PROMYS

Tyler King

June 22, 2021

1 Question 1

Bashing the first few cases yields:
$1^3 + 5^3 + 3^3 = 153$
$16^3 + 50^3 + 33^3 = 165033$
$166^3 + 500^3 + 333^3 = 166500333$
A simple pattern can be noted fairly quickly, as each right hand side output is just the concatenation of each number (before getting cubed).

1.1 Initial Proof Based on Factoring

I quickly began by remembering the property that $x^3 + y^3 + z^3 - 3xyz$ was factorable to $(x + y + z)(x^2 + y^2 + z^2 - xy - yz - zx)$ and thus attempted to manipulate the expressions into this form. I begin by dictating the first case to be $n = 0$, the second case to be $n = 1$, and for each new case, it increments by 1. Therefore, a generalized form of the left-hand side for case n can be constructed:

$$(1\underbrace{66\cdots6}_{\text{n times}})^3 + (5\underbrace{00\cdots0}_{\text{n times}})^3 + (\underbrace{33\cdots3}_{\text{n+1 times}})^3.$$

I furthermore assign $x = 1\underbrace{66\cdots6}_{\text{n times}}$, $y = 5\underbrace{00\cdots0}_{\text{n times}}$, and $z = \underbrace{33\cdots3}_{\text{n+1 times}}$, allowing me to rewrite the generalized left-hand side as

$$x^3 + y^3 + z^3 - 3xyz + 3xyz$$

which then factors to

$$(x + y + z)(x^2 + y^2 + z^2 - xy - yz - zx) + 3xyz.$$

By noting that $x + y + z = 3z = \underbrace{99\cdots9}_{\text{n+1 times}}$ (this can be confirmed by noting that $x + y = 2z$ since the leading 1 in x and the leading 5 in y will always add to 6, and then all the 6s from x along with the leading 6 divided by 2 yield z), we can factor this value out. Additionally, we can regroup $x^2+y^2+z^2-xy-yz-zx = x(x - y) + y(y - z) + z(z - x)$, yielding

$$\underbrace{99\cdots9}_{\text{n+1 times}}(x(x - y) + y(y - z) + z(z - x) + xy).$$

By focusing only on the part of the expression inside the parentheses and substituting in the actual values for x, y, and z for the parts inside the nested parentheses, we obtain

$$x(1\underbrace{66\cdots6}_{\text{n times}}-5\underbrace{00\cdots0}_{\text{n times}})+y(5\underbrace{00\cdots0}_{\text{n times}}-\underbrace{33\cdots3}_{\text{n+1 times}})+z(\underbrace{33\cdots3}_{\text{n+1 times}}-1\underbrace{66\cdots6}_{\text{n times}})+xy$$

which can be rewritten as

$$x(-\underbrace{33\cdots3}_{\text{n times}}4)+y(1\underbrace{66\cdots6}_{\text{n-1 times}}7)+z(1\underbrace{66\cdots6}_{\text{n-1 times}}7)+xy.$$

Remark 1.1. *Note that for the case $n=0$, the above expression is equivalent to $x(-4)+y(2)+z(2)$.*

Again the expression can be rewritten:

$$x(2\cdot-1\underbrace{66\cdots6}_{\text{n-1 times}}7)+y(1\underbrace{66\cdots6}_{\text{n-1 times}}7)+z(1\underbrace{66\cdots6}_{\text{n-1 times}}7)+xy$$

$$2x(-1\underbrace{66\cdots6}_{\text{n-1 times}}7)+y(1\underbrace{66\cdots6}_{\text{n-1 times}}7)+z(1\underbrace{66\cdots6}_{\text{n-1 times}}7)+xy$$

$$1\underbrace{66\cdots6}_{\text{n-1 times}}7(y+z-2x)+xy.$$

Note that $y+z-2x=5\underbrace{00\cdots0}_{\text{n-1 times}}1=y+1$ and $1\underbrace{66\cdots6}_{\text{n-1 times}}7=x+1$, which means the expression thus becomes

$$(y+1)(x+1)+xy$$

$$2xy+x+y+1$$

$$2(1\underbrace{66\cdots6}_{\text{n times}})(5\underbrace{00\cdots0}_{\text{n times}})+1\underbrace{66\cdots6}_{\text{n times}}+5\underbrace{00\cdots0}_{\text{n times}}+1$$

Which is equivalent to:

$$2(8\overbrace{33\cdots3}^{\text{n-1 times}}\underbrace{00\cdots0}_{\text{n+1 times}})+\underbrace{66\cdots6}_{\text{n times}}7$$

$$1\overbrace{66\cdots6}^{\text{n times}}\underbrace{00\cdots0}_{\text{n+1 times}}+\underbrace{66\cdots6}_{\text{n times}}7$$

$$1\underbrace{66\cdots6}_{\text{2n times}}7.$$

Our final goal becomes proving that

$$(\underbrace{99\cdots9}_{\text{n+1 times}})(1\underbrace{66\cdots6}_{\text{2n times}}7)$$

is equivalent to the right hand side. We begin approaching this by noting that $\underbrace{99\cdots9}_{\text{n+1 times}}=10^{n+1}-1$ and thus our left-hand side becomes

$$10^{n+1}(1\underbrace{66\cdots6}_{\text{2n times}}7)-1\underbrace{66\cdots6}_{\text{2n times}}7$$

$$1\overbrace{66\cdots6}^{\text{2n times}}7\underbrace{00\cdots0}_{\text{n+1 times}} - 1\underbrace{66\cdots6}_{\text{2n times}}7.$$

Knowing our final goal on the right hand side, we cut the first term (which has $3n+3$ digits) into 3 parts each with length $n+1$ digits. Furthermore, we parse the second term (which has $2n+2$ digits), into 2 distinct parts, again each with $n+1$ digits. This yields

$$10^{2n+2}(1\underbrace{66\cdots6}_{\text{n times}}) + 10^{n+1}(\underbrace{66\cdots6}_{\text{n times}}7) + \underbrace{00\cdots0}_{\text{n times}} - \Big(10^{n+1}(1\underbrace{66\cdots6}_{\text{n times}}) + \underbrace{66\cdots6}_{\text{n times}}7\Big).$$

By grouping terms with similar number of trailing zeroes and simplifying, we obtain:

$$10^{2n+2}(1\underbrace{66\cdots6}_{\text{n times}}) + 10^{n+1}(5\underbrace{00\cdots0}_{\text{n-1 times}}1) - \underbrace{66\cdots6}_{\text{n times}}7.$$

Utilizing the distributive property yields

$$10^{2n+2}(1\underbrace{66\cdots6}_{\text{n times}}) + 10^{n+1}(5\underbrace{00\cdots0}_{\text{n times}}) + 10^{n+1} - \underbrace{66\cdots6}_{\text{n times}}7$$

$$10^{2n+2}(1\underbrace{66\cdots6}_{\text{n times}}) + 10^{n+1}(5\underbrace{00\cdots0}_{\text{n times}}) + 1\underbrace{00\cdots0}_{\text{n+1 times}} - \underbrace{66\cdots6}_{\text{n times}}7$$

$$10^{2n+2}(1\underbrace{66\cdots6}_{\text{n times}}) + 10^{n+1}(5\underbrace{00\cdots0}_{\text{n times}}) + \underbrace{33\cdots3}_{\text{n+1 times}}$$

or

$$1\overbrace{66\cdots6}^{\text{n times}}5\underbrace{00\cdots0}_{\text{n times}}\overbrace{33\cdots3}^{\text{n+1 times}},$$

which is equivalent to the examples on the right hand side, proving the pattern noticed at the beginning of the question.

2 Question 2

2.1 Properties of Sequence (x_n)

$x_1 = 1$
$x_2 = 2$
$x_3 = \frac{1}{2}$
$x_4 = 1$
$x_5 = 2$
$x_6 = \frac{1}{2}$

It can be quickly noted that the pattern is cyclic with period 3, and can be denoted by the expression:

$$x_n = \begin{cases} n & \text{if } n \equiv 0 \pmod 3 \\ m & \text{if } n \equiv 1 \pmod 3 \\ \frac{1}{nm} & \text{if } n \equiv 2 \pmod 3 \end{cases}$$

assuming $x_1 = n$ and $x_2 = m$. Therefore, the sequence $x_{n-1}x_nx_{n+1} = 1$, given real inputs for both x_1 and x_2, will always yield a sequence with periodicity

3. The only exceptions are when $n = 0$ or $m = 0$, as that sequence does not exist since $\frac{1}{nm}$ is undefined and when $n = 1$ and $m = 1$, as the sequence has a periodicity of 1.

2.2 Properties of Sequence (y_n)

$$y_{n-1}y_{n+1} + y_n = 1$$

$y_1 = 1$
$y_2 = 2$
$y_3 = -1$
$y_4 = 1$
$y_5 = 0$
$y_6 = 1$
$y_7 = k$
$y_8 = 1 - k$
$y_9 = 1$
$y_{10} = 0$
$y_{11} = 1$
where k is any real number. Therefore, we can state that the starting values of $y_1 = 1$ and $y_2 = 2$ eventually yield a sequence with periodicity 5.

Remark 2.1. *Note that I assume k is constant to define the sequence as cyclic, but technically k can be any real value.*

When we test this with inputs $y_1 = n$ and $y_2 = m$ where $n, m \in \mathbb{R}$, we obtain:
$y_1 = n$
$y_2 = m$
$y_3 = \frac{1-m}{n}$
$y_4 = \frac{n+m-1}{nm}$
$y_5 = \frac{1-n}{m}$
$y_6 = n$
$y_7 = m$
By this point it is fair to conclude that the sequence will continue to repeat since y_8 will be calculated in the same way that y_3 was. In this scenario, we can again conclude that the sequence has periodicity of 5, however now we are aware of certain caveats. There are two particular circumstances that must be considered with this sequence:
Circumstance 1: $n = 0$ or $m = 0$
If either of these statements were true, this would yield an undefined value for y_3 and y_4 or y_4 and y_5 depending on $n = 0$ or $m = 0$, respectively. The only exceptions are when $n = 0$ and $m = 1$ OR $m = 0$ and $n = 1$. In these two scenarios, the sequence follows circumstance 2.
Circumstance 2: $n = 1$, $m = 1$, or $n + m = 1$
While this does not change the periodicity of the sequence, it does displace the indices of the sequence that are periodic. For a standard sequence that does not follow circumstance 2, the periodicity occurs mod 5, but when one of either $n = 1$, $m = 1$, or $n + m = 1$ occurs, their indices are shifted up 4, 2, or 3 respectively. These can be confirmed by simply testing these three particular cases (one example can be observed earlier in 2.2 when $n = 1$) and noticing this pattern.

3 Question 3

By taking the initial pattern $t_{n+1} = 3^{t_n}$, we attempt to solve for a recurrence relation by utilizing the characteristic root technique. By taking $\log_3$ on both sides, we obtain

$$\log_3 t_{n+1} = t_n.$$

From here, we make the substitutions $t_n = r^n$ and $t_{n+1} = r^{n-1}$, yielding

$$\log_3 r^{n-1} = r^n.$$

3.1 Last Two Digits of $3^{3^{3^3}}$

However, I realized this would not yield an answer without a graphing calculator, so I decided to change my approach. Similar to question 2, I look for periodicity within the pattern of the last 2 digits for powers of 3. Writing out the first 20 terms (mod 100), it becomes apparent that a pattern can be noticed.

$3^0 = 01$
$3^1 = 03$
$3^2 = 09$
$3^3 = 27$
$3^4 = 81$
$3^5 = 43$
$3^6 = 29$
$3^7 = 87$
$3^8 = 61$
$3^9 = 83$
$3^{10} = 49$
$3^{11} = 47$
$3^{12} = 41$
$3^{13} = 23$
$3^{14} = 69$
$3^{15} = 07$
$3^{16} = 21$
$3^{17} = 63$
$3^{18} = 89$
$3^{19} = 67$
$3^{20} = 01$

By noting that the last 2 digits of 3^n follows a cyclic repetition with periodicity 20, we can conclude that the last two digits of $3^{3^{3^3}}$ can be found by taking 3^{3^3} (mod 20). This is equivalent to 3^{27} (mod 20) $= 3^{3\cdot 9}$ (mod 20) $= (3^3)^9$ (mod 20) $= 27^9$ (mod 20) $= 27^8 \cdot 27^1$ (mod 20) $= 27^8$ (mod 20) $\cdot 27^1$ (mod 20) by modular multiplication rules. To find 27^8 (mod 20), we use the property:

$$(A \cdot B) \bmod C = ((A \bmod C) \cdot (B \bmod C)) \bmod C.$$

Utilizing this technique, we obtain

$$27^2 \bmod 20 = ((27 \bmod 20) \cdot (27 \bmod 20)) \bmod 20 \equiv 9$$

$$27^4 \bmod 20 = ((27^2 \bmod 20) \cdot (27^2 \bmod 20)) \bmod 20 \equiv (9 \cdot 9) \bmod 20 \equiv 1$$

$$27^8 \bmod 20 = ((27^4 \bmod 20) \cdot (27^4 \bmod 20)) \bmod 20 \equiv (1 \cdot 1) \bmod 20 \equiv 1.$$

Therefore,

$$3^{3^3} (\text{mod } 20) = 27^8 (\text{mod } 20) \cdot 27^1 (\text{mod } 20) \equiv 1 \cdot 7 = 7.$$

Using the table of values written out earlier, we can conclude that the final two digits of $3^{3^{3^3}}$ are **87**.

3.2 Last Three Digits of $3^{3^{3^3}}$

Although the last technique can be utilized to solve this question, it creates a list that has a periodicity that is too large to be drawn out. I tested this by starting at 3^0 and continuously multiplying the last 3 digits by 3 until I arrived back at 001. However, the realization that this would be an unrealistic solution led me to the discovery of Euler's theorem:

$$a^{\varphi(n)} \equiv 1 \ (\text{mod } n).$$

Looking for the last 3 digits yields the equation

$$a^{\varphi(1000)} \equiv 1 (\text{mod } 1000)$$

where $\varphi(1000) = 1000(1-\frac{1}{2})(1-\frac{1}{5}) = 400$. If we assume $a = 3$, we can note that now we are now looking for n where $3^{3^3} \equiv n$ (mod 400) since each occurrence of 3^{400} will yield 1 (mod 1000). I initially began by splitting 27 into a sum of powers of 2 ($27 = 16 + 8 + 2 + 1$), but I then realized the powers were so low compared to the divisor that this was not necessary. Beginning with $3^{20} \equiv 1$ (mod 400), we can confirm that $3^{3^3} \equiv 3^7$ (mod 400). Therefore, $3^{3^3} \equiv 187$ (mod 400). Therefore, our final goal becomes observing n where

$$3^{187} \equiv n \ (\text{mod } 1000).$$

Again utilizing the modular multiplication technique from 3.1, we know that

$$3^{40} \bmod 1000 = ((3^{20} \bmod 1000) \cdot (3^{20} \bmod 1000)) \bmod 1000 \equiv 801$$

since $3^{20} \equiv 401$ (mod 1000). Repeating this pattern twice more yields

$$3^{80} \bmod 1000 = ((3^{40} \bmod 1000)^2 \bmod 1000 \equiv (801 \cdot 801) \bmod 1000 \equiv 601$$

$$3^{160} \bmod 1000 = ((3^{80} \bmod 1000)^2 \bmod 1000 \equiv (601 \cdot 601) \bmod 1000 \equiv 201$$

$$3^{180} \bmod 1000 = ((3^{160} \bmod 1000) \cdot (3^{20} \bmod 1000) \bmod 1000 \equiv 201 \cdot 401 = 601$$

$$3^{187} \bmod 1000 = ((3^{180} \bmod 1000) \cdot (3^7 \bmod 1000) \bmod 1000.$$

To find the final solution, we note that $3^7 \equiv 187$ (mod 1000).

$$((3^{180} \bmod 1000) \cdot (3^7 \bmod 1000)) \bmod 1000 = (601 \cdot 187) \bmod 1000 = 387.$$

Therefore the last 3 digits of $3^{3^{3^3}}$ are **387**.

3.3 Last 10 digits of t_k for all $k \geq 10$

I begin by attempting to form a more general form of this idea. I noticed that the last n digits for t_n and t_{n+1} seemed to be the same for smaller cases (i.e. $n = 2, 3$), so I attempted to prove this more general statement. If t_n and t_{n+1} share the same values for the last n digits, then this can be rephased to showing that $t_{n+1} - t_n \equiv 0 \pmod{10^n}$. Factoring out t_n yields $t_n(\frac{t_{n+1}}{t_n} - 1) \equiv 0 \pmod{10^n}$. Using the modular multiplication property outlined earlier, we know that

$$t_n(\frac{t_{n+1}}{t_n} - 1) \bmod 10^n = ((t_n \bmod 10^n) \cdot (\frac{t_{n+1}}{t_n} - 1 \bmod 10^n)) \bmod 10^n.$$

Therefore, our earlier proof becomes demonstrating that $\frac{t_{n+1}}{t_n} - 1 \equiv 0 \pmod{10^n}$ or $\frac{t_{n+1}}{t_n} \equiv 1 \pmod{10^n}$. Note that since $t_{n+1} = 3^{t_n}$, this is equivalent to $\frac{3^{t_n}}{3^{t_{n-1}}} \equiv 1 \pmod{10^n}$ or $3^{t_n - t_{n-1}} \equiv 1 \pmod{10^n}$. By Euler's theorem, we look to prove that $t_n - t_{n-1} \equiv 0 \pmod{\varphi(10^n)}$. Factoring out t_{n-1} yields $t_{n-1}(\frac{t_n}{t_{n-1}} - 1) \equiv 0 \pmod{\varphi(10^n)}$.

Remark 3.1. *I had originally planned to continue repeating this pattern until the calculations were significantly simpler (i.e. once $3(\frac{3^3}{3} - 1)$ was obtained, however, I realized this would not achieve the answer because the final answer would be mod $10^n \cdot (\frac{1}{2} \cdot \frac{4}{5})^n$ (since this pattern repeats exactly n times and 10^n is divisible by 5 exactly n times so the totient function continuously yields a number only divisible by 2 and 5 while being the multiple of enough 5s that it continuously multiplies the original number by the same two fractions). Since this modulo was significantly larger than the values I was working with, I attempted to approach the problem a different way.*

I began by assuming that t_{n-1} and t_n shared the last $n-1$ digits. Therefore, we know that $t_n - t_{n-1} = a \cdot 10^{n-1}$ for integer a. Taking both sides and putting them each to the power of 3 yields:

$$3^{t_n - t_{n-1}} = 3^{a \cdot 10^{n-1}}$$

$$\frac{3^{t_n}}{3^{t_{n-1}}} = 3^{a \cdot 10^{n-1}}$$

$$\frac{t_{n+1}}{t_n} = 3^{a \cdot 10^{n-1}}.$$

Recall earlier when we were attempting to prove

$$\frac{t_{n+1}}{t_n} - 1 \equiv 0 \pmod{10^n}.$$

Substituting in the expression we obtained modifies our proof to be

$$3^{a \cdot 10^{n-1}} - 1 \equiv 0 \pmod{10^n}$$

$$3^{a \cdot 10^{n-1}} \equiv 1 \pmod{10^n}.$$

Since $3^{a \cdot 10^{n-1}}$ and 10^n are relatively prime (since one is only divisible by 3 and the other divisible by only 2 and 5), we can apply Euler's theorem. This modifies our proof into showing that

$$a \cdot 10^{n-1} \equiv 0 \pmod{\varphi(10^n)}.$$

Note that for all $n \geq 1$, $\varphi(10^n) = (\frac{1}{2})(\frac{4}{5}) \cdot 10^n = 4 \cdot 10^{n-1}$. Therefore, our proof becomes looking for $a \equiv 0 \pmod 4$. However this ended me up in the same issue I had earlier, so I decided to backtrack again.

Starting again, I attempted to prove that $t_{n+1} - t_n \equiv 0 \pmod{2^n}$ and $t_{n+1} - t_n \equiv 0 \pmod{5^n}$. since $gcd(2^n, 5^n) = 1$, and $t_{n+1} - t_n \equiv 0 \pmod{2^n}$ is easy to prove since t_n will always be odd $\forall n$, I ended up only focusing on $t_{n+1} - t_n \equiv 0 \pmod{5^n}$. To prove this statement, I attempted to apply a proof by contradiction. I realized at this point I did the $\pmod{2^n}$ hastily and only proved the case $n = 1$, so I went back and redid it. I had been attempting to apply induction for a little while, and this gave me the perfect opportunity to do so. Proving one case allowed me to prove all future cases via induction and with the help of a modified version of Euler's theorem that I found on Wikipedia that states:

$$\text{if } x \equiv y \pmod{\varphi(n)}, \text{then } a^x \equiv a^y \pmod n.$$

Note that since $a = 3$ and 2^n and 5^n are all co-prime, $\varphi(2) = 1$, and $3^3 \equiv 3 \pmod 1$, we know that $3^{3^3} \equiv 3^3 \pmod 2$ or $t_2 \equiv t_1 \pmod 2$. Since $\varphi(2^n) = 2^{n-1}$, we can continue repeating the pattern observed earlier. This proves the first component that $t_{n+1} \equiv t_n \pmod{2^n}$ or $t_{n+1} - t_n \equiv 0 \pmod{2^n}$.

Moving on to the second component of this proof, I apply a similar approach. Again using the modified version of Euler's theorem and induction, we note that the first case $3^3 \equiv 3 \pmod 4$ can be proven easily. Since $\varphi(5) = 4$ (and more generally, $\varphi(5^n) = 4 \cdot 5^{n-1}$, we know that $3^{3^3} \equiv 3^3 \pmod 5$. However, to prove that $3^{3^3} \equiv 3^3 \pmod{4 \cdot 5}$ (and thus continue the proof by induction), I returned back to the proof for the $\pmod{2^n}$ case. Note that since $3^3 \equiv 3 \pmod 2$, we apply the modified version of Euler's theorem to prove $3^{3^3} \equiv 3^3 \pmod 4$. Since $gcd(4, 5^n) = 1$, we apply modular multiplication to obtain $3^{3^3} \equiv 3^3 \pmod{4 \cdot 5}$ or $t_2 \equiv t_1 \pmod{4 \cdot 5}$. For all $n \geq 2$, our prior proof for the $\pmod{2^n}$ case proves that t_n and t_{n+1} are equivalent $\pmod 4$ since 4 is a divisor of 2^n when $n \geq 2$. Since we already know that $\varphi(5^n) = 4 \cdot 5^{n-1}$ and $t_2 \equiv t_1 \pmod{4 \cdot 5}$, we can use induction to prove that $t_{n+1} \equiv t_n \pmod{5^n}$ or $t_{n+1} - t_n \equiv 0 \pmod{5^n}$.

Having proven $t_{n+1} - t_n \equiv 0 \pmod{2^n}$ and $t_{n+1} - t_n \equiv 0 \pmod{5^n}$, we can use modular multiplication (since $gcd(2^n, 5^n) = 1$) to prove:

$$t_{n+1} - t_n \equiv 0 \pmod{10^n}.$$

This means that the last 10 digits of t_k for $k \geq 10$ will always be the same by induction. Since t_{10} and t_{11} share the last 10 digits and any t_{k+1} and t_k will share the last k digits (and thus if $k \geq 10$, then they must at least share the same last 10 digits), concluding the proof.

4 Question 4

While visualizing the question, the most difficult step for me personally to understand was the intersection of the perpendicular bisector of DC meeting the perpendicular bisector of AB. Since a majority of the other steps were pretty clean-cut, I chose to focus on this step first. Approaching this question was reminiscent of the types of problems that appear in Physics, so I began by constructing a right triangle with hypotenuse AD. Furthermore, I called θ the supplementary angle to x. Since x is an obtuse angle (and thus $x > 90$), $\theta < 90$.

I assumed AD had length b (and therefore the length of BC was also b), so the height of the right triangle is $b\sin\theta$ and the width is $b\cos\theta$. I additionally assume AB to have length a.
From here, I apply coordinate geometry. I assume point A falls at $(0,0)$. This means that B falls at $(a,0)$, C falls at (a,b), and D falls at $(-b\cos\theta, b\sin\theta)$. We first find the slope of DC, which is equivalent to $\frac{b-b\sin\theta}{a+b\cos\theta}$. Therefore the perpendicular bisector to DC would have a slope of $\frac{a+b\cos\theta}{b\sin\theta-b}$. Furthermore, we know that this perpendicular bisector intersects the midpoint of DC (which falls at $(\frac{a-b\cos\theta}{2}, \frac{b+b\sin\theta}{2})$, and thus we can utilize these values along with the equation of a linear line ($y = mx + c$ where c is the y-intercept) to obtain the equation of the perpendicular bisector. First, however, we obtain the y-intercept:

$$\frac{b+b\sin\theta}{2} = \Big(\frac{a+b\cos\theta}{b\sin\theta-b}\Big)\Big(\frac{a-b\cos\theta}{2}\Big) + c.$$

Solving for the y-intercept yields

$$c = \frac{b+b\sin\theta}{2} - \Big(\frac{a+b\cos\theta}{b\sin\theta-b}\Big)\Big(\frac{a-b\cos\theta}{2}\Big)$$

$$c = \frac{(b+b\sin\theta)(b\sin\theta-b)}{2(b\sin\theta-b)} - \frac{(a+b\cos\theta)(a-b\cos\theta)}{2(b\sin\theta-b))}$$

$$c = \frac{-b^2+b^2\sin^2\theta-(a^2-b^2\cos^2\theta)}{2(b\sin\theta-b)}$$

$$c = \frac{-b^2+b^2(\sin^2\theta+\cos^2\theta)-a^2}{2(b\sin\theta-b)}$$

$$c = \frac{-\cancel{b^2}+\cancel{b^2}-a^2}{2(b\sin\theta-b)}$$

$$c = \frac{-a^2}{2(b\sin\theta-b)}.$$

Therefore, the equation of the perpendicular bisector is:

$$y = \frac{a+b\cos\theta}{b\sin\theta-b}x - \frac{a^2}{2(b\sin\theta-b)}.$$

We began by looking for inconsistencies within the diagram. My first assumption was that the perpendicular bisector to DC would intersect the perpendicular bisector of AB above line AB. To do so, we again utilize coordinate geometry, looking for the intersection of the lines of the two perpendicular bisectors. The equation of the line that perpendicularly bisects AB is $x = \frac{a}{2}$. Plugging this into the equation of the first perpendicular bisector yields

$$y = \Big(\frac{a+b\cos\theta}{b\sin\theta-b}\Big)\Big(\frac{a}{2}\Big) - \frac{a^2}{2(b\sin\theta-b)}$$

$$y = \cancel{\frac{a^2}{2(b\sin\theta-b)}} + \frac{ab\cos\theta}{2(b\sin\theta-b)} - \cancel{\frac{a^2}{2(b\sin\theta-b)}}.$$

$$y = \frac{a\cos\theta}{2(\sin\theta-1)}.$$

However, since $\theta < 90$ (as outlined earlier in the problem), this result will always be negative (since $\sin\theta < 1$ and $\cos\theta > 0$ for $\theta \in (0, 90)$). At this point, I realized pure analysis would not be enough to obtain an effective result, so I turned to drawing. My first idea was to visualize the extrema of x (i.e. when it was almost 90° and almost 180°). Although it was hard to tell with the case when x was near 90°, it was clear when x was near 180° that the the line segment DP never crossed along line segment AB. Although I drew my initial example by hand, I decided to recreate the example in Geogebra (Figure 1) for clarity and readability.

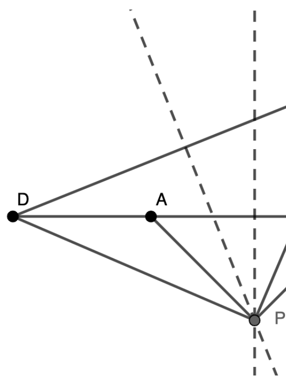


Figure 1: Angle x (or $\angle DAB$) when it is near 180°

From this point, I attempt to prove that the line DA will always intersect the extension of the line AB to a point left of A. Recall that, using the same coordinate points as earlier, we know that $A = (0, 0)$, $D = (-b\cos\theta, b\sin\theta)$, and $P = (\frac{a}{2}, \frac{a\cos\theta}{2(\sin\theta - 1)})$. If we start at point D and move towards point P, we can parameterize the vector $\overrightarrow{DP} = \left\langle -b\cos\theta + t, b\sin\theta + t\frac{\cos\theta}{\sin\theta - 1}\right\rangle$. Since the line AB is the x-axis, we are looking for when the y-component of the vector is equivalent to 0. Using the prior information, we can solve what value of t is necessary to find the intersection between the x-axis and DP:

$$b\sin\theta + \frac{t\cos\theta}{\sin\theta - 1} = 0$$

$$b\sin^2\theta - b\sin\theta + t\cos\theta = 0$$
$$b\sin^2\theta - b\sin\theta = -t\cos\theta$$
$$\frac{b\sin\theta - b\sin^2\theta}{\cos\theta} = t.$$

Having obtained the value of t for the intersection point, we now look to whether or not this point consistently falls to the left of $(0, 0)$ for acute angle x. To solve this, we plug it into the x-component of the vector, yielding:

$$-b\cos\theta + \frac{b\sin\theta - b\sin^2\theta}{\cos\theta}$$

$$\frac{-b\cos^2\theta - b\sin^2\theta + b\sin\theta}{\cos\theta}$$

$$\frac{-b(\cos^2\theta+\sin^2\theta)+b\sin\theta}{\cos\theta}$$
$$\frac{-b+b\sin\theta}{\cos\theta}$$
$$\frac{b(\sin\theta-1)}{\cos\theta}.$$

Recall that for $\theta \in (0, 90)$, this result will always be negative (since $\sin\theta < 1$ and $\cos\theta > 0$). Therefore, line segment DP will always intersect the extension of line AB to the left of point A. Since angle x is supplemental to θ, we conclude that the original proof proposed by the Journal of Irreproducible Results is incorrect because $\angle PAD - \angle PAB \neq \angle PBC - \angle PBA$ since $\angle PAD$ and $\angle PAB$ are adjacent angles instead of overlapping angles.

5 Question 5

I started with the realization that in the 3×3 scenario the primes that only occur once (such as 5 and 7) cannot fall in the same row or column. Without loss of generality, I assumed that these numbers fall along the diagonal. This led me to testing out certain values, and I eventually obtained

7	1	6
2	5	4
3	8	9

Remark 5.1. *While creating this, I felt that there were significant relationships between this question and Legendre's formula. However, I had trouble determining exactly how these two concepts were related (at least in regards to determining whether a number would have to fall on the main diagonal or not)*

I decided to begin testing out whether or not I could find a working scenario for the 5×5 case. I began by noting that all the primes that fell between $\left\lfloor \frac{n^2}{2} \right\rfloor$ and n^2 (noninclusive) must be non-adjacent and therefore fall on the "main diagonal". Furthermore, there are only 3 multiples of 7 (7, 14, 21), so therefore one of these 3 must additionally fall on the same "main diagonal". Without loss of generality, I make the following assumptions:

13	9	mn	10	5
18	17			20
m		19		
n	15		23	
25				7

where m and n represent arbitrary integers. Substituting in a few more numbers yields:

13	9	8	10	5
18	17	11	12	20
2	22	19		
4	15		23	
25	16			7

At this point, the only numbers left are 1, 3, 6, 14, 21, 24.

13	9	8	10	5
18	17	11	12	20
2	22	19	7	6
4	15	21	23	
25	16	1		14

I realized that the 7 in the bottom left would not yield a solution, so I attempted to switch around the 7 and 14 (the 21 cannot be in that position because then the multiples of 3 are no longer able to be split equally into the remaining rows and columns(since there are 8 multiples of 3). However, I realized that this substitution would not yield a solution either. I decided to go back to the first step, changing around some numbers to possibly obtain a better result. I realized after the fact that there were 22 occurrences of 2 across all numbers that could be found in the table (by Legendre's formula). Therefore, I believe that 14 could not have been placed along the main diagonal either since it would mean that the 2s could not be split equally into the remaining rows and columns. I formalize this concept in theorem 5.2 (which is used with respects to both the multiples of 3s and multiples of 2s).

Theorem 5.2. *Given that there exists some $n \times n$ grid filled with the numbers $1, 2, \ldots, n^2$, each prime $p < n^2$ that satisfies*

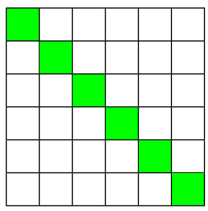
$$\sum_{i=1}^{\infty} \left\lfloor \frac{n^2}{p^i} \right\rfloor \equiv 1 (\text{mod } 2)$$

must not share a horizontal or vertical line along the grid with another prime that satisfies the aforementioned property.

Begin by taking the squares along the "main diagonal" which consist of primes that satisfy

$$\sum_{i=1}^{\infty} \left\lfloor \frac{n^2}{p^i} \right\rfloor \equiv 1 (\text{mod } 2)$$

and coloring them different shades of green (occasionally if the number of primes that satisfy this property is less than n, a prime number that does NOT satisfy this property squared will take a place along the "main diagonal"). For example, in the 6×6 case, this can be viewed as:



I realized at this point that this "theorem" was false by my initial testing on the 3×3 case. I then attempted to understand why my theorem was unsuccessful by studying the aforementioned case. Looking at the spread of the 2s yields the grid

		1
1		2
	3	

I was unable to discover any notable information from that grid, however, and thus decided the best approach for solving the 5×5 case was attempting to solve the 4×4 case and seeing if there were any trends or techniques that could be extrapolated.

5		6	3
	11		15
	12	13	
9	10		

This also proved to be a dead end, however, as I was unable to find a solution for the 4×4 case (one example of my failed efforts can be seen above). I then returned to the original 3×3 case, noticing that each square that doesn't fall on the main diagonal can be divided by the greatest common factor of itself and the square that is mirrored over the main diagonal. When doing this with the original 3×3 case, we obtain:

7	1	2
2	5	1
1	2	9

However, due to the nature of the multiples of 5 in the 5×5 case, it proved to be very difficult to get any meaningful numbers. Since the primes between $\left\lfloor \frac{5^2}{2} \right\rfloor$ and 5^2 (noninclusive) and one of the multiples of 7 could all be placed along the diagonal and there was a way to orient the multiples of 5 succesfully (for reference, note the orientation of the 5s I used in any of the 5×5 pictures), I assume that all the remaining numbers can be placed in the grid in such a way that the set of numbers obtained from multiplying the rows together and the set of numbers obtained from multiplying the columns together would be equivalent.

For the 11×11 case, we note that there are more than 11 primes between $\left\lfloor \frac{11^2}{2} \right\rfloor$ and 11^2 (61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113), which means that these primes cannot all be adjacent to each other. Since there must be some overlap, it is clear that there exists no 11×11 grid to fulfill the monkey's goals.

This concept can be extrapolated to the $n \times n$ case, as if the number of primes between $\left\lfloor \frac{n^2}{2} \right\rfloor$ and n^2 (noninclusive) is greater than n, then the $n \times n$ grid will not fulfill the monkey's desires.

Remark 5.3. *After a significant amount of testing, I eventually found a solution that would work for the* 5×5 *case:*

13	22	1	5	15
11	17	4	18	20
2	8	19	14	6
3	10	21	23	24
25	9	16	12	7

6 Question 6

6.1 $d = 7$

Having utilized Euler's theorem in question 3, I noticed its applicability again in this question. For something to be divisible by 7, it must additionally be equivalent to 0 (mod 7). However, since the equations we are looking for are $3^n - 1 \equiv 0 \pmod 7$ and $5^n - 1 \equiv 0 \pmod 7$, I decided to move the 1 to the right hand side (since 1 mod 7 is just 1), obtaining $3^n \equiv 1 \pmod 7$ and $5^n \equiv 1 \pmod 7$, which was a form reminiscent Euler's theorem. Knowing that

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

(when a, n are coprime) I allow $n = 7$ which leads to $\varphi(7) = 7(1 - \frac{1}{7}) = 6$. Going back to the original question, $(3^n - 1, 5^n - 1)$ are divisible by $d = 7$ when $n = 6$ by Euler's theorem. However, infinitely more cases of n can satisfy this equation by utilizing the property that $x^{ab} = (x^a)^b$. So long as n is a multiple of 6, $(3^n - 1, 5^n - 1)$ will be divisible by $d = 7$. Knowing that

$$a^6 \equiv 1 \pmod 7,$$

we substitute $a := a^n$ where $n \in \mathbb{Z}^{0+}$, yielding

$$(a^n)^6 \equiv 1 \pmod 7$$

$$a^{6n} \equiv 1 \pmod 7.$$

Therefore, as long as $n \equiv 0 \pmod 6$ and $n \in \mathbb{Z}^{0+}$, $(3^n - 1, 5^n - 1)$ are divisible by $d = 7$.

Remark 6.1. *Note that in all cases both prior and in future sections, $gcd(3^n, d) = 1$ and $gcd(5^n, d) = 1$. This allows us to apply Euler's theorem.*

6.2 $d = 11$

Again using Euler's theorem but with $n = 11$ (which leads to $\varphi(11) = 10$ since Euler's totient function is always $n - 1$ for primes) instead of $n = 7$, we obtain the general equation $a^{10} \equiv 1 \pmod{11}$. Going through the same steps in section 6.1, we note that, as long as $n \equiv 0 \pmod{10}$ and $n \in \mathbb{Z}^{0+}$, $(3^n - 1, 5^n - 1)$ are divisible by $d = 11$.

6.3 $d = 13$

The same technique can be applied to $d = 13$ that we applied to $d = 7$ and $d = 11$, which means that as long as $n \equiv 0 \pmod{12}$ and $n \in \mathbb{Z}^{0+}$, $(3^n - 1, 5^n - 1)$ will be divisible by $d = 13$.

6.4 $d = 77$

Given that 77 is a composite and not a prime, a slight deviation exists and thus we compute Euler's totient function: $\varphi(77) = 77(1 - \frac{1}{11})(1 - \frac{1}{7})$. Therefore, $a^{60} \equiv 1 \pmod{77}$ and thus if $n \equiv 0 \pmod{60}$ and $n \in \mathbb{Z}^{0+}$, $(3^n - 1, 5^n - 1)$ will be divisible by $d = 77$.

6.5 $d = 1001$

Similar to $d = 77$, we start by computing Euler's totient function: $\varphi(1001) = 1001(1 - \frac{1}{7})(1 - \frac{1}{11})(1 - \frac{1}{13})$. Therefore, $a^{720} \equiv 1 \pmod{1001}$ and thus if $n \equiv 0 \pmod{720}$ and $n \in \mathbb{Z}^{0+}$, $(3^n - 1, 5^n - 1)$ will be divisible by $d = 1001$.

When bashing certain cases, it felt like obtaining an n that satisfied the conditions of d seemed nearly impossible, but with Euler's theorem, it becomes significantly easier to deal with these questions and break them down into pieces.

7 Question 7

I started with the belief that every single real number in the form $\frac{n}{m}$ would be in set S as long as $\frac{n}{m} \in (0, 1]$. I attempted to prove that all values in the matrix (I chose to display it this way since it was the most similar to how I visualized the question initially)

$$\begin{bmatrix} \frac{1}{1} & \frac{2}{2} & \frac{3}{3} & \cdots & \frac{m}{m} \\ \frac{1}{2} & \frac{2}{3} & \frac{3}{4} & \cdots & \frac{m}{m+1} \\ \frac{1}{3} & \frac{2}{4} & \frac{3}{5} & \cdots & \frac{m}{m+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{2}{n+1} & \frac{3}{n+2} & \cdots & \frac{m}{m+n} \end{bmatrix}$$

would also be in set S. However, when attempting to prove this, I could not find any values below $\frac{1}{2}$. This led me to believe that there was a lower bound of around $\frac{1}{2}$ along with the aforementioned upper bound of 1.
I begin by first proving that $\frac{a+c}{b+d}$ will always be between the values of $\frac{a}{b}$ and $\frac{c}{d}$. Assume that $\frac{a}{b} \leq \frac{c}{d}$ and $mb = d$ for some constant m. Therefore, we are attempting to prove that

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

which yields

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

$$\frac{a}{b} \leq \frac{\frac{a+c}{m+1}}{b} \leq \frac{\frac{c}{m}}{b}.$$

Since $b > 0$, we can clear it from the inequality:

$$a \leq \frac{a+c}{m+1} \leq \frac{c}{m}.$$

Focusing on the two different inequalities means we are attempting to prove $a \leq \frac{a+c}{m+1}$ and $\frac{a+c}{m+1} \leq \frac{c}{m}$. By cross-multiplying, we obtain the inequalities $a(m+1) \leq a+c$ and $(a+c)m \leq (m+1)c$. Cleaning these up results in $am \leq c$ and $am \leq c$, which is known to always be true since we assumed $\frac{a}{b} \leq \frac{c}{mb}$ (which by cross-multiplication confirms $am \leq c$).
However, since all fractions will always be greater than $\frac{1}{2}$ and less than $\frac{1}{1}$ (this will be confirmed later), rule (iii) will only yield fractions with lower bounding $\frac{1}{2}$ and upper bounding $\frac{1}{1}$.

To prove the earlier statement, we begin by noting that $\frac{1}{1}$ becomes $\frac{1}{2}$ after running step (ii) once. Furthermore, every real number that can be expressed in the form $\frac{n}{m}$ where $n, m \in \mathbb{Z}^+$ and $\frac{1}{2} \leq \frac{n}{m} \leq 1$ will additionally be in set S by step (iii), as if we assume $n = x + y$, then $m = x + 2y$ where there are x occurrences of $\frac{1}{1}$ and y occurrences of $\frac{1}{2}$ (this is assuming that $n \geq 2$ as the cases where $n = 1$ that fulfill the outlined conditions have all been shown to be in set S). This helps prove the upper and lower bound restrictions, as if there were 0 occurrences of x, then $\frac{n}{m} = \frac{a}{2a}$ (for $a \in \mathbb{Z}^+$), and if there were 0 occurrences of y, then $\frac{n}{m} = \frac{a}{a}$ (for $a \in \mathbb{Z}^+$). Any other combination of x and y will result in $\frac{1}{2} \leq \frac{n}{m} \leq 1$ by the proof outlined earlier.

Therefore, all values in the set S are in the form $\frac{n}{m}$ where $n, m \in \mathbb{Z}^+$ and $\frac{1}{2} \leq \frac{n}{m} \leq 1$ and $n, m \in \mathbb{Z}^+$.

8 Question 8

Much like most other questions, I approached this by bashing out the first few cases.

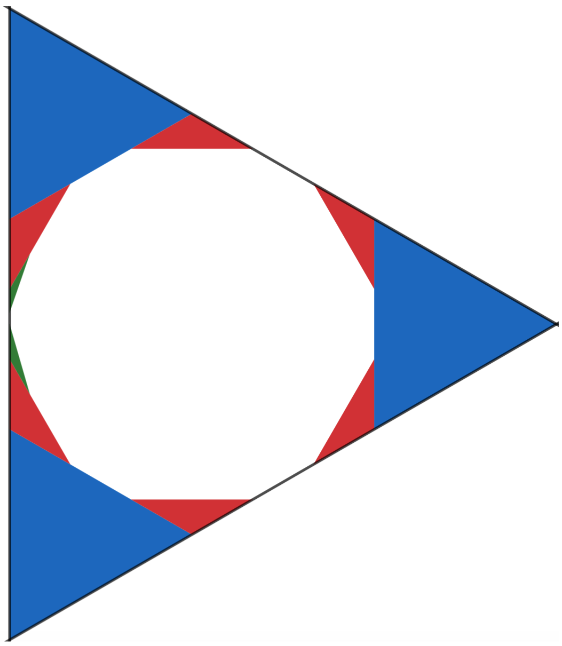


Figure 2: First 3 iterations of P_n and part of area removed for P_4. Area in blue was the area was removed to make P_1. Area in red was removed to make P_2. Area in green was 1/6th of the area removed to make P_3 (note that the green area can be multiplied by 6 and subtracted to find the area of P_3)

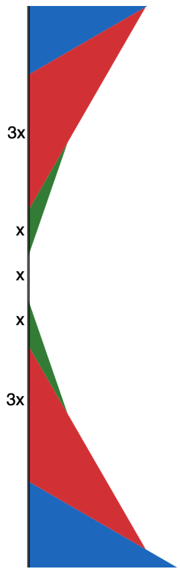
$P_0 = 10$.
The hexagon described in the question makes up $\frac{2}{3}$rds the area of the triangle (this can be confirmed by knowing that the hexagon is regular and that each

small triangle outside the hexagon is equivalent to $\frac{1}{6}$th the area of the hexagon) $P_1 = \frac{20}{3}$.

By subtracting out the 6 triangles whose area can be found by $\frac{1}{2}a^2 \sin 120$ (where $a : \frac{1}{3}x$ where x is the length of a side of the hexagon), we can obtain the area of P_2. Knowing that the original triangle had area 10 $\left(\text{and thus } \frac{1}{2}(3x)(\frac{3x\sqrt{3}}{2}) = 10 \text{ or } x^2 = \frac{40}{9\sqrt{3}}\right)$, we use $a^2 = \frac{1}{9}x^2$ to conclude that $a^2 = \frac{40}{81\sqrt{3}}$ and thus $\frac{1}{2}a^2 \sin 120 = \frac{10}{81}$. Since there are 6 triangles, the total area being subtracted out is $\frac{20}{27}$, resulting in

$P_2 = \frac{160}{27}$.

Solving for the area of P_3 takes a similar approach to solving the area of P_2. Note that there are now 12 triangles being cut off, but the dodecagon now has alternating side lengths. However, this means each one of the 12 triangles being cut off can still be calculated by the same equation.



Note that one of the side lengths of these green triangles is equivalent to $\frac{1}{3}a$ (or in the figure to the right, x where $a = 3x$). Since the side of the green triangles lying along side-length are equivalent to a third of the middle third of the side-length between the two blue triangles, they have side length $\frac{3x}{3}$. Therefore, the smaller side lengths of the green triangles are x or one third of the red triangle.

To calculate the other side length, we note that it is $\frac{1}{3}$rd of the length of the largest side of the red triangle, which is equivalent to $\sqrt{3}$ times the side length of one of the smaller side lengths (by the property of $30° - 30° - 120°$ triangles). Therefore, the larger side length of the green triangle (with respects to the figure on the right) is $x\sqrt{3}$. Furthermore, since the side lengths of this dodecagon are alternating, we know that the polygon is equiangular. Solving for the interior angles gives us 150°. Using this information, we are able to calculate the area of all 12 green triangles. Using the formula $\frac{1}{2}ab \sin C$ to calculate the area of the triangle (where a and b are adjacent sides and C is the angle between them), we obtain $\frac{1}{2}(x)(x\sqrt{3}) \sin 150$, which is equivalent to $\frac{1}{2}\left(\frac{1}{3}a\right)\left(\frac{\sqrt{3}}{3}a\right) \sin 150$ since $a = 3x$. Recalling that $a^2 = \frac{40}{81\sqrt{3}}$ allows us to remove the variable, yielding $\frac{1}{\cancel{2}}\left(\frac{\cancel{\sqrt{3}}}{9}\right)\left(\frac{\cancel{4}\cdot 10}{81\cancel{\sqrt{3}}}\right)\frac{1}{\cancel{2}}$ which simplifies to $\frac{10}{729}$. Since there are 12 triangles, this yields a final removal area of $\frac{40}{243}$. Therefore,

$P_3 = \frac{1400}{243}$.

As I was unable to find a clear pattern from the first few values of P_n, my next step was to find a closed formula for by proving that P_n was a sequence that is Δ^k-constant. To do so, I looked at the differences between the values.

$P_1 - P_0 = \frac{10}{3}$

$P_2 - P_1 = \frac{20}{27}$

$P_3 - P_2 = \frac{40}{243}$

Although the results prove that P_n is not Δ^k-constant, they yield a pattern where the difference between P_n and P_{n-1} is $\frac{2}{9}$ths of the difference between

P_{n+1} and P_n. Since this produces a convergent geometric series, we can evaluate the final area as $10 - \frac{\frac{10}{3}}{1-\frac{2}{9}}$. Simplifying this yields $10 - \frac{10(9)}{7(3)}$, which is equivalent to $10 - \frac{30}{7}$ or $\frac{40}{7}$. This led to my postulate that the area of P_∞ was $\frac{40}{7}$.

8.1 Proof that $P_\infty = \frac{40}{7}$

I begin by attempting to generalize the work I did in finding the area of P_3. For any 2 triangles where 1 is removed to help calculate P_n and the other is removed to help calculate P_{n+1}, we can note that these two triangles will always share an edge along the polygon formed by cutting out P_{n-1} where the two triangles that are separated are adjacent to each other. For example, in figure 2, we note that for each green triangle (P_3), it will share an edge with a red triangle (P_2) that was created by removing the blue triangles (P_1). If we focus on the point of intersection between the edge of P_{n-1} and the triangles being removed to form polygons P_n and P_{n+1}, we can generalize the area for the triangle being removed for P_{n+1} to be $\frac{1}{2}(x)(y)\sin\theta$ while the area being removed for a triangle of P_n is equivalent to $\frac{1}{2}(3x)(3y)\sin(180-\theta)$. Since $\sin\theta = \sin(180-\theta)$, each triangle being removed for P_n is $9 \cdot \frac{1}{2}(x)(y)\sin\theta$. Therefore, each triangle being removed for P_n is 9 times the size of each triangle being removed for P_{n+1}. However, the number of triangles being removed doubles. Therefore, the total area being removed for P_{n+1} is $\frac{2}{9}$ths of the total area being removed for P_n. This confirms the pattern noted earlier, and thus $P_\infty = \frac{40}{7}$.

8.2 Shape of P_∞

Since the hexagon is the last regular polygon, I postulate that the final shape of P_∞ will have hexagonal symmetry. Furthermore, since some parts of the polygon (such as the area between the green triangles in figure 2) are increasing in size less than other areas (such as if there were green triangles along a red triangle), I postulate that the final shape of P_∞ will be a rounded-off hexagon.

9 Question 9

I begin by writing out the cases for situation when there are 3 tries. I assume that trains A, B, and C exist such that $A > B > C$ with respects to their speed. Note that there are 3! possibilities, so I write these all out:
case 1: A, B, C 3 trains
case 2: A, C, B 2 trains
case 3: B, A, C 2 trains
case 4: B, C, A 2 trains
case 5: C, A, B 1 train
case 6: C, B, A 1 train
For these examples, the first train listed is the one in the front and the last train listed is the one in the back. The number of trains after all the linking happens is provided after the initial order of trains. It is clear that, when the slowest train is in the front, there will always be only 1 linked train after all the linking occurs since all the other trains are faster than the one in front, and will eventually link to the front train no matter what. Note that this will also

hold true for any position of the slowest train, as any trains that are behind the train with the slowest speed will always eventually link up to the slowest train. However, the train formed by these linkages will always be at the slowest speed, so they will never hit any of the other trains.

From here on, I refer to average number of all orderings of trains after a long period of time for n initial trains to be R_n. For example, given $n = 3$ initial trains (A, B, C), we can observe that $R_3 = \frac{3+2+2+2+1+1}{6} = \frac{11}{6}$.

Again assume there are n initial trains with the slowest train falls at some position m along the n trains. This means that there are $m - 1$ trains that are in front of the slowest train and $n - m$ trains that fall behind the slowest train. As we noted earlier, the trains that are behind the slowest train will all eventually join together to make a single train after all the linking occurs. The $m-1$ trains in front, however, all move at different speeds. The average number of orderings of trains after a long period of time of these $m-1$ trains can thus be represented by R_{m-1}. Therefore, the total average for the slowest train falling at some slot m with n initial trains is $R_{m-1} + 1$.

Now we attempt to find the average of all orders of n initial trains. To calculate this, we take the average of the average number of trains remaining for the slowest train falling at each position of the n initial trains. This yields the expression:

$$R_n = \frac{1}{n}\Big((R_0 + 1) + (R_1 + 1) + \cdots + (R_{n-2} + 1) + (R_{n-1} + 1)\Big)$$

which can be rewritten as

$$R_n = \frac{1}{n}\sum_{i=0}^{n-1}(R_i + 1).$$

Note that since there are n 1s being summed together, we can further simplify this by rewritting the expression as:

$$R_n = \frac{1}{n}\sum_{i=0}^{n-1} R_i + \frac{1}{n}n$$

$$R_n = 1 + \frac{1}{n}\sum_{i=0}^{n-1} R_i.$$

At this point I attempt to find a closed form without using recursion. Recalling the knowledge I learned from Discrete Mathematics in creating a closed form of recurrence relations, I attempt to apply these techniques here. However, this was not calculatable while the summation existed, so I attempted to find a way to cancel that out initially.

$$R_{n+1} - R_n = \cancel{1} + \frac{1}{n+1}\sum_{i=0}^{n} R_i - \left(\cancel{1} + \frac{1}{n}\sum_{i=0}^{n-1} R_i\right)$$

Note that since

$$\sum_{i=0}^{n} R_i = R_n + \sum_{i=0}^{n-1} R_i$$

by taking the last term of the sum and putting it outside the summation, we can make this substitution, yielding:

$$R_{n+1} - R_n = \frac{1}{n+1}\left(R_n + \sum_{i=0}^{n-1} R_i\right) - \frac{1}{n}\sum_{i=0}^{n-1} R_i$$

$$R_{n+1} - R_n = \frac{1}{n+1}R_n + \frac{1}{n+1}\sum_{i=0}^{n-1} R_i - \frac{1}{n}\sum_{i=0}^{n-1} R_i$$

$$R_{n+1} - R_n = \frac{1}{n+1}R_n + \left(\frac{1}{n+1} - \frac{1}{n}\right)\sum_{i=0}^{n-1} R_i$$

$$R_{n+1} - R_n = \frac{1}{n+1}R_n + \left(\frac{n}{n(n+1)} - \frac{n+1}{n(n+1)}\right)\sum_{i=0}^{n-1} R_i.$$

Recall that since

$$R_n = 1 + \frac{1}{n}\sum_{i=0}^{n-1} R_i$$

we can make this substitution in the right-hand side, yielding

$$R_{n+1} - R_n = \frac{1}{n+1}\left(1 + \frac{1}{n}\sum_{i=0}^{n-1} R_i\right) - \frac{1}{n(n+1)}\sum_{i=0}^{n-1} R_i$$

$$R_{n+1} - R_n = \frac{1}{n+1} + \cancel{\frac{1}{n(n+1)}\sum_{i=0}^{n-1}} R_i - \cancel{\frac{1}{n(n+1)}\sum_{i=0}^{n-1}} R_i$$

$$R_{n+1} - R_n = \frac{1}{n+1}.$$

Re-indexing by substituting $n-1$ for n yields

$$R_n - R_{n-1} = \frac{1}{n}.$$

Since the summation has been removed, it is now possible to calculate a closed form for this recurrence relation (which is equivalent to the one calculated earlier). Lets say we were attempting to find the value of R_n. This is equivalent to

$$R_n = (R_n - R_{n-1}) + (R_{n-1} - R_{n-2}) + \cdots + (R_2 - R_1) + (R_1 - R_0) + R_0.$$

Note that since $R_0 = 0$ (since there are 0 initial trains and thus 0 average final trains), we can rewrite the expression as

$$R_n = \frac{1}{n} + \frac{1}{n+1} + \cdots + \frac{1}{2} + \frac{1}{1}$$

$$R_n = \sum_{i=1}^{n} \frac{1}{i}.$$

At this point, we can now additionally evaluate R_4 and R_5. Knowing our general formula for n trains, we substitute in the values of $n = 4$ and $n = 5$.

$$R_4 = \sum_{i=1}^{4} \frac{1}{i} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = \frac{25}{12}$$

$$R_5 = \sum_{i=1}^{5} \frac{1}{i} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} = \frac{137}{60}$$

10 Question 10

Before I even began testing the configurations of 8 gems, I played a few thought experiments attempting to find any noticeable patterns. For these experiments I let X represent the smaller value between R and D and Y represent the larger value between R and D after every turn is made. The reason this can be done is the removal of gems are symmetric (i.e. there are no rules you can apply exclusively to diamonds that cannot be applied to rubies or vice versa).

I began with the case where $X = 1$ and $Y = 2$. Note that whoever is able to make this configuration first will automatically win. The reason being is the opponent can clear not win off this since they must either take 1 from X (which results in the player taking both from Y and winning), taking 1 from Y (which results in player taking 1 from both and winning), taking 2 from Y (which results in playing taking 1 from X and winning), or taking 1 from both (which results in the player taking 1 from Y and winning). With respects to the question, this means that whenever the starting configuration has either a 1 or a 2 for R or D (and the other value is greater than 2 or 1, respectively) or the difference between two numbers is 1 (and $X > 1$), the opponent will win the game since they can force an $X = 1$ and $Y = 2$ configuration in 1 move.

Now I look for what happens if initially one of the values is 3. Note that if $Y = 3$, then $X \leq 3$, and for all those combinations, the opponent will win (3-3 or 3-0 opponent removes all and wins immediately, 3-1 or 3-2 opponent forces an $X = 1$ and $Y = 2$ positioning). This means that when $X = 3$ the smallest position that isn't known to immediately end in an opponent's victory is R_{35} (since R_{34} has difference 1 and leads to R_{12}). However, this position results in an opponent losing. Any reduction of X or reduction of X and Y leads to X being either 0, 1, or 2, which allows the player to either finish the game or force an R_{12} board. However, if the opponent chooses to try to only reduce Y, this results in either R_{03}, R_{13}, R_{23}, R_{33}, or R_{34}, where the first and fourth configurations are known to lead to immediate victory while the other 3 allow the player to force an R_{12} board. However, if the starting configuration is instead $R_{3,5+n}$, $R_{3+n,5}$, or $R_{3+n,5+n}$, the opponent will instead win as they can force an R_{35} board in one move and win.

I proceeded to look for a case where $X = 4$ that shares the same properties as R_{35}. Obviously R_{46} would not be an acceptable alternative as that would immediately reduce to R_{35}, and thus be an $R_{3+n,5+n}$ configuration that results in an opponent victory when they go first. I thus attempted to increment Y by 1, obtaining the configuration R_{47} (note that the gap between X and Y was 1 more than the R_{35} starting configuration, which in turn was 1 greater than the R_{12} configuration). This can be shown to be a board that will always yield a

win for the player (as long as they are playing optimally):
Case 1: reducing Y with $Y > X$
The gap betwen X and Y decreases, leading to either $R_{3+n,5+n}$ or $R_{1+n,2+n}$. Thus, if the opponent tries to make the R_{47} into either R_{46} or R_{45} the player can counter by reducing it to an R_{35} or R_{12} configuration, allowing the player to win.
Case 2: reducing Y where $Y = X$
The player immediately reduces to an R_{00} board and wins.
Case 3: reducing Y where $Y < X$
R_{04} is a clear win for the player if the opponent makes this configuration while R_{14}, R_{24}, and R_{34}, all allow the player to force an R_{12} board.
Case 4: reducing X and Y
Note that since the gap between X and Y in the R_{47} configuration is larger than the gap in either the R_{12} or R_{35} cases, any equivalent reduction in X and Y will result in some value of X where the respective value of Y is higher than a configuration that is known to win for the player. Thus, when the player obtains such a configuration, all they must do is reduce the value of Y to one of these boards, allowing them to continue maintaining winning boards. For example, with the R_{47} case, the opponent can try to reduce this to something like R_{36}, but since the gap between X and Y is still 3, the player can reduce the Y to R_{35}, allowing the player to continue maintaining their winning position.
Case 5: reducing only X
Note that any of these moves that the opponent will make will result in an even higher Y value than in Case 4, which means that the exact same strategy in case 4 can be applied here.

Remark 10.1. *I attempted to make my statements in these cases as broad as possible so that they could be applied again later (particularly with respects to section 10.2).*

10.1 Initial Incorrect Solution

I noticed that both R_{35} and R_{47} followed a general pattern of $R_{n,2n-1}$, so I initially constructed a solution based around this concept. Although this seemed somewhat promising initially, I came to realize that there were significant issues with other cases. It all began unwinding with the starting configuration of R_{59}, as this configuration can clearly be reduced to R_{35} on the opponent's turn. Since this configuration is known to automatically lose for whoever starts with it, this means the starting configuration of R_{59} for the opponent would not yield a win for the player. Although I initially considered this to only be an edge case, I realized after further exploration ($R_{7,13}$ and $R_{11,21}$) that this pattern would not hold up.

10.2 Solution when Opponent Starts and Player Wins

While looking at the $R_{6,11}$ starting configuration, I realized that a reduction down to $R_{6,10}$ would allow the opponent to create a board that allows them to win. This conclusion helped spur on my proposed solution, which consists of initial configurations that result in the player's victory being part of set $\mathcal{P} = \{P_1, P_2, ..., P_n, ...\}$ where each $P_n = R_{m,m+n}$ and m is the smallest positive

integer value that has not already been used in some starting configuration P_a for $a \in [1, n-1]$. Writing out the first 5 terms of this sequence yields:
$P_1 = R_{12}$
$P_2 = R_{35}$
$P_3 = R_{47}$
$P_4 = R_{6,10}$
$P_5 = R_{8,13}$
This pattern can be proven by proving that no matter what move the opponent makes, the player can reduce the starting P_n board into some new board P_a where $a < n$ and $a \in \mathbb{Z}^+$. I attempt to replicate the casework outlined in the beginning of the problem to prove this result.

Definition 10.2. *For some boards P_n and P_a, $a < n$ and $a, n \in \mathbb{Z}^+$.*

Case 1: reducing Y with $Y > X$
Since P_n has a gap between X and Y of n, the operation outlined in the case will only decrease this gap. Since the value of X, Y of P_n is larger than the value of X, Y (respectively) for some board P_a, we can note that the player can easily subtract a constant from both X and Y from the board the opponent makes, forcing another board P_a that follows the same properties as the inital P_n board.
Case 2: reducing Y where $Y = X$
The player immediately reduces to an R_{00} board and wins.
Case 3: reducing Y where $Y < X$
For some initial configuration $P_n = R_{m,m+n}$, note that if Y is reduces to a value of $m - n$ or lower it is easy for the player to reduce the board into a P_a configuration. This is because all values lower than m exist in some P_a configuration already and even if $m - n$ takes the slot of X, the related Y value must be below $m - n + n$ or m (assuming an n gap between X and Y which couldn't be true for any P_a board). However, for boards where the reduction of the initial value of Y leads to a value greater than $m - n$, note that the gap between the original value of X and this new value of Y is less than $(m) - (m-n)$ or n, which means that if the opponent's makes a move that reduces Y to a value above $m - n$, the gap between the new values of X and Y will allow for the player to reduce the board into some P_a configuration.
Case 4: reducing X and Y
When the opponent reduces the initial values of X and Y for some P_n, note that the new value of X will always exist in a configuration of P_a. This means that any move the opponent makes that obeys the operation outlined in the case will result in a position that the player can reduce into P_a by reducing Y since the gap of P_n is greater than the gap of all P_a boards.
Case 5: reducing only X
Note that any of these moves that the opponent will make will result in an even higher Y value than in Case 4, which means that the exact same strategy in case 4 can be applied here.

Thus, we have proven that for any initial configuration (assuming the opponent goes first) that falls in set $\mathcal{P}$, the player has a guaranteed path to victory so long as they play optimally.

10.3 Starting configurations $R_{a,b}$ where $a + b = 8$

This can be done simply by testing the 9 possibilities when there are 8 total gems (where R denotes rubies and D denotes diamonds):
$R = 0$ and $D = 8$:
Opponent wins when opponent is first
Player wins when player is first
$R = 1$ and $D = 7$:
Opponent wins when opponent is first
Player wins when player is first
$R = 2$ and $D = 5$:
Opponent wins when opponent is first
Player wins when player is first
$R = 3$ and $D = 5$:
Player wins when opponent is first
Oppponent wins when player is first
$R = 4$ and $D = 4$:
Opponent wins when opponent is first
Player wins when player is first
Note that by symmetry you can switch the values of R and D and achieve the same result (i.e. the starting configuration is $R = 7$ and $D = 1$ yields the same outputs as the starting configuration $R = 1$ and $D = 7$).

Websites used:
https://www.khanacademy.org/computing/computer-science/cryptography/modarithmetic/a/fast-modular-exponentiation
https://en.wikipedia.org/wiki/Euler%27s_theorem
https://www.geogebra.org/