PROMYS Problem Set Solutions

Adithya Prabha

January 2023

1 Problem 1

Problem One

Consider the sequence

$$a_1 = 2^1 - 3 = -1$$

$$a_2 = 2^2 - 3 = 1$$

$$a_3 = 2^3 - 3 = 5$$

$$a_4 = 2^4 - 3 = 13$$

...

$$a_n = 2^n - 3$$

...

defined for positive integers n. Which elements of this sequence are divisible by 5? What about 13? Are any elements of this sequence divisible by $65 = 513$? Why or why not?

Note: *This solution was largely motivated by Fermat's Little Theorem to find the* 1*-residue. However, I used a more basic approach to solve Part 1, and used a Fermat's based solve for Part 2.*

Part 1: We start by solving the first question. $5|a_n$ when $5|(2^n - 3)$ which is also $2^n \equiv 3 \pmod 5$. We can say that if $2^n \equiv 3 \pmod 5$, 2^n can have a units digit of either 3 or 8. However, 2^n is always even and its units digit is always in the set $\{2, 4, 8, 6\}$. Thus, 2^n has a units digit of 8 which happens for $n \equiv 3 \pmod 4$.

Part 2: The second question asks me to find all n such that $13|(2^n - 3)$. Testing the first 13, $2^n \pmod{13}$,

$2^1 \equiv 2 \pmod{13}, 2^2 \equiv 4 \pmod{13}, 2^3 \equiv 8 \pmod{13}, 2^4 \equiv 3 \pmod{13}, 2^5 \equiv 6 \pmod{13}, 2^6 \equiv 12 \pmod{13}, 2^7 \equiv 11 \pmod{13}, 2^8 \equiv 9 \pmod{13} \ldots, 2^9 \equiv 5 \pmod{13}, 2^{10} \equiv 10 \pmod{13}, 2^{11} \equiv 7 \pmod{13}, 2^{12} \equiv 1 \pmod{13}, 2^{13} \equiv 2 \pmod{13}$

The above was to demonstrate how the residue of $2^n \pmod{13}$ repeats in a cycle of 13. We essentially need to find all n such that $2^n \equiv 3 \pmod{13}$.

Now, using Fermat's Little Theorem (or just by inspection) we see that $2^{12} \equiv 1 \pmod{13}$. This can be seen by inspection as well because $2^6 \equiv 12 \equiv -1 \pmod{13}$, squaring both sides we reach $2^{12} \equiv 1 \pmod{13}$.

It is known from the problem statement itself that $2^4 \equiv 3 \pmod{13}$, and we also know that $2^{12} \equiv 1 \pmod{13}$. 1 to any power will always remain one so $2^{12x} \equiv 1^n \equiv 1 \pmod{13}$. Thus, $2^{4+12x} \equiv 3 \pmod{13}$ for all positive integers x. This means that $n \equiv 4 \pmod{12}$.

Part 3: Chinese Remainder Theorem Because we know $n \equiv 4 \pmod{12}$ and $n \equiv 3 \pmod{4}$, we can use the Chinese Remainder Theorem(CRT) to solve the system of linear congruence.

$$S = \begin{cases} n \equiv 3 \pmod{4} \\ n \equiv 4 \pmod{12} \end{cases} \tag{1}$$

It's easy to see that using the second congruence, $n = 12k+4$, using this in the first congruence, $12k + 4 \equiv 3 \pmod{4}$. Because $4|12k + 4$ for all integer k, $12k + 4 \equiv 0 \pmod{4}$, which results in $0 \equiv 3 \pmod{4}$. This is, however, not possible. So, we have our answer: there is no element in this sequence divisible by $65 = 5 \cdot 13$.

Why Isn't There a Solution Set?: This is because there are two different cycles: a cycle every 4 and a cycle every 12. These cycles never coincide because of the starting number. In other words, the cycle of 4 is odd for all n because 3 (which is odd) will always be odd if added to another even. On the other hand, the cycle of 12 is always even. If the starting number was to change for the system of linear congruence (or what n is congruent to for each of the modulo) such that their parity is the same, they would intersect for some $\mod 65$ or be in the form $a + 65 \cdot b$.

Observation: The modulo for 5 and 13 are both cycles of 4 and 12. What is interesting about this is that for both it is $n - 1$. When thinking about this, I realized that it was because it both 5 and 13 are prime. Now using either, Fermat's Little Theorem or Euler's Theorem. Say there is a prime p,

Fermat's Theorem is,

$$a^{p-1} \equiv 1 \pmod{p} \text{ or } a^p \equiv a \pmod{p}$$

Euler's Theorem is,

$$a^{\varphi(p)} \equiv 1 \pmod{p}$$

For any prime p, $\varphi(p) = p - 1$. Using the latter, $a^{p-1} \equiv 1 \pmod{p}$, and similarly, using Fermat's, $a^{p-1} \equiv 1 \pmod{p}$. We get the same thing. This influences the length of the cycles for specifically primes p: the length of the cycle is $p - 1$. This is because when using the two theorems, we cycle back to a 1 at $p - 1$ and at the same time $a^0 \equiv 1 \pmod{p}$ so the length is from 0 to $p - 2$ which has length $p - 1$. Another way to get the length of the cycle is by noticing by the second form of Fermat's that it cycles from 1 to $p-1$ which also has length $p-1$.

What if the number divisible is composite?: If the number is even it can never happen because of parity. If the number is composite and odd, it depends on the number. For example if it is 9, it will never be congruent to 3 for the modulo. However, numbers like $125, 253, etc.$ will work. I couldn't get any considerable conclusion for this case.

2 Problem 2

Problem Two

To get the echo of a positive integer, we write it twice in a row without a space. For example, the echo of 2023 is 20232023. Is there a positive integer whose echo is a perfect square? If so, how many such positive integers can you find? If not, explain why not.

We start by looking at the number given in the problem statement. $20232023 = 2023 * 10001 = 2023 \cdot (10^4 + 1)$. A generalization we can make,

Generalization: For a number N with n digits, $\text{echo}(N) = N \cdot (10^n + 1)$.
Now there are three cases that I originally thought of to how the echo can be a perfect square.

Case 1: N and $10^n + 1$ are both perfect squares.
This cannot be true because $10^n + 1$ cannot be a perfect square. This can be proved by taking $\pmod 3$ of the expression, $10^n + 1 \equiv 2 \pmod 3$. However, it is known that a perfect square cannot be in the form: $2 \pmod 3$. Consider the modulo-3 residues,

If $N \equiv 0 \pmod 3$, $N^2 \equiv 0 \pmod 3$.
If $N^2 \equiv 1 \pmod 3$, $N^2 \equiv 1 \pmod 3$.
If $N^2 \equiv 2 \pmod 3$, $N^2 \equiv 1 \pmod 3$.

All numbers are either $0, 1, 2$ modulo-3, and thus, N cannot be a perfect square. **Case 2:** $N = 10^n + 1$.
This can also never be possible because $10^n + 1$ always has $n + 1$ digits and we defined N to have only n digits, and thus, this is never possible.
Case 3: $N \cdot (10^n + 1)$ is a perfect square but neither of N and $10^n + 1$ are. Because we have disproved case 2, we can conclude that if there is an echo number that is a perfect square, the only way that could be possible is if N contains a perfect square. The main idea is that if $10^n + 1 = a^2 \cdot b$, and $N = b \cdot c^2$ where $a \neq 1$ as this leads back to case 2.
Playing around with $10^n + 1$, we calculate the prime factorization of the first few:

$$10^1 + 1 = 1 \cdot 11$$
$$10^2 + 1 = 1 \cdot 101$$
$$10^3 + 1 = 11 \cdot 91$$
$$10^4 + 1 = 73 \cdot 137$$
$$10^5 + 1 = 11 \cdot 9091$$
$$10^6 + 1 = 101 \cdot 9901$$
$$10^7 + 1 = 11 \cdot 909091$$

One obvious observation you can get from this is that 11 is a factor for all $10^n + 1$ for all odd n, and this can be "proved" by noting that any odd number is in the form $2k + 1$. Plugging this in as n, $10^{2k+1} + 1 = 10 \cdot 10^{2k} + 1 = 10 \cdot (100)^k + 1 = 10 \cdot (99 + 1)^k + 1$ Noting from the binomial theorem, $(99 + 1)^k$ is divisible by 11 or 99 for all terms in the expansion other than $1^k = 1$. Plugging this back into the starting $10 \cdot (100)^k + 1$, we see that $10 \cdot 1 + 1 \equiv 0 \pmod{11}$. So we have shown that 11 is a factor of all $10^n + 1$ where n is odd (which is equivalent to n being in the form $2k + 1$). At the same time, a much easier way is by using the divisibility rule for 11 we have seen if the difference between the sum of the alternating terms is divisible by 11, the number is divisible by 11. If n is odd, it simply boils down to $1 - 1 = 0$ which is divisible by 11.

Another less obvious observation is that for all odd n, $10^n + 1$ is in the form $11 \cdot 9090\ldots91$. A generalization we can make for this is that, $10^n + 1$ equals to $11 \cdot x$ where x can be constructed from repeating 90 a total number of $\frac{n-3}{2}$ and then append a 91 at the end. The reason $\frac{n-3}{2}$ works is because n is always odd and thus the fraction is always even; however, it works for all odd numbers greater than one.

Sidenote: The reason I chose this way of pairing up the digits was because it was easy to explain and pairing them up differently makes it really complicated.

Going back to the start of case 3, we need to find a number in the form $10^n + 1$ that is not square-free. Our best bet is to go with 11 because half of our numbers are already divisible by 11 and we simply need to find a number such that $11 | 9090\ldots91$. We can again use the divisibility rule for 11, The number of digits in $9090\ldots91$ is equal to $n-1$ and thus $\frac{n-1}{2}$. It is easy to see that the sum of the other set of alternating terms is 1, so we need to find some n such that $9 \cdot \frac{n-1}{2} - 1 \equiv 0 \pmod{11}$.

$$9 \cdot (n-1) \cdot \frac{1}{2} \equiv 1 \pmod{11}$$

$\frac{1}{2} \equiv 6 \pmod{11}$ by noting that $2 \cdot 6 \equiv 1 \pmod{11}$ and thus,

$$54 \cdot (n-1) \equiv 10 \cdot (n-1) \equiv 1 \pmod{11}$$

This can be converted to an equation, $10 \cdot (n-1) = 11w + 1$ An obvious value of w that results in an integer value of n is $w = 9$ as $10 \cdot (n-1) = 100$ meaning that $n = 11$ would work. Note, for the next part, I used a calculator frequently.

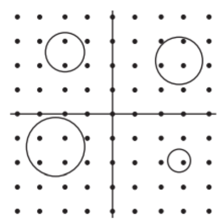
We have seen that $n = 11$ creates a not square-free $10^{11} + 1$. Meaning that $10^{11} + 1 = 11^2 \cdot 23 \cdot 4093 \cdot 8779 = 11^2 \cdot 826446281$. Note that 826446281 has 9 digits. At the very start of the problem, we defined N to have n digits, so our number needs to have 11 digits and be in the form $826446281 \cdot a^2$ such that it has 11 digits with the former having 9 digits. The easiest a to think about is $a = 10$ because it adds the two digits by multiplying by 100. Thus, our N is 82644628100. Now using a calculator once again, we see that
$\text{echo}(82644628100) = 90909090910^2$.

We can also see that for some number of "90"s, $9090\ldots91$ will be divisible by 11 and thus if its not square-free, there is a number N with n digits that will make $\text{echo}(N)$ a perfect square again because there are infinite numbers such that $11^2 \mid 10^n + 1$ which can be understood because $9090\ldots91$ just needs to follow the divisibility rule for 11, and there is always a perfect square such that you can multiply $9090\ldots91/11$ with it to get a number with the correct number of digits.

3 Problem 3

Problem Three

A lattice point is a point (x, y) in the plane, both of whose coordinates are integers. It is easy to see that every lattice point can be surrounded by a small circle which excludes all other lattice points from its interior. It is not much harder to see that it is possible to draw a circle which has exactly two lattice points in its interior, or exactly 3, or exactly 4, as shown in the picture below.



The main idea is to prove that for some center (a, b), the distance from all points are different meaning that increasing the radius a tiny bit will only include one and not two lattice points. This also means that the distance from the center and every single lattice point is different. The reasoning behind this is because the distance from the center to a lattice point (if it were to be included) is the radius and all lattice points included for a smaller radius are obviously included as well.

Trying a random center say $(3, 4)$, there are multiple reasons why this doesn't work. Quite simply $(3, 3)$ and $(3, 5)$ are equidistant and two new lattice points are introduced once the radius reaches two. We can make a generalization that if the center is any lattice point, this will not work. This is because there are lattice points that clearly have the same distance from the center (which is also a lattice point). Trying for (any) rational numbers, we see the same conclusion except that it isn't as obvious.

Take two points $A(x_1, y_1)$ and $B(x_2, y_2)$ (where both are lattice points where x_1, y_1, x_2, y_2 are all integers) and a center (a, b). We are trying to find some (a, b) such that the following equality is not possible.

$$(x_1 - a)^2 + (y_1 - b)^2 = (x_2 - a)^2 + (y_2 - b)^2$$

There are multiple ways to continue from here:
Way 1: We are trying to show that the distance from the center to both of the points cannot be equal. After experimenting a bit, we see that if one of the coordinates is irrational and the other is rational two lattice points cannot have the same distance, here is the proof:
Trying some random center that follows the above idea: $(a, b) = (\frac{5}{7}, \sqrt{7})$

$$(x_1 - \frac{5}{7})^2 + (y_1 - \sqrt{7})^2 = (x_2 - \frac{5}{7})^2 + (y_2 - \sqrt{7})^2$$

$$x_1^2 - \frac{10}{7} \cdot x_1 + \frac{25}{49} + y_1^2 - 2\sqrt{7} \cdot y_1 + 7 = x_2^2 - \frac{10}{7} \cdot x_2 + \frac{25}{49} + y_2^2 - 2\sqrt{7} \cdot y_2 + 7$$

$$x_1^2 - \frac{10}{7} \cdot x_1 + y_1^2 - 2\sqrt{7} \cdot y_1 = x_2^2 - \frac{10}{7} \cdot x_2 + y_2^2 - 2\sqrt{7} \cdot y_2$$

Because $2\sqrt{7}$ is irrational, $y_1 = y_2$ and thus equating and simplifying, it is easily seen that $x_1 = x_2$. Thus the points A and B are the same.

Way 2: We are trying to show the same thing as Way 1, but in a different way. We choose the same center $(\frac{5}{7}, \sqrt{7})$.

$$(x_1 - \frac{5}{7})^2 + (y_1 - \sqrt{7})^2 = (x_2 - \frac{5}{7})^2 + (y_2 - \sqrt{7})^2$$

$$(x_1 - \frac{5}{7})^2 - (x_2 - \frac{5}{7})^2 = (y_2 - \sqrt{7})^2 - (y_1 - \sqrt{7})^2$$

Simplifying we see,

$$x_1^2 - x_2^2 - \frac{10}{7} \cdot (x_1 - x_2) = y_2^2 - y_1^2 - 2\sqrt{7} \cdot (y_2 - y_1)$$

$$x_1^2 - x_2^2 - \frac{10}{7} \cdot (x_1 - x_2) - y_2^2 + y_1^2 = 2\sqrt{7} \cdot (y_2 - y_1)$$

We should remember that x_1, y_1, x_2, y_2 are all integers meaning that one side is rational while the other side is irrational and a rational and irrational can never be equal. The only way to get rid of the irrational part on the RHS is by equating $y_2 = y_1$

$$x_1^2 - x_2^2 - \frac{10}{7} \cdot (x_1 - x_2) - y_2^2 + y_1^2 = 0$$

$$x_1^2 - x_2^2 - \frac{10}{7} \cdot (x_1 - x_2) = y_2^2 - y_1^2$$

$$x_1^2 - x_2^2 - \frac{10}{7} \cdot (x_1 - x_2) = 0$$

$$x_1^2 - x_2^2 = \frac{10}{7} \cdot (x_1 - x_2)$$

$$x_1 + x_2 = \frac{10}{7}$$

However, we defined all x_1, y_1, x_2, y_2 to be integers and two integers cannot have a sum of a fraction. Thus, this is a contradiction. We can state this more generally. Pretend we don't have a center but we know that the one coordinate is irrational while the other is rational. We have, $(x_1 - a)^2 + (y_1 - b)^2 = (x_2 - a)^2 + (y_2 - b)^2$. Now there are two options, if a is irrational, $x_1 = x_2$. This leads to $(y_1 - b)^2 = (y_2 - b)^2$ and then either $y_1 - b = y_2 - b \Rightarrow y_1 = y_2$ or $y_1 - b = -y_2 + b \Rightarrow \dfrac{(y_1 + y_2)}{2} = b$ which means that $2 \not| b$.

Another kind of proof which is kind of like a proof by contradiction is that. If there are two lattice points that are introduced and are on the circumference of the circle, if we draw the line between the two points, a good observation is that the center lies on the perpendicular bisector of the that line. One thing we should note is that the line of the perpendicular bisector has a slope of the negative reciprocal of the line between the two lattice points. The line between the two lattice points has a rational slope (because they are all integers) and so the negative reciprocal will be a rational number as well. The perpendicular bisector obviously goes through the midpoint of the line (which also has rational coordinates) and thus the line can be in the form where the slope is rational and so is y-intercept. Thus, the only way for one of the coordinates to be irrational is if the other coordinates is irrational. Thus, we know that the center cannot be in the form of two irrational coordinates.

Conclusion: It is possible for there to be n lattice points inside a circle for all n if the center of the circle has one irrational and one rational coordinate where the rational coordinate cannot have a factor of 2 in the numerator. I also proved that the coordinates of the center cannot be in the form of two irrationals by using the perpendicular bisector to link two lattice points.

4 Problem 4

Problem Four

Calculate each of the following:
$1^3 + 5^3 + 3^3 =??$
$16^3 + 50^3 + 33^3 =??$
$166^3 + 500^3 + 333^3 =??$
$1666^3 + 5000^3 + 3333^3 =??$
$\cdots =??$
What do you see? Can you state and prove a generalization of your observations?

Note: *This problem took me a really long time to think about or to make considerable progress. I thought a while about how to link* $16, 50, 33$*;* $166, 500, 333$ *and I continue my solution based on that.*

We first start by evaluating the expressions in the problem,
$1^3 + 5^3 + 3^3 = 153$
$16^3 + 50^3 + 33^3 = 165033$
$166^3 + 500^3 + 333^3 = 166500333$

It is easy to see that in simple terms, the addition signs and the powers are removed and it is appended to each other. It can be generalized that,

Generalization: $(\underbrace{16\ldots6}_{n\text{ digits}})^3 + (\underbrace{50\ldots0}_{n\text{ digits}})^3 + (\underbrace{33\ldots3}_{n\text{ digits}})^3 = \underbrace{16\ldots6}_{n}\underbrace{50\ldots0}_{n}\underbrace{33\ldots3}_{n}$. What I thought of was that $166 = \frac{10^3-4}{6}, 500 = \frac{10^3}{2}, 333 = \frac{10^3-1}{3}$ and this works for all the following numbers just increase the exponent of 10. Say the number of digits of n we can make an equation. The following is the LHS of the Generalization above with the use of the "link" of $16\ldots6, 50\ldots0, 33\ldots3$ to 10^n. Here is an algebraic interpretation of the following idea:

$$\left(\frac{(10^n-4)}{6}\right)^3 + \left(\frac{10^n}{2}\right)^3 + \left(\frac{(10^n-1)}{3}\right)^3$$
$$= \frac{(10^n-4)}{6}\cdot 10^{2n} + \frac{10^n}{2}\cdot 10^n + \frac{(10^n-1)}{3}$$

Starting with the LHS,

$$\frac{10^{3n}-12\cdot 10^{2n}+48\cdot 10^{n}-64}{216}+\frac{10^{3n}}{8}+\frac{10^{3n}-3\cdot 10^{2n}+3\cdot 10^{n}-1}{27}$$

Simplfying,

$$\frac{10^{3n}-12\cdot 10^{2n}+48\cdot 10^{n}-64+27\cdot 10^{3n}+8\cdot[10^{3n}-3\cdot 10^{2n}+3\cdot 10^{n}-1]}{216}$$

Simplfying further,

$$\frac{36\cdot 10^{3n}-36\cdot 10^{2n}+72\cdot 10^{n}-72}{216}=\frac{10^{3n}}{6}-\frac{10^{3n}}{6}+\frac{10^{n}}{3}-\frac{1}{3}.$$

Now looking at the RHS,

$$\frac{(10^{n}-4)}{6}\cdot 10^{2n}+\frac{10^{n}}{2}\cdot 10^{n}+\frac{(10^{n}-1)}{3}=\frac{10^{3n}}{6}-\frac{2}{3}\cdot 10^{2n}+\frac{10^{2n}}{2}+\frac{10^{n}}{3}-\frac{1}{3}$$

Simplfying, we do see that

$$\frac{36\cdot 10^{3n}-36\cdot 10^{2n}+72\cdot 10^{n}-72}{216}=\frac{10^{3n}}{6}-\frac{10^{3n}}{6}+\frac{10^{n}}{3}-\frac{1}{3}$$

Finally,

$$\frac{10^{3n}}{6}-\frac{10^{3n}}{6}+\frac{10^{n}}{3}-\frac{1}{3}=\frac{10^{3n}}{6}-\frac{10^{3n}}{6}+\frac{10^{n}}{3}-\frac{1}{3}.$$

So now we have proved the generalization made above, for all n,

$$(\underbrace{16\ldots6}_{n\ \textbf{digits}})^3+(\underbrace{50\ldots0}_{n\ \textbf{digits}})^3+(\underbrace{33\ldots3}_{n\ \textbf{digits}})^3=\underbrace{16\ldots6}_{n}\underbrace{50\ldots0}_{n}\underbrace{33\ldots3}_{n}\mathbf{.}$$

5 Problem 5

Problem Five

Any positive integer can be written in binary (also called base 2). For example, 37 is 100101 in binary (because 37 = 25 + 22 + 20), and 45 is 101101 in binary. Let's say that a positive integer is "scattered" if, in its binary expansion, there are never two ones immediately next to each other. For example, 37 is scattered but 45 is not. How many scattered numbers are there less than 4? Less than 8? Less than 2^n?

Note: *An n-digit scattered number starts with the digit* 1 *and that there are never two ones adjacent to each other. For example, a* 6*-digit scattered number can be* 100101_2 *but can't be* 010101_2*. Also, let n denote the number of digits in the binary representation of a number.*

First finding the binary representation of all base-10 numbers less than 8,

$0=0_2, 1=1_2,\ 2=10_2,\ 3=11_2,\ 4=100_2,\ 5=101_2,\ 6=110_2,\ 7=111_2$

Part 1: Inspection By inspection, the scattered numbers less than 4 are $\{0, 1, 2\}$, and the scattered numbers less than 8 are $\{0, 1, 2, 4, 5\}$. We can see that the number of scattered numbers less than $4 = 2^2$ are simply the sum of the number of scattered numbers that have a one digit binary representation and a two digit binary representation. Similarly, the number of scattered numbers less than $8 = 2^3$ are the sum of the number of scattered numbers that have a one digit binary representation, a two digit binary representation, and a three digit binary representation.

Generalization: The number of scattered numbers less than 2^n is equal to the sum of the number of scattered numbers that have a 1 digit binary representation, a 2 digit binary representation,. . . , an n digit binary representation.

Let n denote the number of digits in the binary representation of a number. Testing for the number of scattered numbers for the first five n.

If $n = 1$, there are 2 valid scattered numbers 2: $\{1_2\}$.
If $n = 2$, there is 1 valid scattered number $\{10_2\}$.
If $n = 3$, there are 2 valid scattered numbers: $\{100_2, 101_2\}$.
If $n = 4$, there are 3 valid scattered numbers: $\{1000_2, 1001_2, 1010_2\}$.
If $n = 5$, there are 5 valid such numbers: $\{10000_2, 10001_2, 10010_2, 10100_2, 10101_2\}$.

It is easy to see that the number of valid scattered numbers for each n is the Fibonacci numbers and so is the number of scattered numbers less than 2^n.

Definition I: Let a_n denote the denote the number of scattered numbers which has n digits in its binary representation.

Definition II: Let s_n denote the number of scattered numbers less than 2^n.

Calculating the first few s_n, $s_2 = 4, s_3 = 4, s_4 = 7, s_5 = 12$. Now its easy to see that the first few s_n are the Fibonacci numbers minus 1. Also, the first few a_n form the Fibonacci sequence exactly.

Why is it Fibonacci? We start by proving a_n is Fibonacci. To start, we essentially want to show that $a_n = a_{n-1} + a_{n-2}$. Start with the case $n = 4$, we want to show how we can build a_n from a_{n-1} and a_{n-2}.

Notice that the valid scattered numbers for $n = 3$ is in $n = 4$ with a 0 appended to the number. Similarly, the scattered numbers for $n = 2$ is in $n = 4$ with a 01 appended to the number. This can be seen to work with $n = 5$ and any $n \geq 4$. This can be generalized to any $n \geq 4$: The number of valid scattered numbers with n digits in its binary representation can be constructed by appending the digit 0 to all elements in the subset corresponding to $n - 1$ and as well as appending the digit 01 to the elements in the subset corresponding to $n - 2$. Which can be proven with a one-to-one correspondence.

One key observation in the proof to show why a_n is Fibonacci is noting that $n = 1$ has an extra term (this will cancel out later on). Start by calculating the sum of the first n terms in a Fibonacci sequence denoted by S_n which we know is $s_n - 1$

$S_1 = 1, S_2 = 2, S_3 = 4, S_4 = 7, \ldots$. It can be observed that $S_n = a_{n+2} - 1$ but we need to prove this. We proceed by induction. Starting with the base case n=1, $S_1 = a_3 - 1 = 1$ which is true. Then assume that $S_n = a_{n+2} - 1$, we need to prove that $S_{n+1} = a_{n+3} - 1$. We can do this by simply adding a_{n+1} on both sides to $S_n = a_{n+2} - 1$ which results in $S_{n+1} = a_{n+3} - 1$. Thus, we have proved using induction that the sum of the first n Fibonacci numbers in the sequence a_n sum to $a_{n+3} - 1$. Now because there was an extra 1 in the n equals one case way above (for the valid scattered numbers with 1 digit), we can say that $s_n = S_n + 1 = a_{n+2}$.

Conclusion + Generalization: The number of scattered numbers less than $2^n = a_{n+2}$ where the series a_n is the Fibonacci sequence.

6 Problem 6

Problem Two

The set S contains some real numbers, according to the following three rules.(i) $\frac{1}{1}$ is in S. (ii) If $\frac{a}{b}$ is in S, where $\frac{a}{b}$ is written in lowest terms (that is, a and b have highest common factor 1), then $\frac{b}{2a}$ is in S. (iii) If $\frac{a}{b}$ and $\frac{c}{d}$ are in S, where they are written in lowest terms, then $\frac{a+c}{b+d}$ is in S. These rules are exhaustive: if these rules do not imply that a number is in S, then that number is not in S. Can you describe which numbers are in S?

We start by creating bounds for the values in the set S which I claim is $\frac{1}{2}$ and $\frac{1}{1}$ where there is no fraction in the set S greater than 1. We start by noting the only thing we can do to get an element in S with only having $\frac{1}{1}$ is using part ii to add $\frac{1}{2}$ to the set. It is a known fact that for two fractions, $\frac{a}{b}$ and $\frac{c}{d}$, the fraction $\frac{a+c}{b+d}$ lies between these two fractions (I will provide a proof at the end). There are two possible ways to construct $\frac{c}{d}$ either by case 2 or case 3. The second case says that that we can add $\frac{b}{2a}$ to the set S if $\frac{a}{b}$ is in the set and the only fraction we are given is $\frac{1}{1}$ and $\frac{b}{2a}$ will always be equal to either $\frac{1}{2}$. We should note that the greater $\frac{a}{b}$ is, the smaller $\frac{b}{2a}$ and we said that the greatest. Essentially, the use of case 3 will always be in the middle of the two fractions and thus the second fraction needs to be greater than 1 for the use of case 3 to be greater than one. However, there is no rule that allows a number to be greater than 1 directly or indirectly because $\frac{b}{2a}$ is always less than or equal to one for the bounds and if it is greater, that means the fraction is less than 1/2 and there is no way to produce that. The main idea of this part is showing that going outside of the bounds are not possible unless another number goes outside the bounds all the way until 1/2 and 1/1 which are in the bounds. Like I would think of this like a domino effect where if one of them are greater than 1 or less than 1/2 there is room for others to be, but if none of them are, they cannot be as they are dependent on the previous fractions. Now we have proved that the bounds for the rational numbers in the set S is $[\frac{1}{2}, 1]$. The only thing left is to show that all rational numbers can be derived from the rules given in the problem statement.

Proof that $\frac{a}{b} \leq \frac{a+c}{b+d} \leq \frac{c}{d}$. We start by dividing this up into two inequalities: we see three cases

$$\frac{a}{b} \leq \frac{c}{d}$$

$$\frac{a}{b} \leq \frac{a+c}{b+d}$$

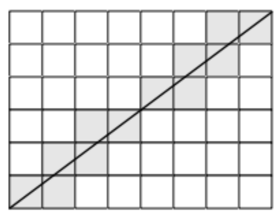
$$\frac{a+c}{b+d} \leq \frac{c}{d}$$

From the first inequality, we see that $ad \leq bc$. The second inequality yields $ad \leq bc$. The third yields the same thing. Thus, dividing, we see $\frac{a}{b} \leq \frac{c}{d}$.

7 Problem 7

Problem Seven

The rectangular floor of a bathroom is covered with square tiles (all of the same size). A spider starts at one corner of the bathroom, and walks to the corner diagonally opposite. For example, the figure below shows a 6×8 bathroom, in which the spider touches 12 tiles on its path. (A spider doesn't touch a tile if it just walks over the grout at the corner of a tile.) For an m by n bathroom, how many tiles does the spider touch on its walk?



Note: *This solution is based on the lattice points the diagonal intersects and is built up from there In advance, s is the width, r is the height of the rectangle. The lattice points are important because that is where we over count because there is only one square where the diagonal intersects a horizontal and vertical line at the same time, which will be the primary focus of the solution.*

Start by placing the figure given in the problem on the coordinate plane.

Let the bottom-left vertex have coordinates $(0,0)$, top-left vertex have coordinates $(0,r)$, bottom-right vertex $(s,0)$, top-right vertex have coordinates (s,r). The equation for the line of the diagonal of the rectangle of vertices above is,

$$y = \frac{r}{s}x.$$

If r is relatively prime to s, $\frac{r}{s}$ cannot be simplified and thus it only goes through two lattice points, the bottom-left and the top-right vertex.

If r and s share a factor, the number of lattice points it goes through is $\gcd(r,s)+1$. This results from the slope.

The reason is that, $\frac{r}{s}x = \frac{\frac{r}{\gcd(r,s)}}{\frac{s}{\gcd(r,s)}}x$. We know that the bounds on x are $0 \leq x \leq s$, and thus for it to intersect a lattice point, $x = \frac{s}{\gcd(r,s)} \cdot k$ for some k such that x is still in the bounds

$0 \leq x \leq s \Rightarrow 0 \leq \frac{s}{\gcd(r,s)} \cdot k \leq s$. We can see that the two equality cases happen when $k = 0$ and when $k = \gcd(r, s)$. Thus, the number of lattice points it goes through is $\gcd(r, s) + 1$ or the number of integers between 0 and $\gcd(r, s)$. Using this, we can see that the number of lattice points it goes through that **are not vertices** are $\gcd(r, s)$. This is because that extra one is the top-right vertex. We will get back to this.

Now we can continue to the number of squares it goes through. Take two cases: the vertical and horizontal lines. Starting with the vertical lines, every time the diagonal intersects a vertical line, it goes through the square that contains the same height in the column before. If we take the 8x6 rectangle given in the problem statement, starting from the second vertical line (because there are no squares that are to the left of the first vertical line), there are 8 intersections and thus 8 squares. Now considering vertically, every time the diagonal intersects a horizontal line that is above the base, it intersects the square that is in the same row that contains the x coordinate of the intersection. This means that it intersects 6 times in the given figure. However, we need to count for overlap. This happens at $(4, 3)$ and $(8, 6)$ and because its counted double, we need to subtract two. This leaves us with $8 + 6 - 2 = 14$ which gets us the same answer. A generalization shows us that its $a + b - \gcd(a, b)$ for sides of length a and b. The $\gcd(a, b)$ is there because we don't want to double count the lattice point intersection excluding the starting point. The reason we exclude the starting point is because the way we set up the "algorithm" already includes it but only once.

Thus, we can create a general algorithm:
i) Number of vertical lines that exclude the left most vertical line
ii) Number of horizontal lines that exclude the bottom most line.
iii) The union of the i + ii; the number of intersection at lattice points.

It just so happens that the first part is equivalent to the length of the width of the rectangle, the second part is equivalent to the length of the height of the rectangle, and the third part is equal to the $\gcd(a, b)$ with side lengths a and b.

We can conclude that the number of squares it goes through is equal to
$a + b - \gcd(a, b)$ which is because of the algorithm made above that can be used for any dimensions of a rectangle.

Note that I have seen a similar idea before. I am not sure where I have seen it, but I have seen a similar idea used in some problem before.

8 Problem 8

Problem Eight

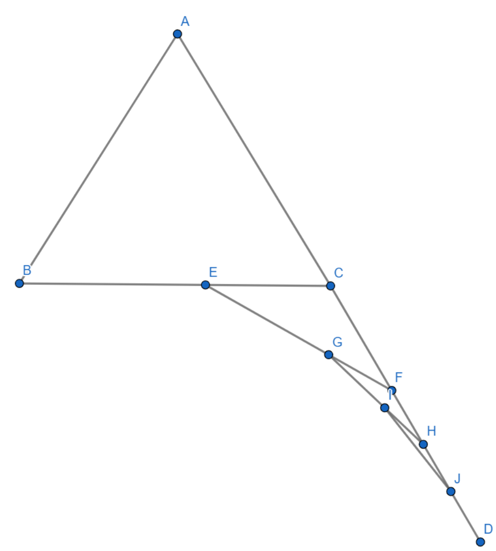
Let P_0 be an equilateral triangle of area 10. Each side of P_0 is trisected into three segments of equal length, and the corners of P_0 are snipped off, creating a new polygon (in fact, a hexagon) P_1. What is the area of P_1? Now repeat the process to P_1 – i.e. trisect each side and snip off the corners – to obtain a new polygon P_2. What is the area of P_2? Now repeat this process infinitely to create an object P_∞. What can you say about the shape P_∞? What is the area of P_∞?

Note: *The main idea of the solution is the idea of a "nested" infinite geometric series which will calculate the area of the remaining portion. The diagram that I made for the problem essentially shows the starting of the infinite number of triangles.*

We can start by noting that when drawing the diagram, the number of corners that are cut off doubles each cycle. When making the hexagon, only 3 triangles are cut off but when making the 12-gon, 6 triangles are cut off, and so on. This can also be viewed as the number of triangles cut off are the same number as half of the number of sides of the resulting polygon.

Now, we start by calculating the side length of largest equilateral triangle which has area 10. This can be done so using the formula for the area of an equilateral triangle, $\frac{s^2\sqrt{3}}{4} = 10$, $s = \sqrt{\frac{40}{\sqrt{3}}}$. Now we can use this to calculate the area of P_1. This is equal to, in words, the total area minus three times the area of the equilateral triangles. The side length of the three smaller equilateral triangles is 1/3rd of the side length of the larger equlateral triangle, and thus we can see that the area of $\frac{10}{9}$. There is a smarter way to do this by simply noting that for similar triangles, the area of one is the scale factor squared multiplied by the other, and thus here the scale factor is $\frac{1}{3}$ and thus, the area of the smaller equilateral triangle is $\frac{10}{9}$ make the area of the hexagon $10 - 3 \cdot \frac{10}{9} = \frac{20}{3}$.

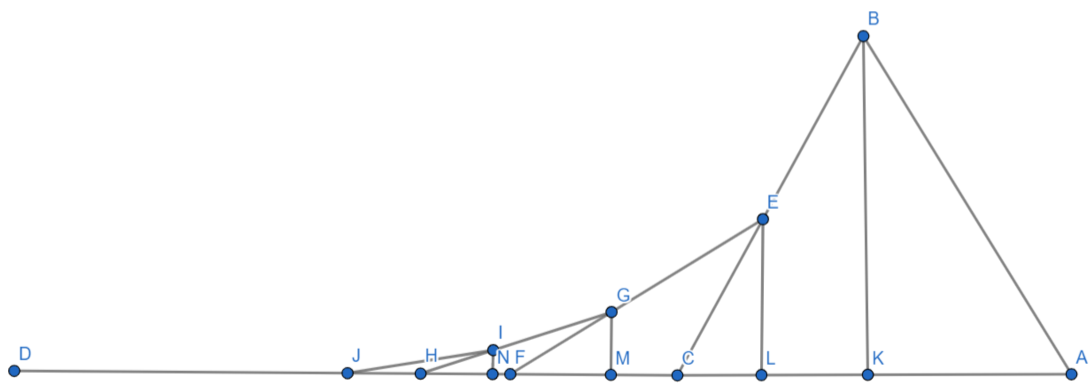
Now we go onto the triangles that are chopped off from the hexagon to form the 12-gon, which is in an isosceles triangle with two sides having length $\frac{s}{9}$ and angle 120. We can use the sine-area formula to calculate the area which is $\frac{1}{2} \cdot \frac{\sqrt{3}}{2} \cdot \frac{s^2}{81} = \frac{\sqrt{3}}{4} \cdot \frac{40}{81\sqrt{3}} = \frac{10}{81}$. Now is where the figure below comes in use,



The first area we found (of the equilateral triangle) is of ABC, while the second area we found is of ECF. Now it seems like it looks like an infinite geometric series with $r = \frac{1}{9}$. However, we need to prove this, which I think can be done by examining the diagram above.

Now looking at the third triangle down, we can try to calculate its area. We know from supplementary angles that the obtuse angle in the third triangle is 150. It is worth noting that this triangle is **not** isosceles so we need to find the two distinct values of the sides. One side is simply $\frac{s}{27}$ while the other is the the length GF which is $\frac{1}{3} \cdot EF$. Drawing the altutitude from C to EF of the CEF, we see that the triangle is two 30-60-90 triangles stuck together, and because we know $EC = \frac{s}{9}$, $EF = 2 \cdot \frac{s\sqrt{3}}{18} = \frac{s\sqrt{3}}{9}$ which means that $GF = \frac{s\sqrt{3}}{27}$.

Now it is easy to calculate the area of GFH using the sine-area formula. $\frac{1}{2} \cdot \sin 150 \cdot \frac{s\sqrt{3}}{27} \cdot \frac{s}{27} = \frac{1}{4} \cdot \frac{s^2\sqrt{3}}{3^6} = \frac{1}{4} \cdot \frac{\frac{40}{\sqrt{3}} \cdot \sqrt{3}}{3^6} = \frac{1}{4} \cdot \frac{40}{3^6} = \frac{10}{729}$. We can again see that it is multiplied by $\frac{1}{9}$. Now I'm



pretty sure that the problem is based on an infinite geometric series with common ratio of $\frac{1}{9}$.

"Proof" of $\frac{1}{9}$ factor: Using the diagram above, our main idea is using the heights. The diagram has an infinite number of triangles in the range from one vertex to a midpoint of an equilateral triangle containing the point.
For example, in some equilateral triangle BYZ, A is the midpoint of either BY or BZ. Where A and B are the same points as in the diagram above.

It is simple to see once the altitudes are drawn why the area ratios are $\frac{1}{9}$ (it is easiest to see when the base of the area of the triangle is on AD). When the altitudes are drawn, it creates a pair of similar triangles. Take the example of EFC and ABC. In this, the bases already have a ratio of $\frac{1}{3}$. Using similar triangles(AA similarity) on ELC and BKC, we see that heights have a ratio of $\frac{1}{3}$ as well (which happens because 3*EC = BC). Now we can see this works for any two consecutive triangles on line AD, and they all have a area ratio of $\frac{1}{9}$. More generally, the smaller base is a 1/3rd of the larger base and the smaller height is 1/3rd of the larger height.

Now that we have somewhat proved why the common ratio for the geometric series is $\frac{1}{9}$, we can proceed by doing some algebra. The general form of the series is,

$$10 - \left(3 \cdot \frac{10}{9} + 6 \cdot \frac{10}{81} + 12 \cdot \frac{10}{729} + \dots\right)$$
$$= 10 - 30\left(\frac{1}{9} + \frac{2}{81} + \frac{4}{729} + \dots\right)$$
$$= 10 - 30\left(\frac{1}{9} + \frac{2^1}{9^2} + \frac{2^2}{9^3} + + \frac{2^3}{9^4} \dots\right)$$

By the formula for the sum of an infinite geometric series, we easily see that the inside sum is equal to,

$$\frac{\frac{1}{9}}{1 - \frac{2}{9}} = \frac{1}{7}.$$

And thus, plugging this back in, we see

$$10 - 30 \cdot \frac{1}{7} = \boxed{\frac{40}{7}}.$$

9 Problem 9

Problem Two

A monkey has filled in a 3×3 grid with the numbers 1, 2, . . . , 9. A cat writes down the three numbers obtained by multiplying the numbers in each horizontal row. A dog writes down the three numbers obtained by multiplying the numbers in each vertical column. Can the monkey fill in the grid in such a way that the cat and dog obtain the same lists of three numbers? What if the monkey writes the numbers 1, 2, . . . , 25 in a 5×5 grid? Or 1, 2, . . . , 121 in a 11×11 grid? Can you find any conditions on n that guarantee that it is possible or any conditions that guarantee that it is impossible for the monkey to write the numbers 1, 2, . . . , n^2 in an $n \times n$ grid so that the cat and the dog obtain the same lists of numbers?

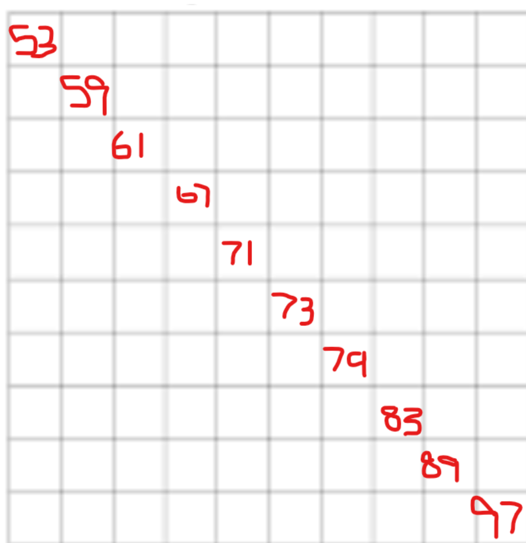
Citation on Wikipedia page on Prime Number Theorem for extra research
One thing that is definitely true when looking at the numbers is that if the number of primes between $\frac{n^2}{2}$ and n^2 is greater than n, then there cannot be a a way for the cats and dogs to have the same product. This is because the prime factor is only repeated once if it is in that range, and thus the row which those prime factors are in and the column which those prime factors are in are the ones which have the same product. Essentially imagine the squares that go diagonally from one end to the other. This is the place they will be placed (not necessarily true if the number of primes in the range is less than n) but the maximum number of spaces is n which is the diagonal. Also another thing we know about n is that $n \neq 2$ because when you try out all possible arrangements, none of them follow the requirements. Listing out all of the prime numbers ≤ 121.

$$\{2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113\}$$

This is important because my main idea is based on the number of primes between $\frac{n^2}{2}$ and n^2. There are two ways to continue from here, on my first try I just brute-forced the first 11. The number of primes between $\frac{n^2}{2}$ and n^2 for the first 11 n (excluding 1 and 2 because we have already done so)

$n = 3,\ \frac{n^2}{2} \le 5, 7 \le n^2$
$n = 4,\ \frac{n^2}{2} \le 11, 13 \le n^2$
$n = 5,\ \frac{n^2}{2} \le 13, 17, 19, 23 \le n^2$
$n = 6,\ \frac{n^2}{2} \le 19, 23, 29, 31 \le n^2$
$n = 7,\ \frac{n^2}{2} \le 29, 31, 37, 41, 43, 47 \le n^2$
$n = 8,\ \frac{n^2}{2} \le 37, 41, 43, 47, 53, 59, 61 \le n^2$
$n = 9,\ \frac{n^2}{2} \le 41, 43, 47, 53, 59, 61, 67, 71, 73, 79 \le n^2$
$n = 10,\ \frac{n^2}{2} \le 53, 59, 61, 67, 71, 73, 79, 83, 89, 97 \le n^2$
$n = 11,\ \frac{n^2}{2} \le 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113 \le n^2$

Counting up, we see that $n = 9$ has 10 numbers and using the explanation above, we know that $n = 9$ does not work and similarly $n = 11$ also doesn't work as it has 15 numbers.
To illustrate,



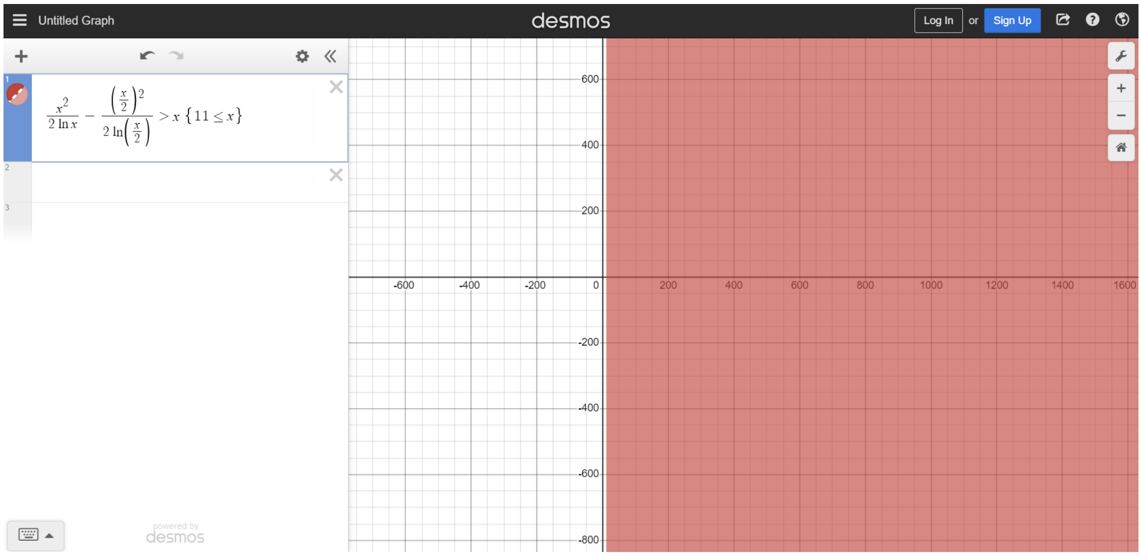
Note: I am not saying this is the order, however, those primes in the range need to be on that diagonal (or another diagonal) for there to be a solution. Take the 59 row and column for example. The cat multiples out the second from the top row while the dog multiplies out the 2nd from the left column and those are supposed to have the same product. This applies to all of the numbers on the diagonal.

Now we use the Prime Number Theorem to show that there isn't a square that works for all $n > 11$. We do this by noting that the Prime Number Theorem essentially says that the number of primes less than or equal to a certain number can be calculated by $\frac{n}{ln(n)}$. We can use the idea we developed at the very starting of the solution and apply it here for $n \geq 11$

$$\frac{n^2}{ln(n^2)} - \frac{(\frac{n}{2})^2}{ln((\frac{n}{2})^2)} > n$$

$$\frac{n^2}{2ln(n)} - \frac{(\frac{n}{2})^2}{2ln(\frac{n}{2})} > n$$

The above figure illustrates (using Desmos) that using PNT, all $n \geq 11$ cannot result in the



dog and the cat having the same numbers after all the products are done by showing that there are too many primes than there are spaces that would work. However, when graphing without the range on Desmos, we see that there is some type of error largely due to the function in PNT not being the exact value but an approximation.

10 Problem 10

Problem Two

On a strange railway line, there is just one infinitely long track, so overtaking is impossible. Any time a train catches up to the one in front of it, they link up to form a single train moving at the speed of the slower train. At first, there are three equally spaced trains, each moving at a different speed. You watch, and eventually (after all the linking that will happen has happened), you count the trains. You wonder what would have happened if the trains had started in a different order (but each of the original three trains had kept its same starting speed). On average (averaging over all possible orderings), how many trains will there be after a long time has elapsed? What if at the start there are 4 trains (all moving at different speeds)? Or 5? Or n?

In advance, my solution is based on casework for where the the notation we will use in the problems are the capital letters in the alphabets, where A is has the smallest speed, and as we progress through the alphabet, the speed of that specific train increases. A key idea that would be good to notice is that all trains behind A will at some point merge into A, and no train in front of A will connect with the train A. Another thing we should note before we start the solution is that I calculate the total value and then divide based on the probability. This is because the probability of any happening is equal (for example, ABC has the same probability as CAB to be the trains configuration). Something else we need to define early on is that when multiplying, I will always put the number of ways and then the number of trains, for example: $a \cdot b$ is where there are a ways to get the a b number of trains. Another thing we should note is that the factorial will be used frequently because the number of ways we place n trains in some order is $n!$

So we start by seeing the smaller cases. When $n = 2$, there are two possible cases AB and BA. In the first case, there is only one train and on the other case there are two. Thus, the expected value is $\frac{1}{2} \cdot (1 + 2) = \frac{3}{2}$.

Now when $n = 3$, there are 6 cases but thinking about it, the factorial increases very fast and it would be very hard to list out $5! = 120$ cases. Thus, we employ early on what I talked about earlier: casework on the slowest train A. If A is the first train (with the other two trains behind it), there are two ways to place the other two trains with both having the same outcome. If A is the middle train, the train behind it will merge with A but the new train will not merge with the starting (because A is the slowest). Thus, there are two ways that both end up with 2 trains at the end: $2 \cdot 2 = 4$. If A is the last train, it is harder to calculate. There are two ways again, CBA and BCA. In the first way, there is no merging so there always will end up being 3 trains and for the latter B and C will merge so it will end up with 2 trains and thus this totals up to $\frac{2\cdot1+2\cdot2+1\cdot2+1\cdot3}{3!} = \frac{11}{6}$.

Now when $n = 4$, we do casework on where A is once again. If A is the first train, there are 6 ways to organize the other trains in order, and we know that the outcome will always be 1. If A is the 2nd train, all trains before A will merge with A and thus there are two trains for all possible arrangements of the list of trains in order: $6 \cdot 2$. If A is the third train, the last train will merge with A and thus the first two are the only ones that matter. There are 6 options to arrange the first two and only 3 where they will merge. BC, BD, CD and 3 where they won't merge CB, DB, DC. And thus, $3 \cdot 2 + 3 \cdot 3$ is the total when A is the third train in the line. If A is at the end, there are three empty spaces ahead of it. There are two ways we can calculate this case: either by "bashing" it out or by using a previous thing we calculated, when $n = 3$. We need to manipulate the numerator of the fraction right before simplification. We need to add one to each of the values (aka. the right side of each product as we said at the start of the solution) by adding one to account for the new A at the end. Meaning that $2\cdot1+2\cdot2+1\cdot2+1\cdot3$ becomes $2\cdot(1+1)+2\cdot(2+1)+1\cdot(2+1)+1\cdot(3+1)$. Other than that it is "mostly" recursion. Using this, we see that it is $2\cdot2+2\cdot3+1\cdot3+1\cdot4 = 17$. Summing it all up, we see that the expected value if there are 4 trains is $\frac{50}{24}$.

First thought that comes to my mind was maybe it wasn't really recursion and we build it from n itself. It is clear that the denominators of each fraction is a factorial (assume that we don't simplify) and I don't think you can form it from the n because the expected value when $n = 4$ or $\frac{50}{24}$ is greater than 2 and the others are less than 2 meaning that you can't multiply something to n (easily) or anything related to n to produce the numerator.

Now, the next thing we should try is recursion.Looking at the fractions $\frac{3}{2}, \frac{11}{6}, \frac{50}{24}$, we are looking to find a similarity or some form of recursion (aka. building the next term from the previous term).

$$\frac{11}{6} = \frac{3 \cdot 3 + 2!}{2! \cdot 3} = \frac{3}{2} + \frac{1}{3}$$
$$\frac{50}{24} = \frac{11 \cdot 4 + 3!}{3! \cdot 4} = \frac{11}{6} + \frac{1}{4}$$

Just for reference, I also calculated when there were 5 trains but I decided not to include it because it took nearly a whole page $\Rightarrow \frac{274}{120}$

$$\frac{274}{120} = \frac{50 \cdot 5 + 4!}{5 \cdot 4!} = \frac{50}{4!} + \frac{1}{5}$$

Now the idea seems clearer. Declare the function f to be the expected value number of trains. $f(2) = \frac{3}{2}$, $f(3) = \frac{11}{6}$, $f(4) = \frac{50}{24}$, $f(5) = \frac{274}{120}$ we can generalize this based on the 4 values that were calculated:

$$f(n+1) = f(n) + \frac{1}{n+1}$$

To further explores if and why this works, we will examine the $n = 5$ and the specific case when the slowest train is at the very back (this is because this is the easiest to see if there is recursion because we have 4 blank spaces where trains can be "inserted" which kind of seems like when $n = 4$. Now the main idea is to place the second slowest train and do casework on where it is. If the train B is the first train, there are 6 ways to organize the 3 remaining trains and the number of trains at the end will be 2 so $6 \cdot 2$. If B is the second left-most train, there are 6 ways to arrange the other trains and where all end up with 3 trains so $6 \cdot 3$. If B is the third train from the left, there are 3 options where there are 3 trains at the end and 3 options where there are 4 trains in the end (the reason for this is for the two trains in front of B, if the slower one is first, there would be 3 trains in the end but if the faster one is first, there would be 4 trains in the end and there is an equal number of ways for both of these to happen and thus, $3 \cdot 3 + 3 \cdot 4$. If BA are the last two trains in the list of 5, we do casework on where C the next slowest train is: if C is first, there are 2 ways that can happen and both end up with three trains at the end, if C is second from the left, there are two ways and both end up with 4 trains in the end, if C is the last train, there are two ways where one way has 4 trains at the end and one with 5 trains at the end. This totals up to $74 = 50 + 4!$. This leaves $200 = 50 \cdot 4$ for the other cases of where A is placed. However, doing the same thing for $n = 4$, we see that $17 = 11 + 3!$.

Another recursive formula (recursion for all previous expected value so a summation) is the following:

$$f(n+1) = \frac{1}{n+1} \cdot \left(\sum_{i=1}^{n} f(i) \right)$$

TI realized this worked when I randomly added all of the expected values trying to find a way to construct $n = 5$ by just adding all the previous ones and luckily I realized it differed by $120 = 5!$ which I then tested on the others and it worked as well. This also resulted in working. So now I have two formulas for $f(n+1)$.

This resulted in me noticing that $\frac{1}{n+1} \cdot \left(\sum_{i=1}^{n} f(i) \right) = f(n) + \frac{1}{n+1}$. I couldn't get more progress.

Thank you for taking your time to look at my solutions for the problem set! These problems were very fun to solve.