbers p/q such that $|\alpha - p/q| < 1/q^2$.*
3. *If α is irrational, then there are infinitely many pairs of positive integers (m, n) such that $|m\alpha - n| < 1/m$.*

More Farey Sequences, III

1. There exists a rational number p/q such that $0 < q \leq n$ and $\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q(n+1)}$.

Proof:

- By replacing α with $\alpha - \lfloor \alpha \rfloor$ as necessary, we can assume $\alpha \in [0, 1]$.
- Now consider the Farey sequence of level n, and let $\frac{a}{b}$ and $\frac{c}{d}$ be two consecutive terms such that $\frac{a}{b} \leq \alpha \leq \frac{c}{d}$.
- By our earlier results, we know that $bc - ad = 1$ and $b + d \geq n + 1$.

More Farey Sequences, IV

1. There exists a rational number p/q such that $0 < q \leq n$ and $\left|\alpha - \dfrac{p}{q}\right| \leq \dfrac{1}{q(n+1)}$.

<u>Proof</u> (continued):

- The number α either lies in $\left[\dfrac{a}{b}, \dfrac{a+c}{b+d}\right]$ or in $\left[\dfrac{a+c}{b+d}, \dfrac{c}{d}\right]$.
- In the first case, $\left|\alpha - \dfrac{a}{b}\right| \leq \left|\dfrac{a}{b} - \dfrac{a+c}{b+d}\right| = \dfrac{|ad-bc|}{b(b+d)} \leq \dfrac{1}{b(n+1)}$.
- In the second case, $\left|\alpha - \dfrac{c}{d}\right| \leq \left|\dfrac{c}{d} - \dfrac{a+c}{b+d}\right| = \dfrac{|ad-bc|}{d(b+d)} \leq \dfrac{1}{d(n+1)}$.
- Hence, in either case, we obtain a rational number p/q such that $\left|\alpha - \dfrac{p}{q}\right| \leq \dfrac{1}{q(n+1)}$.

More Farey Sequences, IV

2. If α is irrational, then there are infinitely many distinct rational numbers p/q such that $|\alpha - p/q| < 1/q^2$.

Proof:

- Apply (1) to the Farey sequence of level n for each n: this yields a collection of rational numbers p_n/q_n such that $\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n(n+1)} < \frac{1}{q_n^2}$, and with $q_n \leq n$.
- Since α is irrational, none of these differences can be zero.
- Thus, there must be infinitely many different terms p_n/q_n, since the distances $\left|\alpha - \frac{p_n}{q_n}\right|$ become arbitrarily small, but remain nonzero.

More Farey Sequences, V

3. If α is irrational, then there are infinitely many pairs of positive integers (m, n) such that $|m\alpha - n| < 1/m$.

Proof:

- Clear denominators in (2).

This result was first proven by Dirichlet and is sometimes known as Dirichlet's Diophantine approximation theorem.

More Farey Sequences, VI

We can illustrate these results with a typical irrational $\alpha = \sqrt{2} \approx 1.4142136\ldots$ for various n.

- For example, with $n = 5$ (in the Farey sequence of level 5) the two entries surrounding $\sqrt{2} - 1$ are $2/5$ and $1/2$.
- We can see that $\left|\sqrt{2} - 7/5\right| \approx 0.0142 < \frac{1}{5 \cdot 5}$, so $7/5$ has the desired property in (1) of the proposition.
- In fact, $3/2$ also has the desired property, since $\left|\sqrt{2} - 3/2\right| \approx 0.0858 < \frac{1}{5 \cdot 2}$.
- Taking an increasing sequence of values of n up to $n = 100$ then yields various $\dfrac{p}{q}$ with $\left|\sqrt{2} - p/q\right| < 1/q^2$ per (2): specifically, we obtain the sequence 1, 2, $3/2$, $4/3$, $7/5$, $10/7$, $17/12$, $24/17$, $41/29$, $58/41$, $99/70$, $140/99$, ...

Summary

We finished our discussion of the Frobenius coin problem.

We introduced the Farey sequences and established some of their properties.

Next lecture: Rational approximation and continued fractions.