



Problem of the Month

Problem 0: September 2023

In this problem, f will always be a function defined by $f(r) = \frac{ar + b}{cr + d}$ where a , b , c , and d are integers. These integers will vary throughout the parts of the problem.

Given such a function f and a rational number r_1 , we can generate a sequence r_1, r_2, r_3, \dots by taking $r_n = f(r_{n-1})$ for each $n \geq 2$. That is, $r_2 = f(r_1)$, $r_3 = f(r_2)$, $r_4 = f(r_3)$, and so on. Unless there is some point in the sequence where $f(r_{n-1})$ is undefined, a sequence of this form can be made arbitrarily long.

These sequences behave in different ways depending on the function f and the starting value r_1 . This problem explores some those behaviours.

(a) Suppose $f(r) = \frac{2r - 1}{r + 2}$.

(i) With $r_1 = \frac{3}{2}$, compute r_2 , r_3 , and r_4 .

(ii) Find a rational number r_1 with the property that r_2 is defined, but r_3 is not defined.

(b) Suppose $f(r) = \frac{r + 3}{2r - 1}$.

(i) With $r_1 = \frac{3}{7}$, compute r_2 , r_3 , r_4 , and r_5 .

(ii) Determine all rational values of r_1 with the property that there is some integer $n \geq 1$ for which $f(r_n)$ is undefined. For all other values of r_1 , find simplified formulas for r_{2023} and r_{2024} in terms of r_1 .

(c) Suppose $f(r) = \frac{r + 2}{r + 1}$.

(i) With $r_1 = 1$, compute r_2 through r_9 . Write down decimal approximations of r_2 through r_9 (after computing them exactly).

(ii) Suppose r is a positive rational number. Prove that

$$\left| \frac{f(r) - \sqrt{2}}{r - \sqrt{2}} \right| = \left| \frac{1 - \sqrt{2}}{r + 1} \right|$$

(iii) Suppose r_1 is a positive rational number. Prove that $|r_n - \sqrt{2}| < \frac{1}{2^{n-1}} |r_1 - \sqrt{2}|$ for each $n \geq 2$. Use this result to convince yourself that as n gets large, r_n gets close to $\sqrt{2}$, regardless of the choice of the positive value r_1 . Can you modify f slightly so that the sequence always approaches $\sqrt{3}$?

(d) Explore the behaviour of the sequences generated by various values of r_1 for each of the functions below. Detailed solutions will not be provided, but a brief discussion will.

$$f(r) = \frac{r - 3}{r - 2}, \quad f(r) = \frac{r - 1}{5r + 3}, \quad f(r) = \frac{r - 1}{r + 2}, \quad f(r) = \frac{2r + 2}{3r + 3}, \quad f(r) = \frac{r + 1}{r - 2}$$



Problem of the Month

Problem 0: September 2023

Hint

- (a) $f(r)$ is undefined only when $r = -2$. For what value of r is $f(r) = -2$?
- (b) The sequence in part (i) is periodic. Can you show that the sequence is periodic for other values of r_1 ?
- (c) (ii) After substituting the expression for $f(r)$, multiply the numerator and denominator by $r + 1$. Try to find a common factor in the numerator and denominator.
- (iii) Use (ii) and the fact that when r is positive, $\left| \frac{1 - \sqrt{2}}{r + 1} \right| < \frac{1}{2}$. Try to establish the given inequality for a few small values of n and observe how knowing the inequality for n can help you to deduce it for $n + 1$.
- (d) Three of these sequences are periodic, one of them is constant (after the first term), and one of them always approaches the fixed value $\frac{3 - \sqrt{13}}{2}$ as long as there are no undefined values in the sequence.
-
-



Problem of the Month

Solution to Problem 0: September 2023

(a) (i)

$$\begin{aligned}r_2 = f(r_1) &= \frac{2r_1 - 1}{r_1 + 2} \\ &= \frac{2\left(\frac{3}{2}\right) - 1}{\frac{3}{2} + 2} \\ &= \frac{4}{7}\end{aligned}$$

$$\begin{aligned}r_3 = f(r_2) &= \frac{2\left(\frac{4}{7}\right) - 1}{\frac{4}{7} + 2} \\ &= \frac{1}{18}\end{aligned}$$

$$\begin{aligned}r_4 = f(r_3) &= \frac{2\left(\frac{1}{18}\right) - 1}{\frac{1}{18} + 2} \\ &= -\frac{16}{37}\end{aligned}$$

(ii) Observe that $f(r)$ is only undefined when $r = -2$, so to choose r_1 so that $r_3 = f(r_2)$ is undefined, we need $r_2 = -2$. Therefore, $r_2 = f(r_1) = -2$, which gives rise to the equation $\frac{2r_1 - 1}{r_1 + 2} = -2$. Multiplying both sides by $r_1 + 2$ gives $2r_1 - 1 = -2(r_1 + 2)$, which can be rearranged to get $4r_1 = -3$, so $r_1 = -\frac{3}{4}$. Indeed, if $r_1 = -\frac{3}{4}$, then $r_2 = -2$ and r_3 is undefined.

(b) (i)

$$\begin{aligned}r_2 = f(r_1) &= \frac{\frac{3}{7} + 3}{2\left(\frac{3}{7}\right) - 1} \\ &= -24.\end{aligned}$$

$$\begin{aligned}r_3 = f(r_2) &= \frac{-24 + 3}{2(-24) - 1} \\ &= \frac{-21}{-49} \\ &= \frac{3}{7} \\ &= r_1\end{aligned}$$

We have that $r_3 = r_1$, so this means $r_4 = f(r_3) = f(r_1) = r_2$, and using these facts, $r_5 = f(r_4) = f(r_2) = r_3 = r_1$. This will continue to get that $r_n = \frac{3}{7}$ when n is odd and $r_n = -24$ when n is even.

- (ii) First observe that $f(r)$ is undefined only when $2r - 1 = 0$, or $r = \frac{1}{2}$. Therefore, if $r_1 = \frac{1}{2}$, then $r_2 = f(r_1)$ is undefined.

Now suppose $f(r) = \frac{1}{2}$. Then $\frac{r+3}{2r-1} = \frac{1}{2}$, which can be rearranged to get the equation $2r+6 = 2r-1$. This equation implies $6 = -1$, which is nonsense, and so we conclude that there is no r such that $f(r) = \frac{1}{2}$.

The only way that the sequence can fail to be defined somewhere is if $r_n = \frac{1}{2}$ for some n . This happens if $r_1 = \frac{1}{2}$, but since $\frac{1}{2}$ is never the output of f , it is not possible that $r_n = \frac{1}{2}$ unless $n = 1$. Therefore, $r_1 = \frac{1}{2}$ is the only starting value for which the sequence is undefined somewhere.

Assuming $r \neq \frac{1}{2}$, we will now compute a general expression for $f(f(r))$. Since $r \neq \frac{1}{2}$ and $f(r) \neq \frac{1}{2}$ regardless of r , there will be no issues with the expression being undefined.

$$\begin{aligned} f(f(r)) &= \frac{\frac{r+3}{2r-1} + 3}{2\left(\frac{r+3}{2r-1}\right) - 1} \\ &= \frac{r+3 + 3(2r-1)}{2(r+3) - (2r-1)} && \text{(multiply through by } 2r-1\text{)} \\ &= \frac{7r}{7} \\ &= r \end{aligned}$$

and so we have that $f(f(r)) = r$ for all $r \neq \frac{1}{2}$. We can use this equation to get $r_3 = f(r_2) = f(f(r_1)) = r_1$. Next, we can compute $r_5 = f(r_4) = f(f(r_3)) = r_3 = r_1$. Continuing, we see that $r_n = r_1$ for all odd n . Similarly, $r_4 = f(r_3) = f(f(r_2)) = r_2$, and we can continue with this reasoning to get that $r_n = f(r_1) = r_2 = \frac{r_1+3}{2r_1-1}$ for all even n .

To answer the given question, since 2023 is odd, $r_{2023} = r_1$ and since 2024 is even, $r_{2024} = \frac{r_1+3}{2r_1-1}$.

- (c) (i) The values along with their decimal approximations are in the table below.

r_1	1	1
r_2	$\frac{3}{2}$	1.5
r_3	$\frac{7}{5}$	1.4
r_4	$\frac{17}{12}$	1.416667
r_5	$\frac{41}{29}$	1.413793
r_6	$\frac{99}{70}$	1.414286
r_7	$\frac{239}{169}$	1.414201
r_8	$\frac{577}{408}$	1.414216
r_9	$\frac{1393}{985}$	1.414213

(ii) We will work with the quantity $\frac{f(r) - \sqrt{2}}{r - \sqrt{2}}$ and worry about the absolute value later.

$$\begin{aligned}
\frac{f(r) - \sqrt{2}}{r - \sqrt{2}} &= \frac{\frac{r+2}{r+1} - \sqrt{2}}{r - \sqrt{2}} \\
&= \frac{r + 2 - \sqrt{2}(r + 1)}{(r + 1)(r - \sqrt{2})} && \text{(multiply through by } r + 1) \\
&= \frac{r - \sqrt{2} + 2 - \sqrt{2}r}{(r + 1)(r - \sqrt{2})} \\
&= \frac{r - \sqrt{2} - \sqrt{2}(r - \sqrt{2})}{(r + 1)(r - \sqrt{2})} \\
&= \frac{(r - \sqrt{2})(1 - \sqrt{2})}{(r + 1)(r - \sqrt{2})} \\
&= \frac{1 - \sqrt{2}}{r + 1} && \text{(cancel } r - \sqrt{2})
\end{aligned}$$

The result now follows by taking the absolute value of both sides.

(iii) Observe that if r is positive, then $f(r) = \frac{r+2}{r+1}$ is also positive since $r+2$ and $r+1$ are both positive. It follows that if r_1 is positive, then r_n is positive for all n . Since $r_n > 0$ for all n , $r_n + 1 > 1$ for all n , and taking reciprocals, we get $\frac{1}{r_n + 1} < 1$ for all n . Since $r_n + 1$ is positive, $\left| \frac{1}{r_n + 1} \right| = \frac{1}{r_n + 1}$, so we actually get that $\left| \frac{1}{r_n + 1} \right| < 1$ for all n .

We will now show that $|1 - \sqrt{2}| < \frac{1}{2}$ (you may already believe this to be true, but the proof presented does not assume that we have a known approximation of $\sqrt{2}$). To see this, observe that $8 < 9$, and so $\sqrt{8} < \sqrt{9}$ which is the same as $2\sqrt{2} < 3$. Dividing both sides by 2, we get $\sqrt{2} < \frac{3}{2}$. Subtracting $\frac{1}{2} + \sqrt{2}$ from both sides gives $-\frac{1}{2} < 1 - \sqrt{2}$. Now observe that $1 < 2$, so $\sqrt{1} < \sqrt{2}$ or $1 < \sqrt{2}$. Therefore, $1 - \sqrt{2} < 0$.

We have shown that $-\frac{1}{2} < 1 - \sqrt{2} < 0$, which implies that $|1 - \sqrt{2}| < \frac{1}{2}$. Combining this with $\left| \frac{1}{r_n + 1} \right| < 1$, we get

$$\left| \frac{1 - \sqrt{2}}{r_n + 1} \right| = \left| \frac{1}{r_n + 1} \right| |1 - \sqrt{2}| < 1 \times \frac{1}{2} = \frac{1}{2}$$

$$\text{so } \left| \frac{1 - \sqrt{2}}{r_n + 1} \right| < \frac{1}{2}.$$

Since $r_n = f(r_{n-1})$ for all $n \geq 2$, we can apply part (ii) to get

$$\left| \frac{r_n - \sqrt{2}}{r_{n-1} - \sqrt{2}} \right| = \left| \frac{f(r_{n-1}) - \sqrt{2}}{r_{n-1} - \sqrt{2}} \right| = \left| \frac{1 - \sqrt{2}}{r_{n-1} + 1} \right| < \frac{1}{2}$$

where the final inequality comes from applying what we showed above.

This implies that for all $n \geq 2$ we have

$$|r_n - \sqrt{2}| < \frac{1}{2}|r_{n-1} - \sqrt{2}| \quad (*)$$

Now let's return to the inequality in the question, which is $|r_n - \sqrt{2}| < \frac{1}{2^{n-1}}|r_1 - \sqrt{2}|$.

When $n = 2$, this inequality is $|r_2 - \sqrt{2}| < \frac{1}{2}|r_1 - \sqrt{2}|$, which is exactly $(*)$ when $n = 2$. We have already shown that $(*)$ is true for all n , so this means the desired inequality is true for $n = 2$.

When $n = 3$, we can apply $(*)$ to get $|r_3 - \sqrt{2}| < \frac{1}{2}|r_2 - \sqrt{2}|$, but we have just shown that $|r_2 - \sqrt{2}| < \frac{1}{2}|r_1 - \sqrt{2}|$. Therefore,

$$|r_3 - \sqrt{2}| < \frac{1}{2}|r_2 - \sqrt{2}| < \frac{1}{2} \left(\frac{1}{2}|r_1 - \sqrt{2}| \right) = \frac{1}{2^2}|r_1 - \sqrt{2}|$$

which shows that the desired inequality holds for $n = 3$.

By similar reasoning, we can use the fact that the inequality holds for $n = 3$ to prove that it holds for $n = 4$, then we can use that it holds for $n = 4$ to prove that it holds for $n = 5$, and so on to show that the inequality holds for all positive integers $n \geq 2$. We can formalize this using *mathematical induction*.

Assume that $k \geq 2$ is an integer for which the inequality $|r_k - \sqrt{2}| < \frac{1}{2^{k-1}}|r_1 - \sqrt{2}|$ is true. Using $(*)$ with $n = k + 1$, we have the following

$$|r_{k+1} - \sqrt{2}| < \frac{1}{2}|r_k - \sqrt{2}|$$

and now using the *inductive hypothesis*, that $|r_k - \sqrt{2}| < \frac{1}{2^{k-1}}|r_1 - \sqrt{2}|$, we get

$$|r_{k+1} - \sqrt{2}| < \frac{1}{2} \left(|r_k - \sqrt{2}| \right) < \frac{1}{2} \left(\frac{1}{2^{k-1}}|r_1 - \sqrt{2}| \right) = \frac{1}{2^k}|r_1 - \sqrt{2}|$$

but $k = (k + 1) - 1$, so we have shown that the inequality holds for the integer $k + 1$.

To summarize, we have shown that the inequality holds for $n = 2$, and we have shown that if the inequality holds for an integer, then it holds for the next integer. This shows that the inequality holds for all integers $n \geq 2$.

Finally, since $|r_1 - \sqrt{2}|$ is a fixed quantity, the quantity $\frac{1}{2^{n-1}}|r_1 - \sqrt{2}|$ must get closer and closer to 0 as n gets larger and larger. We also have that $|r_n - \sqrt{2}| < \frac{1}{2^{n-1}}|r_1 - \sqrt{2}|$ for all n , which means the quantity $|r_n - \sqrt{2}|$, which is positive, is also getting closer and closer to 0 as n gets larger and larger. It follows that $r_n - \sqrt{2}$ is very close to zero for very large n , which means r_n is very close to $\sqrt{2}$ for very large n . Keep in mind that this is true for *any* positive starting value r_1 .

(d) Here is a brief summary of the behaviours of the sequences.

When $f(r) = \frac{r-3}{r-2}$, the sequence will repeat with period 3 as long as at least three terms in the sequence are defined. The only values of r_1 for which the sequence has an undefined term are $r_1 = 2$ and $r_1 = 1$.

When $f(r) = \frac{r-1}{5r+3}$, the sequence will repeat with period 4 as long as at least four terms are defined. The only values of r_1 for which the sequence has an undefined term are $r_1 = -\frac{3}{5}$, $r_1 = -\frac{1}{5}$, and $r_1 = \frac{1}{5}$.

When $f(r) = \frac{r-1}{r+2}$, the sequence will repeat with period 6 as long as at least six terms are defined. The only values of r_1 for which the sequence has an undefined term are $r_1 = -2$, $r_1 = -1$, $r_1 = -\frac{1}{2}$, $r_1 = 0$, and $r_1 = 1$.

When $f(r) = \frac{2r+2}{3r+3}$, the sequence is undefined after r_1 if $r_1 = -1$. Otherwise, the sequence has $r_n = \frac{2}{3}$ for all $n \geq 2$, regardless of the value of r_1 .

Note: These examples, together with the one in part (b), show that the sequences can be periodic with period 1, 2, 3, 4, or 6. Do you think that any other periods are possible?

When $f(r) = \frac{r+1}{r-2}$, the sequence converges to $\frac{3-\sqrt{13}}{2}$ unless the sequence has an undefined term. To get an idea of why, try solving the equation $\frac{r+1}{r-2} = r$ for r (this can be rearranged to a quadratic equation in r). There are infinitely many values of r_1 for which r_n is undefined. The first few of them are

$$2, 5, \frac{11}{4}, \frac{26}{7}, \frac{59}{19}, \frac{137}{40}, \frac{314}{97}, \frac{725}{217}, \dots$$

Can you determine how these values are calculated? You might be interested in computing decimal approximations of these values and looking for a pattern.

As a final remark, we note that not all sequences are “well-behaved” (either periodic or approaching some value). For an example of a more chaotic sequence, try exploring the example in part (a) a little further. Can you see any pattern at all?



Problem of the Month

Problem 1: October 2023

Given a positive integer n , the *digit sum* of n is the sum of the base 10 digits of n . We will denote the digit sum of n by $D(n)$. For example, $D(1409) = 1 + 4 + 0 + 9 = 14$.

Suppose that m is a positive integer. We will call a list of consecutive positive integers

$$a, a + 1, a + 2, \dots, a + k$$

an m -list if none of $D(a)$, $D(a + 1)$, $D(a + 2)$, and so on up to $D(a + k)$ is a multiple of m . For example, the list 997, 998, 999, 1000, 1001, 1002 is a 4-list because the digit sums of the integers in the list are 25, 26, 27, 1, 2, and 3, respectively, none of which is a multiple of 4.

This problem explores the maximum length of an m -list for a few values of m .

- (a) Show that the maximum length of a 2-list is 2. To do this, you must show that there is a 2-list of length 2 and you must also show that no list of three or more consecutive positive integers can be a 2-list.
 - (b) Show that the maximum length of a 7-list is 12.
 - (c) Determine the maximum length of a 9-list.
 - (d) Determine the maximum length of an 11-list.
-
-



Problem of the Month

Problem 1: October 2023

Hint

For an integer a , think about how $D(a)$ and $D(a + 1)$ compare to each other. The relationship is most interesting when $a + 1$ is a multiple of 10.

Our solution will make use of some basic properties of remainders. For positive integers a and b , the remainder when dividing a by b can be found by first determining qb , the greatest multiple of b with $qb \leq a$, then computing the remainder as $r = a - qb$. The integer q can be found by computing $\frac{a}{b}$ and rounding down. For example, the remainder when dividing 38 by 5 is 3 because 35 is the greatest multiple of 5 that does not exceed 38 (in this case, $q = 7$), and so the remainder is $38 - 35 = 3$. We assume this idea is familiar, but here are a couple of things to think about and keep in mind.

- For integers a and b with $b > 0$, the remainder, r , when dividing a by b satisfies $0 \leq r < b$.
- For a fixed positive integer b , the remainders when dividing the positive integers $1, 2, 3, 4, 5, \dots$ cycle through the integers $1, 2, 3, \dots, b - 2, b - 1, 0$ repeatedly in that order. Importantly, if we think of 0 as coming “after” $b - 1$, the remainders for consecutive integers are consecutive.
- The remainder when dividing a by b is 0 exactly when a is a multiple of b .

If you have seen modular arithmetic before, it could be useful in writing a solution. Our solution will not use modular arithmetic, but we suggest reading about it anyway since it is quite useful.

Below are some specific hints for the given questions.

- (a) Show that if $D(a)$ and $D(a + 1)$ are both odd, then $a + 1$ is a multiple of 10.
 - (b) First, try to find the maximum length of a 7-list that does not contain a multiple of 10. For an integer a , how do the remainders when $D(a)$ and $D(a + 1)$ are divided by 7 relate to each other?
 - (c) By a rather famous divisibility rule, a positive integer is a multiple of 9 if and only if the sum of its digits is a multiple of 9.
 - (d) First show that an 11-list that starts with a multiple of 10 has length at most 29. Think about the remainders when $D(a)$ and $D(a + 1)$ are divided by 11.
-
-



Problem of the Month

Solution to Problem 1: October 2023

Before presenting solutions to the various parts of this question, we present a few general facts that will be used throughout the solutions. As well, we will often denote a list of integers enclosed in brackets. For example, we might write $[1, 2, 3, 4, 5]$ instead of $1, 2, 3, 4, 5$. This is to emphasize when we want to think of a list as a single object to which we can refer.

Fact 1: If n is a positive integer with units digit different from 9, then $D(n + 1) = D(n) + 1$.

Proof. Suppose that n is a k -digit positive integer with digits $A_1, A_2, \dots, A_{k-1}, A_k$, when read from left to right. That is, n is $A_1A_2 \cdots A_{k-1}A_k$. Then $D(n) = A_1 + A_2 + \cdots + A_{k-1} + A_k$. Since $A_k \neq 9$, no “carry” occurs when adding 1 to n . In other words, $n + 1$ is a k -digit positive integer with digits $A_1, A_2, \dots, A_{k-1}, A_k + 1$ when read from left to right. Therefore, $D(n + 1) = A_1 + A_2 + \cdots + A_{k-1} + (A_k + 1) = (A_1 + A_2 + \cdots + A_{k-1} + A_k) + 1 = D(n) + 1$. \square

Fact 2: Suppose that $L = [n, n + 1, n + 2, \dots, n + k]$ is a list of consecutive non-negative integers. If none of $n + 1, n + 2, \dots, n + k$ is a multiple of 10, then the integers $D(n), D(n + 1), \dots, D(n + k)$ are consecutive.

Proof. None of $n, n + 1, \dots, n + k - 1$ has a units digit of 9. This is because the alternative would imply that one of $n + 1, n + 2, \dots, n + k$ has a units digit of 0 and so is a multiple of 10. Therefore, we can repeatedly apply Fact 1 to get that $D(n + 1)$ is one more than $D(n)$, $D(n + 2)$ is one more than $D(n + 1)$, and so on up to $D(n + k)$ is one more than $D(n + k - 1)$. \square

Fact 3: Suppose m is an integer with $m > 1$ and that $[a, a + 1, a + 2, \dots, a + (m - 2)]$ is a list of $m - 1$ consecutive integers with the property that no integer in the list is a multiple of m . Then the remainder after a is divided by m is 1 and the remainder after $a + m - 2$ is divided by m is $m - 1$.

Proof. In the hint, it was stated that the remainders after dividing the positive integers $1, 2, 3, 4, 5, \dots$ by m are

$$1, 2, 3, \dots, m - 2, m - 1, 0, 1, 2, 3, \dots, m - 2, m - 1, 0, 1, 2, \dots$$

That is, the remainders cycle through the values $1, 2, 3, \dots, m - 2, m - 1, 0$ in that order.

Consider the list $[a, a + 1, a + 2, \dots, a + (m - 2)]$. Since none of the $m - 1$ integers in this list are multiples of m , none of the remainders after dividing by m can be 0, so the remainders must be $1, 2, 3, \dots, m - 2, m - 1$ in that order. In particular, the remainder after dividing a by m is 1 and the remainder after dividing $a + (m - 2)$ by m is $m - 1$. \square

For a positive integer n with units digit d , we say that n has k trailing d s if the rightmost k digits of n are d but the rightmost $k + 1$ digits are not equal to d . For example, the integer 123 has one trailing 3, 45000 has three trailing 0s, and 29199 has two trailing 9s.

By Fact 1, it is “usually” easy to predict $D(n + 1)$ from $D(n)$ since if the units digit of n is not 9, then $D(n + 1) = D(n) + 1$. However, if the units digit of n is 9, then adding 1 causes a carry to happen, so the digit sum can go down. For example, the digit sum of 2499 is $D(2499) = 2 + 4 + 9 + 9 = 24$, but the digit sum of $2499 + 1 = 2500$ is $D(2500) = 2 + 5 + 0 + 0 = 7$. Fact 4 establishes a relationship between $D(n)$ and $D(n + 1)$ when n has a units digit of 9.

Fact 4: Suppose n is a positive integer with t trailing 9s. Then $D(n + 1) - D(n) = 1 - 9t$.

Proof. Suppose n has t trailing 9s, which means we can assume that n takes the form $n = A_1A_2A_3 \cdots A_{r-1}A_r99 \cdots 99$ where the A_i are the first r digits and $A_r \neq 9$. Since A_r is a digit that is not 9, it must be less than 9. Therefore, when we add 1 to n , the first $r - 1$ digits do not change, the r^{th} digit increases by 1, and all of the trailing 9s become trailing 0s. In symbols, $n + 1$ is the integer $A_1A_2 \cdots A_{r-1}(A_r + 1)00 \cdots 00$.

Computing digit sums, $D(n) = A_1 + A_2 + \cdots + A_r + 9t$ and $D(n + 1) = A_1 + A_2 + \cdots + A_{r-1} + (A_r + 1)$. Subtracting, we get cancellation of A_1 through A_r and are left with $D(n + 1) - D(n) = 1 - 9t$. \square

Remark: We will repeatedly use the observation that if $L = [a, a + 1, a + 2, \dots, a + k]$ is an m -list for some m , then every sublist of L is also an m -list. One important consequence of this fact is that if we can show that there does not exist an m -list of length k , then we can conclude that there is no m -list that is longer than k . Thus, in general, if we want to prove that k is the maximum length of an m -list, it suffices to do the following:

- Produce an example of an m -list of length k .
 - Show that there does not exist an m -list of length $k + 1$.
- (a) The list $[9, 10]$ has $D(9) = 9$ and $D(10) = 1$, both of which are odd. Therefore, $[9, 10]$ is a 2-list of length 2. We will now show that a list of three consecutive positive integers cannot be a 2-list.

Suppose a is a positive integer with the property that $D(a)$ and $D(a + 1)$ are both odd. Since $D(a)$ is odd, $D(a) + 1$ is even. We are assuming that $D(a + 1)$ is odd, and since $D(a) + 1$ is even, we conclude that $D(a + 1) \neq D(a) + 1$.

By Fact 1, the units digit of a must be 9 (otherwise $D(a + 1) = D(a) + 1$), so the units digit of $a + 1$ is 0. Again by Fact 1, $D(a + 2) = D(a + 1) + 1$, and since $D(a + 1)$ is odd, $D(a + 1) + 1 = D(a + 2)$ is even.

Now consider the list $[a, a + 1, a + 2]$. If $D(a)$ and $D(a + 1)$ are both odd, we have shown that $D(a + 2)$ is even. This means it is impossible for $D(a)$, $D(a + 1)$, and $D(a + 2)$ to all be odd, so we conclude that a 2-list of length 3 cannot exist.

- (b) The arguments presented in this part and in part (d) rely on examining the location of multiples of 10 in lists of consecutive integers. For this part, because of Fact 2, seven consecutive integers will have seven consecutive digit sums unless a multiple of 10 causes a carry. To build a 7-list of length more than six, there needs to be a multiple of 10 somewhere in the middle to avoid ever having seven consecutive digit sums. You might be able to guess from this that the maximum length should be around 12.

Suppose $L = [a, a + 1, a + 2, \dots, a + 12]$ is an arbitrary list of 13 consecutive non-negative integers. We will show that L cannot be a 7-list and this will show that a 7-list cannot have length 13.

Since L contains 13 consecutive integers, it must contain at least one multiple of 10. Thus, there must be some k with $0 \leq k \leq 12$ such that $a + k$ is a multiple of 10.

Case 1: $0 \leq k \leq 6$.

Since $k \leq 6$, $a + k + 6 \leq a + 12$. The list L contains the integers a through $a + 12$, so we conclude that the seven integers $a + k, a + k + 1, \dots, a + k + 6$ are all in L . Moreover, the smallest multiple of 10 after $a + k$ is $a + k + 10$, so none of $a + k + 1, a + k + 2, \dots, a + k + 6$ is a multiple of 10 since they are all strictly between two consecutive multiples of 10.

By Fact 2, the integers $D(a + k), D(a + k + 1), \dots, D(a + k + 6)$ are consecutive. Any list of seven consecutive integers must contain a multiple of 7, so we conclude that L is not a 7-list.

Case 2: $7 \leq k \leq 12$.

Since $7 \leq k$, $a + 7 \leq a + k$ which can be rearranged to get $a \leq a + k - 7$. By similar reasoning to the beginning of Case 1, the seven integers $a + k - 7, a + k - 6, \dots, a + k - 1$ are all in L . As well, they are strictly between $a + k - 10$ and $a + k$, which are consecutive multiples of 10. Therefore, none of $a + k - 7, a + k - 6, \dots, a + k - 1$ is a multiple of 10.

By Fact 2, $D(a + k - 7), D(a + k - 6), \dots, D(a + k - 1)$ are consecutive, so one of them must be a multiple of 7. Therefore, L is not a 7-list.

Cases 1 and 2 address all possibilities for the location of a multiple of 10 in L . In both cases, L is not a 7-list. We conclude that a list of 13 consecutive non-negative integers cannot be a 7-list.

To help find a 7-list of length 12, we can modify the reasoning above slightly. To that end, suppose $M = [a, a + 1, a + 2, \dots, a + 11]$ is a 7-list. In the previous paragraphs, we essentially used the fact that if seven consecutive non-negative integers do *not* include a multiple of 10, then one of their digit sums must be a multiple of 7. Therefore, for M to be a 7-list, every sublist of seven-consecutive integers must contain a multiple of 10. We leave it to the reader to verify that this forces $a + 6$ to be a multiple of 10.

Since $a + 6$ is a multiple of 10, the closest multiples of 10 to $a + 6$ are $a - 4$ and $a + 16$. Therefore, the list $a, a + 1, \dots, a + 5$ does not contain a multiple of 10, and the only multiple of 10 in the list $a + 6, a + 7, \dots, a + 11$ is $a + 6$. By Fact 2,

$$D(a), D(a + 1), D(a + 2), D(a + 3), D(a + 4), D(a + 5)$$

and

$$D(a + 6), D(a + 7), D(a + 8), D(a + 9), D(a + 10), D(a + 11)$$

are lists of six consecutive integers. We are assuming M is a 7-list, so no integer in either of these lists is a multiple of 7. By Fact 3, the remainder after dividing $D(a + 5)$ by 7 is $7 - 1 = 6$. Also by Fact 3, the remainder after dividing $D(a + 6)$ by 7 is 1.

To summarize, if $M = [a, a + 1, \dots, a + 11]$ is a 7-list, then the following must be true.

- $a + 6$ is a multiple of 10.
- The remainder after dividing $D(a + 5)$ by 7 is 6.
- The remainder after dividing $D(a + 6)$ by 7 is 1.

The first condition implies that $a + 5$ has a units digit of 9, so suppose that $n + 5$ has t trailing 9s. By Fact 4, $D(n + 6) - D(n + 5) = 1 - 9t$. The second two conditions imply that there are integers m and n such that $D(a + 5) = 7n + 6$ and $D(a + 6) = 7m + 1$. Substituting into $D(a + 6) - D(a + 5) = 1 - 9t$, we obtain $(7n + 1) - (7m + 6) = 1 - 9t$ which can be rearranged to get $2t - 6 = 7(m - n - t)$. We know that t is a positive integer, and this shows that $2t - 6$ must be a multiple of 7. The smallest positive integer t with the property that $2t - 6$ is a multiple of 7 is $t = 3$ since $2(3) - 6 = 0$, which is a multiple of 7.

We are looking for an integer a such that $a + 5$ ends in three 9s and has the property that $D(n + 5)$ is six more than a multiple of 7. Thus, $a + 5$ will have the form $x999$ where x is some string of digits. In fact, $D(999) = 9 + 9 + 9 = 27$ has a remainder of 6 when divided by 7, so we set $a + 5 = 999$ or $a = 994$. Indeed, if we set

$$M = [994, 995, 996, 997, 998, 999, 1000, 1001, 1002, 1003, 1004, 1005]$$

the digit sums of the integers in M are 22, 23, 24, 25, 26, 27, 1, 2, 3, 4, 5, 6, none of which is a multiple of 7. Therefore, M is a 7-list of length 12.

This is actually the earliest 7-list of length 12. Others can be found by prepending digits to 994 in such a way that the number of trailing 9s of $a + 5$ does not change, and the remainder after dividing the digit sum by 7 does not change. This can be done by prepending any digits that have a sum that is a multiple of 7, provided the rightmost new digit is not a 9.

For example, we could take a to be 7994, 16994, 25994, 662994, and so on.

- (c) Following the reasoning from (b), one might think that we can find a 9-list of length $8 + 8 = 16$. Specifically, we might expect that we can find a list of the form $[a, a + 1, \dots, a + 15]$ where $a + 8$ is a multiple of 10 so that there are never nine consecutive integers without a multiple of 10. In fact, the maximum length of a 9-list is eight, so the approach in part (b) cannot be applied here. You might like to try to follow the reasoning from part (b) through to see exactly what goes wrong. The proof given here uses the well-known fact that a positive integer n is a multiple of 9 if and only if $D(n)$ is a multiple of 9.

A list of nine consecutive integers must contain a multiple of 9, so it must contain an integer whose digit sum is a multiple of 9. Therefore, a list of 9 consecutive integers cannot be a 9-list.

Any list of eight consecutive integers, none of which is a multiple of 9, must be a 9-list since none of the digit sums are multiples of 9. An example of a 9-list of length 8 is $[1, 2, 3, 4, 5, 6, 7, 8]$.

- (d) It might come as a surprise, but the maximum length of an 11-list is 38.

To get an idea of how one might guess this, we will first determine the maximum length of an 11-list that starts with a multiple of 10.

Suppose $L = [a, a + 1, a + 2, \dots, a + 19]$ is an 11-list of length 20 with the property that a is a multiple of 10. By Fact 2, $D(a), D(a + 1), \dots, D(a + 9)$ and $D(a + 10), D(a + 11), \dots, D(a + 19)$ are lists of ten consecutive integers. Since we are assuming that L is an 11-list, no integer in either of these lists is a multiple of 11. Therefore, by Fact 3, the remainder after $D(a + 9)$ is divided by 11 is 10 and the remainder after $D(a + 10)$ is divided

by 11 is 1. This implies the existence of integers m and n such that $D(a + 9) = 11m + 10$ and $D(a + 10) = 11n + 1$.

Since a is a multiple of 10, the units digit of $a + 9$ is 9. Suppose $a + 9$ has t trailing 9s. By Fact 4, $D(a + 10) - D(a + 9) = 1 - 9t$. Substituting $D(a + 9) = 11m + 10$ and $D(a + 10) = 11n + 1$ into this equation, we get

$$(11n + 1) - (11m + 10) = 1 - 9t$$

which can be rearranged to get $9t = 11(m - n) + 10$. This means t is a positive integer with the property that $9t$ is ten more than a multiple of 11. It can be checked that $t = 6$ is the smallest positive integer with this property.

We have shown the following: if $[a, a + 1, a + 2, \dots, a + 19]$ is an 11-list of length 20 with the property that a is a multiple of 10, then $a + 9$ has at least six trailing 9s. We can use this fact to prove the following:

- (i) An 11-list that starts with a multiple of 10 cannot have length 30.
- (ii) An 11-list cannot have length 39.

To prove (i), assume that $L = [a, a + 1, a + 2, \dots, a + 29]$ is an 11-list of length 30 and that a is a multiple of 10. We will show that this implies something that is plainly false, which will allow us to conclude that such an 11-list cannot exist.

The fact that L is an 11-list implies that the lists $M = [a, a + 1, \dots, a + 19]$ and $N = [a + 10, a + 11, \dots, a + 29]$ are both 11-lists. Moreover, a is a multiple of 10, so $a + 10$ is a multiple of 10. Thus, M and N are both 11-lists of length 20 that start with a multiple of 10. By what was established earlier, both $a + 9$ and $a + 10 + 9 = a + 19$ have at least six trailing 9s.

The difference between two integers with at least six trailing 9s must have at least six trailing 0s. However, $(a + 19) - (a + 9) = 10$ has only one trailing 0. We have shown that if an 11-list starts with a multiple of 10 and has length 30, then 10 has at least six trailing 0s. Since 10 has only one trailing 0, we conclude that it must be impossible for an 11-list to have length 30 and start with a multiple of 10. **Note:** The technique just used is known as a *proof by contradiction*.

To prove (ii), we can apply (i). Suppose $L = [a, a + 1, a + 2, \dots, a + 38]$ is a list of 39 consecutive positive integers. Among the first ten integers in L , there must be a multiple of 10. Suppose k is an integer with $0 \leq k \leq 9$ such that $a + k$ is a multiple of 10. Observe that $a + k + 29 \leq a + 9 + 29 = a + 38$, which means the list $M = [a + k, a + k + 1, \dots, a + k + 29]$ is completely contained in L . The list M has length 30 and starts with a multiple of 10, so it cannot be an 11-list by (i). Since L contains a sub-list that is not an 11-list, it cannot be an 11-list. Therefore, an 11-list cannot have length 39.

We can now use what we have done to find an 11-list of length 38. Suppose that $L = [a, a + 1, a + 2, \dots, a + 37]$ is an 11-list. There must be an integer k with $0 \leq k \leq 9$ such that $a + k$ is a multiple of 10. If $0 \leq k \leq 8$, then $a + k + 29 \leq a + 37$, so $[a + k, a + k + 1, \dots, a + k + 29]$ is entirely contained in L . This would imply the existence of an 11-list of length 30 that starts with a multiple of 10, which is impossible by (i). Since $0 \leq k \leq 8$ is impossible, we conclude that $k = 9$.

We assumed that $L = [a, a + 1, a + 2, \dots, a + 37]$ was an 11-list and deduced that $a + 9$ is a multiple of 10. The list $[a + 9, a + 10, \dots, a + 28]$ is completely contained in L , has length 20, and starts with a multiple of 10. Therefore, from the argument at the beginning of the solution to this part, $a + 9 + 9 = a + 18$ has at least six trailing 9s. As well, Since $a + 9$ is a multiple of 10, we can apply Fact 2 to $a + 9, a + 10, \dots, a + 18$ to get that $D(a + 9), D(a + 10), \dots, D(a + 18)$ is a list of 10 consecutive integers. Since L is an 11-list, none of them is a multiple of 11. By Fact 3, the remainder after $D(a + 18)$ is divided by 11 is 10.

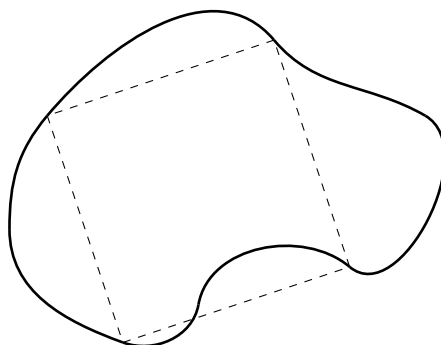
Therefore, we will look for an integer a with the property that $a + 18$ has 6 trailing 9s and $D(a + 18)$ is 10 more than a multiple of 11. The integer 999999 has $D(999999) = 54$ which is 10 more than a multiple of 11, so we guess that $a + 18 = 999999$, or $a = 999981$. It can be checked that the list $[999981, 999982, \dots, 1000018]$ is an 11-list of length 38. It also follows from our reasoning that this is the first such 11-list.



Problem of the Month

Problem 2: November 2023

This month's problem is based on the following general question: If you draw a "loop" in the Cartesian plane, is it always possible to find four points on that loop that are the vertices of a square? For example, the diagram below has a loop (the solid line) and a square drawn (the dashed line) with its four vertices on the loop.



Although it is a bit informal, it should be sufficient to think of a "loop" as a curve that you could draw by starting your pencil somewhere on a page and moving the pencil around the page eventually ending up where it started. Such a loop could be "smooth" (like a circle), "jagged" (like a polygon), or some combination of the two.

(a) In each of parts (i) through (v), find four points on the loop that are the vertices of a square.

(i) the circle with equation $x^2 + y^2 = 1$

(ii) the ellipse with equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ where a and b are fixed positive real numbers

(iii) the polygon with vertices $(1, 0)$, $(\frac{1}{2}, \frac{\sqrt{3}}{2})$, $(-\frac{1}{2}, \frac{\sqrt{3}}{2})$, $(-1, 0)$, $(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$, and $(\frac{1}{2}, -\frac{\sqrt{3}}{2})$

(iv) the boundary of the region enclosed by the parabola with equation $y = -\frac{1}{2}x^2 + \frac{1}{6}x + \frac{16}{9}$ and the line with equation $y = x$

(v) the boundary of the region enclosed by the parabolas with equations $y = x^2 + \frac{2}{3}x - \frac{4}{3}$ and $y = -x^2 + \frac{2}{3}x + \frac{4}{3}$

(b) Show that for every acute triangle there are exactly three squares whose vertices all lie on the perimeter of the triangle.



Problem of the Month

Problem 2: November 2023

Hint

- (a) (i) Because of the rotational symmetry of circles, there are lots of squares with their vertices lying on the circle with equation $x^2 + y^2 = 1$. In fact, given any point on the circle, there is a square with that point as one of its vertices and the other three vertices somewhere else on the circle. It might be useful for part (v) to spend some time thinking about this.
- (ii) Try looking for a square with the property that all four of its sides are parallel to either the x -axis or the y -axis. What are the coordinates of a square centred at the origin with its sides parallel to the axes?
- (iii) It will be helpful to have an accurate sketch of the hexagon. Similar to part (ii), there is a square with its sides parallel to the axes.
- (iv) After sketching the region, it shouldn't be hard to believe that there is a square with one of its sides on the line with equation $y = x$. The opposite side will lie on a line with the same slope.
- (v) The figure has 180° rotational symmetry about the origin, so try looking for a square that is centred at the origin. If such a square exists, it should also have 108° rotational symmetry about the origin. Observations from part (i) might be useful when thinking about such squares.
- (b) In our solution, we found it useful to coordinatize and assume that one the sides of the the triangle is on the x -axis in such a way that one of the vertices is at the origin.
-
-



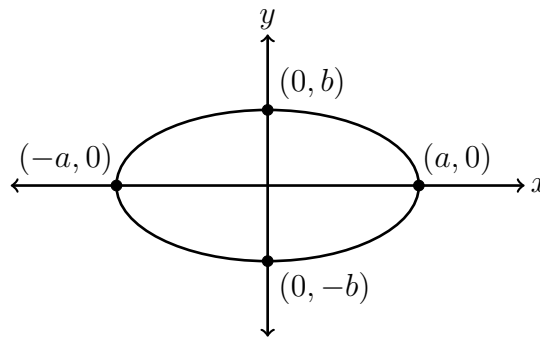
Problem of the Month

Solution to Problem 2: November 2023

- (a) (i) Because a circle has so much symmetry, there are many squares whose four vertices all lie on the circle. For example, the points $(1, 0)$, $(0, 1)$, $(-1, 0)$, and $(0, -1)$ are the vertices of a square and all lie on the circle with equation $x^2 + y^2 = 1$. As well, the four points $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$, and $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ are the vertices of a square and all lie on the circle.

Note: In general, if we start with a point P and rotate P repeatedly by 90° counterclockwise about the origin, the four points obtained are always the vertices of a square. Starting with the point with coordinates (a, b) , the coordinates of the point obtained by a rotation of 90° counterclockwise about the origin are $(-b, a)$. If we rotate 90° counterclockwise two more times, we get the points with coordinates $(-a, -b)$ and $(b, -a)$. Thus, for any real numbers a and b , the points with coordinates (a, b) , $(-b, a)$, $(-a, -b)$, and $(b, -a)$ are the vertices of a square. This will be useful in later parts. In addition, if (a, b) is a point on the circle with equation $x^2 + y^2 = 1$, then $a^2 + b^2 = 1$. Then since $(\pm a)^2 + (\pm b)^2 = a^2 + b^2 = 1$, the four points (a, b) , $(-b, a)$, $(-a, -b)$, and $(b, -a)$, which are the vertices of a square, all lie on the circle.

- (ii) Observe that the four points $(a, 0)$, $(0, b)$, $(-a, 0)$, and $(0, -b)$ are on the ellipse with equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, as shown in the illustration below.



Because $x^2 = (-x)^2$ and $y^2 = (-y)^2$, if (s, t) is a point on the ellipse, then $(s, -t)$ and $(-s, t)$ are both on the ellipse. Therefore, the ellipse has reflective symmetry in the x -axis and the y -axis, which is also evident from the diagram. It is reasonable to expect there to be a square inscribed in the ellipse that also has reflective symmetry in the two axes. The vertices of such a square should be of the form (t, t) , $(-t, t)$, $(-t, -t)$, and $(t, -t)$ for some real number $t \neq 0$.

Notice that the point (t, t) is on the line with equation $y = x$, so assuming such a square exists, we should be able to find one of its vertices by solving the system of

equations

$$y = x$$
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Substituting the first equation into the second and finding a common denominator gives

$$\frac{b^2x^2 + a^2x^2}{a^2b^2} = 1$$

which rearranges to give

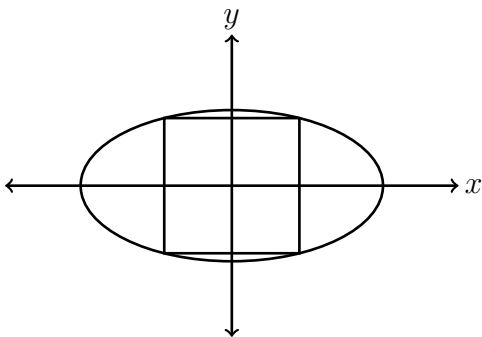
$$x^2 = \frac{a^2b^2}{a^2 + b^2}.$$

Since a and b are positive, $\sqrt{a^2b^2} = ab$, and so we get that $x = \frac{ab}{\sqrt{a^2 + b^2}}$. We therefore expect that the four points

$$\left(\frac{ab}{\sqrt{a^2 + b^2}}, \frac{ab}{\sqrt{a^2 + b^2}} \right), \left(-\frac{ab}{\sqrt{a^2 + b^2}}, \frac{ab}{\sqrt{a^2 + b^2}} \right)$$
$$\left(-\frac{ab}{\sqrt{a^2 + b^2}}, -\frac{ab}{\sqrt{a^2 + b^2}} \right), \left(\frac{ab}{\sqrt{a^2 + b^2}}, -\frac{ab}{\sqrt{a^2 + b^2}} \right)$$

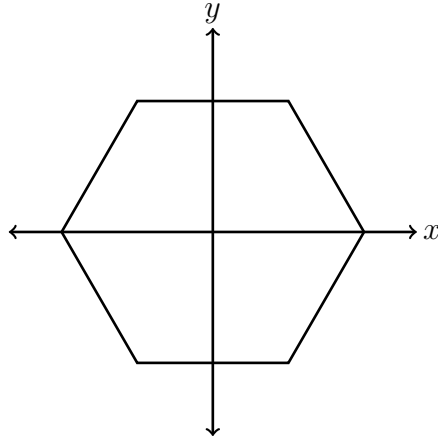
are all on the ellipse and are the vertices of a square. As discussed in the solution to part (i), these points are of the form (t, t) , $(-t, t)$, $(-t, -t)$, and $(t, -t)$, so they are the vertices of a square. It is an exercise to verify that they are all on the ellipse.

Below is a picture of the ellipse with a square inscribed.



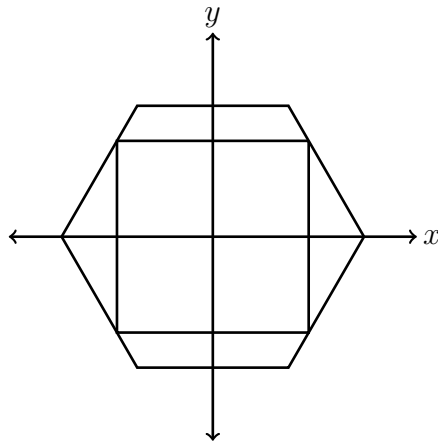
Follow up question: If $a = b$, then the ellipse is a circle and there are infinitely many inscribed squares. If $a \neq b$, is the square in the solution above the only square inscribed in the ellipse?

(iii) In this case, the loop is the regular hexagon pictured below.



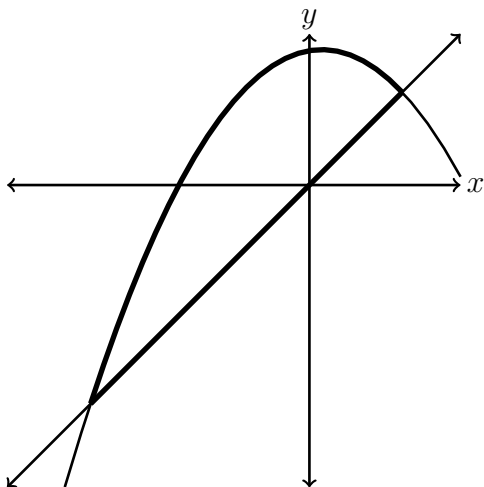
Similar to the situation in part (ii), the loop has reflective symmetry in both the x and y axes. Thus, we might again guess that there is an inscribed square with reflective symmetry in both axes. If such a square exists, then one of the vertices will have the form (t, t) . Thus, we are looking for the points (there are two of them) where the line with equation $y = x$ intersects the hexagon.

The line with equation $y = x$ must intersect the part of the hexagon that is in the first quadrant. Notice that $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ is above the line with equation $y = x$, so the horizontal line segment in the first quadrant will not intersect the line with equation $y = x$. Therefore, the intersection point we seek is on the line segment joining $(1, 0)$ to $(\frac{1}{2}, \frac{\sqrt{3}}{2})$, which has equation $y = -\sqrt{3}x + \sqrt{3}$. To find the point of intersection, we set $x = -\sqrt{3}x + \sqrt{3}$ and solve to get $x = \frac{\sqrt{3}}{1+\sqrt{3}}$. Recall that we hope the square has vertices at (t, t) , $(-t, t)$, $(-t, -t)$, and $(t, -t)$. You can check that with $t = \frac{\sqrt{3}}{1+\sqrt{3}}$, these four points lie on the given hexagon. Below is a picture of the hexagon with the inscribed square.

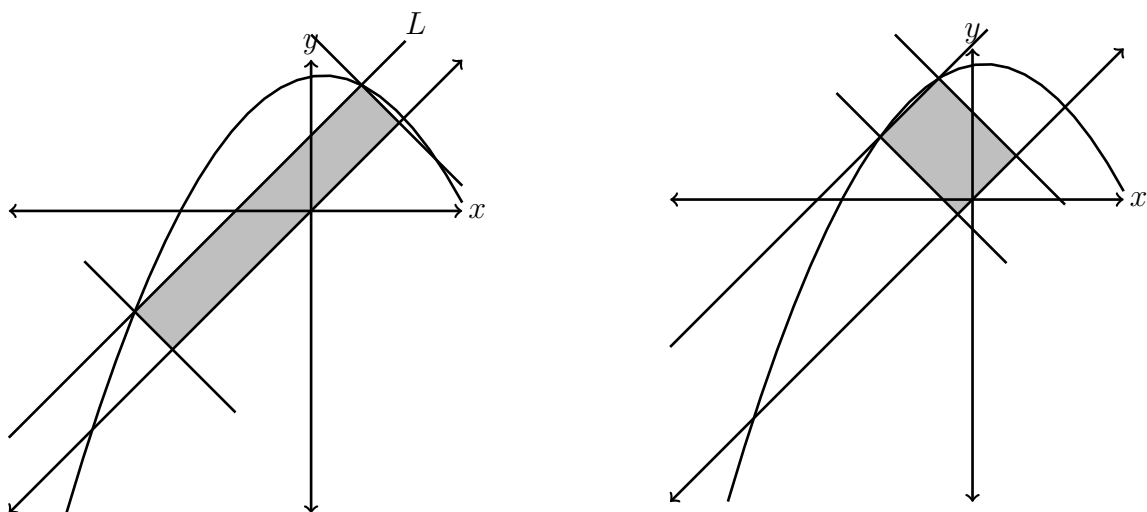


Follow up question: There are two other squares inscribed in the hexagon. They arise from rotating the square in the solution by 60° clockwise or 60° counterclockwise about the origin. Are these the only possible squares?

- (iv) The graphs of the relations $y = -\frac{1}{2}x^2 + \frac{1}{6}x + \frac{16}{9}$ and $y = x$ are pictured on the same axes below. The loop, which is the boundary of the region enclosed by the two graphs, is bold.



One approach is to try to find a square with two adjacent vertices on the line with equation $y = x$ and the other two vertices on the parabola. Why should we expect such a square even exists? Imagine drawing a line L that is parallel to the line with equation $y = x$, lies above the line with equation $y = x$, and intersects the parabola twice. By drawing lines with slope -1 through these two points of intersection, we get a rectangle with two vertices on the parabola and two vertices on the line. The diagrams below show such configurations with the rectangle shaded in each case.



If the points of intersection of L with the parabola are far apart, as in the left diagram, then the two sides of the rectangle that are parallel to the line $y = x$ are longer than the other two sides. If the points of intersection are close together, as in the right diagram, then the two sides of the rectangle that are parallel to the line $y = x$ are shorter than the other two sides. It seems reasonable to guess that somewhere in between those two extremes is a happy medium where the rectangle is a square. Formally, this kind of argument is an application of something called the *Intermediate Value Theorem*.

Let's use this thinking to find the coordinates of such a square. Let L be the line with equation $y = x + b$ where b is the positive real number that will cause the rectangle to be a square. The vertices of the square we hope for are the points of intersection

of L with the parabola. To find these points, we set $x + b$ equal to $-\frac{1}{2}x^2 + \frac{1}{6}x + \frac{16}{9}$ and rearrange to get.

$$\begin{aligned}x + b &= -\frac{1}{2}x^2 + \frac{1}{6}x + \frac{16}{9} \\0 &= \frac{1}{2}x^2 + \frac{5}{6}x + b - \frac{16}{9} \\0 &= 9x^2 + 15x + (18b - 32)\end{aligned}$$

where the last line is obtained by multiplying the second-last line by 18.

We can now use the quadratic formula to get

$$\begin{aligned}x &= \frac{-15 \pm \sqrt{15^2 - 4(9)(18b - 32)}}{18} \\&= \frac{-5 \pm \sqrt{5^2 - 4(18b - 32)}}{6} \\&= \frac{-5 \pm 3\sqrt{17 - 8b}}{6}\end{aligned}$$

Note: Since we are assuming that the line $y = x + b$ has two points of intersection with the parabola, this quadratic equation must have two distinct real solutions. This means we must have $17 - 8b > 0$ and so $b < \frac{17}{8}$. Revisiting the diagrams, it should be clear that there is an upper bound on the values of b that can result in the line intersecting the parabola. The diagrams also indicate that $b > 0$.

Therefore, the x -coordinates of the points of intersection are $\frac{-5 - 3\sqrt{17 - 8b}}{6}$ and $\frac{-5 + 3\sqrt{17 - 8b}}{6}$.

The difference between the x -coordinates of these points is

$$\frac{-5 + 3\sqrt{17 - 8b}}{6} - \frac{-5 - 3\sqrt{17 - 8b}}{6} = \sqrt{17 - 8b}$$

and since the points are on a line of slope 1, the distance between the y -coordinates is also $\sqrt{17 - 8b}$. Therefore, the distance between the two points is

$$\begin{aligned}\sqrt{\left(\sqrt{17 - 8b}\right)^2 + \left(\sqrt{17 - 8b}\right)^2} &= \sqrt{2(17 - 8b)} \\&= \sqrt{34 - 16b}\end{aligned}$$

We will leave it as an exercise to show that the perpendicular distance between L and the line with equation $y = x$ is equal to $\frac{b}{\sqrt{2}}$. This means our square has sides of length $\sqrt{34 - 16b}$ and $\frac{b}{\sqrt{2}}$. A square has four equal sides, so this implies the equation $\sqrt{34 - 16b} = \frac{b}{\sqrt{2}}$. Squaring both sides, we have $34 - 16b = \frac{b^2}{2}$ which can

be rearranged to get $b^2 + 32b - 68 = 0$. Factoring gives $(b - 2)(b + 34) = 0$, so the possible values of b are 2 and -34 .

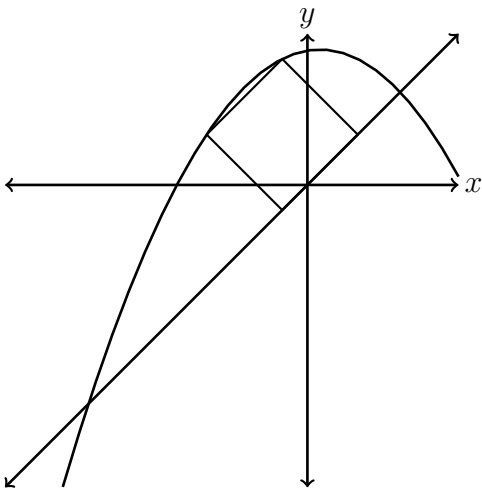
Since we need b to be positive, we conclude that $b = 2$, so the line L has equation $y = x + 2$. The x -coordinates of the points of intersection of L with the parabola were computed earlier. Substituting $b = 2$ into these expressions, we get x -coordinates of $\frac{-5 - 3\sqrt{17 - 8b}}{6} = -\frac{4}{3}$ and $\frac{-5 + 3\sqrt{17 - 8b}}{6} = -\frac{1}{3}$. These points lie on the line with equation $y = x + 2$, so the coordinates of two of the vertices of the square are $(-\frac{4}{3}, \frac{2}{3})$ and $(-\frac{1}{3}, \frac{5}{3})$.

To find the other two vertices, we can find the equations of the lines of slope -1 through each of these points, then find their intersection points with the line with equation $y = x$. We will not include the calculations here, but the resulting points are $(-\frac{1}{3}, -\frac{1}{3})$ and $(\frac{2}{3}, \frac{2}{3})$.

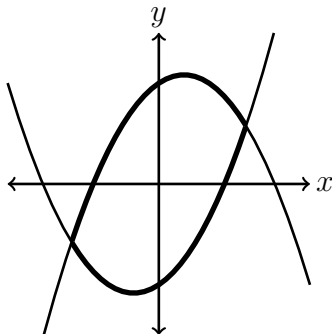
Indeed, it is not difficult to check that the points

$$\left(-\frac{4}{3}, \frac{2}{3}\right), \left(-\frac{1}{3}, -\frac{1}{3}\right), \left(\frac{2}{3}, \frac{2}{3}\right), \left(-\frac{1}{3}, \frac{5}{3}\right)$$

are the vertices of a square. Below is a plot of the loop along with the square with these four points as its vertices.



(v) The two parabolas are pictured below. The loop is bold.



By the discussion at the end of the solution to part (i), the four points

$$\left(1, \frac{1}{3}\right), \left(-\frac{1}{3}, 1\right), \left(-1, -\frac{1}{3}\right), \left(\frac{1}{3}, -1\right).$$

are the vertices of a square. Furthermore, you can check that these four points all lie on the loop.

For the rest of the solution to this part, we will show how one might find these four points. The task is a bit trickier here because, unlike in earlier parts, it is not easy to guess what the slope of the sides of the square should be. However, there is still some symmetry that we can use. Specifically, the loop has 180° rotational symmetry about the origin because the two parabolas are each the result of rotating the other 180° about the origin. This is not difficult to believe from the diagram, but it can also be verified algebraically.

This symmetry suggests two things. The first is that we should look for a square that has 180° rotational symmetry about the origin. The second is that each of the two parabolas should contain two vertices of the square.

A square that has 180° rotational symmetry about the origin must also have 90° rotational symmetry about the origin (you might want to take some time to convince yourself of this). Again by the discussion from part (i), the vertices of the square should be at the points with coordinates (a, b) , $(-b, a)$, $(-a, -b)$, and $(b, -a)$ where a and b are some real numbers.

The parabola with equation $y = x^2 + \frac{2}{3}x - \frac{4}{3}$ should contain two vertices of the square with the property that one of these vertices is the result of rotating the other vertex 90° about the origin. Thus, for some a and b , this parabola should contain the points (a, b) and $(-b, a)$. This implies the following two equations in a and b :

$$\begin{aligned} b &= a^2 + \frac{2}{3}a - \frac{4}{3} \\ a &= (-b)^2 + \frac{2}{3}(-b) - \frac{4}{3} \end{aligned}$$

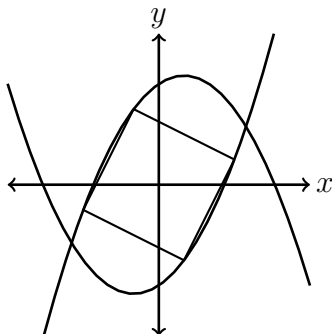
The second equation gives an expression for a in terms of b , so we can substitute this expression into the first equation and simplify as follows:

$$\begin{aligned} b &= \left((-b)^2 + \frac{2}{3}(-b) - \frac{4}{3}\right)^2 + \frac{2}{3}\left((-b)^2 + \frac{2}{3}(-b) - \frac{4}{3}\right) - \frac{4}{3} \\ &= \left(b^2 - \frac{2}{3}b - \frac{4}{3}\right)^2 + \frac{2}{3}\left(b^2 - \frac{2}{3}b - \frac{4}{3}\right) - \frac{4}{3} \\ 9b &= (3b^2 - 2b - 4)^2 + 2(3b^2 - 2b - 4) - 12 && \text{(multiply by 9)} \\ 9b &= 9b^4 - 12b^3 - 14b^2 + 12b - 4 \\ 0 &= 9b^4 - 12b^3 - 14b^2 + 3b - 4 \end{aligned}$$

In general, we should not expect to easily find a closed expression for the root of a quartic polynomial. However, this question was designed to work out nicely. After some checking, you might notice that $b = -1$ is a root of this equation. Indeed, $9(-1)^4 - 12(-1)^3 - 14(-1)^2 + 3(-1) - 4 = 0$. We could factor to get

$$0 = (b + 1)(9b^3 - 21b^2 + 7b - 4)$$

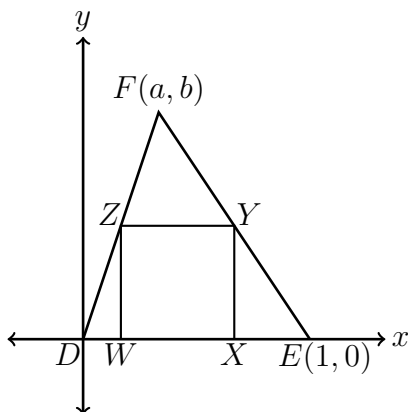
but $b = -1$ is the value we seek, so the other roots of the quartic are not of any use in this solution (though you might want to think about whether they have any “geometric” meaning). With $b = -1$, the equation $a = (-b)^2 + \frac{2}{3}(-b) - \frac{4}{3}$ implies $a = \frac{1}{3}$, so one of the points is $(-b, a) = (1, \frac{1}{3})$. This is the first of the four points listed at the start of the solution to this part. Rotating the point $(1, \frac{1}{3})$ by 90° , 180° , and 270° about the origin gives the other three points. You can check that each of the two parabolas contains two of these points. Below is a diagram of the loop with the square included.



- (b) Given $\triangle ABC$, there exists $\triangle DEF$ so that $\triangle ABC$ is similar to $\triangle DEF$ and DE is horizontal with length 1. So we proceed by examining such a triangle $\triangle DEF$ with vertices $D(0,0)$, $E(1,0)$, and $F(a,b)$ where $a > 0$ and $b > 0$. Moreover, if $\triangle ABC$ is acute, then so is $\triangle DEF$ and we must have $0 < a < 1$.

We will show that there is a unique square whose vertices are on the perimeter of $\triangle DEF$ with two of its vertices are on DE . To find such a square, we will assume there is such a square, deduce conditions on the location of the vertices of such a square, and then confirm that the four points we find actually satisfy the given conditions.

Suppose W, X, Y, Z are the vertices of the square, with W and X on DE . The diagram below shows what this picture should look like.



The equations of the lines joining $(0,0)$ to (a,b) and $(1,0)$ to (a,b) are $y = \frac{b}{a}x$ and $y = \frac{b}{a-1}(x-1)$ respectively.

The coordinates of W , X , Y , and Z must then take the form

$$\begin{aligned} W &= (s, 0) \\ X &= (t, 0) \\ Y &= \left(t, \frac{b}{a-1}(t-1) \right) \\ Z &= \left(s, \frac{b}{a} \right) \end{aligned}$$

for some s and t with $0 < s < t < 1$. Since we want these to be the vertices of a square, we have the following pair of simultaneous equations:

$$\begin{aligned} \frac{b}{a}s &= \frac{b}{a-1}(t-1) \\ t - s &= \frac{b}{a}. \end{aligned}$$

The first equation comes from insisting that the line joining Y to Z is horizontal, and the second comes from insisting that the width of the resulting rectangle is equal to its height. Rearranging the second equation gives $t = \left(\frac{b}{a} + 1\right)s$ and substituting this into the first equation gives

$$\frac{b}{a}s = \frac{b}{a-1} \left(\left(\frac{b}{a} + 1 \right) s - 1 \right).$$

Solving for s gives

$$s = \frac{a}{b+1}$$

and therefore

$$t = \frac{b+a}{b+1}.$$

Therefore, if such a square exists, it must have vertices

$$\begin{aligned} W &= \left(\frac{a}{b+1}, 0 \right) \\ X &= \left(\frac{b+a}{b+1}, 0 \right) \\ Y &= \left(\frac{b+a}{b+1}, \frac{b}{b+1} \right) \\ Z &= \left(\frac{a}{b+1}, \frac{b}{b+1} \right). \end{aligned}$$

We leave it as an exercise to show that these four vertices indeed are the vertices of a square and that they are all on the perimeter of the triangle. Since the assumption that they were the vertices of a square led to specific coordinates of the four points, we can also conclude that there is exactly one square with vertices on $\triangle DEF$ so that two of them are on DE .

By the argument above, in $\triangle DEF$, there is always exactly one square with two vertices on DE and one vertex on each of DF and EF . In $\triangle ABC$, this corresponds to exactly

one square with two vertices on AB and one on each of AC and BC . Since there are three sides in $\triangle ABC$, there are exactly three squares.

Note that if $\triangle ABC$ is obtuse, then the preceding arguments can be modified to show that there is exactly one square with vertices on the perimeter of $\triangle ABC$. Two of the vertices will always be on the longest side. What do you think happens in a right-angled triangle?



Problem of the Month

Problem 3: December 2023

This month's problem is an extension of Problem 3 from Part B of the 2023 Canadian Intermediate Mathematics Contest. The original problem was stated as follows:

The positive integers are written into rows so that Row n includes every integer m with the following properties:

- (i) m is a multiple of n ,
- (ii) $m \leq n^2$, and
- (iii) m is not in an earlier row.

The table below shows the first six rows.

Row 1	1
Row 2	2, 4
Row 3	3, 6, 9
Row 4	8, 12, 16
Row 5	5, 10, 15, 20, 25
Row 6	18, 24, 30, 36

- (a) Determine the smallest integer in Row 10.
- (b) Show that, for all positive integers $n \geq 3$, Row n includes each of $n^2 - n$ and $n^2 - 2n$.
- (c) Determine the largest positive integer n with the property that Row n does not include $n^2 - 10n$.

If you have not already done so, we suggest thinking about the parts above before proceeding.

- (a) For each positive integer k , determine the largest positive integer n with the property that Row n does not include $n^2 - kn$. (This generalizes part (c) from the original problem.)

In the remaining questions, $f(n)$ is defined for each $n \geq 1$ to be the largest non-negative integer m such that $m \leq n$ and mn is *not* in Row n . For example, Row 6 is 18, 24, 30, 36, so $f(6) = 2$ since $2 \times 6 = 12$ is not in Row 6 but 3×6 , 4×6 , 5×6 , and 6×6 are all in Row 6.

- (b) Show that $f(p) = 0$ for all prime numbers p . (Looking closely at the definition of $f(n)$, $f(p) = 0$ means that every positive multiple of p from p through p^2 appears in Row p .)
 - (c) Find an expression for $f(pq)$ where p and q are prime numbers. Justify that the expression is correct.
 - (d) Find an expression for $f(p^d)$ where p is a prime number and d is a positive integer.
 - (e) Take some time to explore the function f further on your own. Are there other results you can prove about the function beyond what is done in (b), (c) and (d)? Is there a nice way to compute $f(n)$ in general without computing each of the first $n - 1$ rows?
-
-



Problem of the Month

Problem 3: December 2023

Hint

For positive integers m and n with $m \leq n$, the integer mn is *not* in Row n if and only if there are integers a and b with $m < a \leq b < n$ such that $mn = ab$. This fact will likely be useful in all parts of this problem.

- (a) With the above fact in mind, convince yourself that if $n^2 - kn$ is not in Row n , then there must be positive integers x and y such that $x \leq y < k$ and $n^2 - kn = (n - x)(n - y)$. Now solve for n and try to determine how to choose x and y to maximize this expression for n .
- (b) If integers a , b , n , and p satisfy $mp = ab$ and p is prime, then at least one of a and b must be a multiple of p .
- (c) The general formula is $f(pq) = (p-1)(q-1)$. In this part and the rest of the parts, you might find the following observation useful: If u and v are integers with $u < v$, then $u \leq v - 1$.
- (d) Consider the cases when d is even and when d is odd separately.
- (e) To formulate a guess at how to find $f(n)$ in general, consider some values of n for which you know the value of $f(n)$ and list the positive factors of n in increasing order. If you are so inclined, you could write some computer code to compute $f(n)$ for some moderately sized values of n .

The value of $f(n)$ depends on how n factors, so it is probably unreasonable to expect a general algebraic expression for $f(n)$ similar to $f(pq) = (p-1)(q-1)$. Instead, try to find a simple procedure to compute $f(n)$ assuming that you already know all the positive factors of n .



Problem of the Month

Solution to Problem 3: December 2023

For solutions to the original problem from the Canadian Intermediate Mathematics Contest, please refer to the solutions on our website.

We will start by presenting two general facts. It is worth spending some time digesting them before reading the rest of the solution.

Fact 1: For all positive integers m and n with $m \leq n$, if the integer mn is not in Row n , then there are positive integers a and b such that $ab = mn$ and $m < a \leq b < n$.

Proof. Assume that m and n are positive integers with $m \leq n$ such that mn is not in Row n . Then mn must be in an earlier row since $mn \leq n^2$. Thus, mn is in Row b for some integer $b < n$. The integers in Row b are multiples of b that are no larger than b^2 , which means that $mn = ab$ for some positive integer a with $a \leq b$. Rearranging $mn = ab$, we get $\frac{a}{m} = \frac{n}{b}$. Since $b < n$, $\frac{n}{b} > 1$, and so $\frac{a}{m} > 1$ which implies $a > m$. Therefore, there are positive integers a and b such that $mn = ab$ and $m < a \leq b < n$. \square

Fact 2: For all positive integers m and n with $m \leq n$, if there are positive integers a and b such that $ab = mn$ and $m < a \leq b < n$, then the integer mn is not in Row n .

Proof. Assume that m and n are positive integers with $m \leq n$ and that there are positive integers a and b with $mn = ab$ and $m < a \leq b < n$. We are assuming $a \leq b$ and that a and b are positive, so $ab \leq b^2$, which implies $mn \leq b^2$ since we are assuming $mn = ab$. Therefore, mn is a positive multiple of b that does not exceed b^2 , so mn cannot appear any later than Row b . Since $b < n$, mn is not in Row n . \square

- (a) Suppose k is a positive integer. If n is a positive integer with $n \leq k$, then $n^2 - kn \leq 0$, and so $n^2 - kn$ is not in Row n . Therefore, we might as well assume that $n > k$ as we look for the largest n such that $n^2 - kn$ is not in Row n .

Assume now that $n^2 - kn = n(n - k)$ is not in Row n . By Fact 1 with $m = n - k$, there must be integers a and b such that $ab = n(n - k)$ and $n - k < a \leq b < n$.

Because a and b are integers strictly between $n - k$ and n , there must exist integers x and y with $1 < x \leq y < k$ such that $a = n - y$ and $b = n - x$. We are assuming that $n^2 - nk = ab$, so we can substitute to get $n^2 - nk = (n - y)(n - x) = n^2 - (x + y)n + xy$. This can be simplified and rearranged to get $(x + y - k)n = xy$. The integers x and y are both positive, so xy is positive, which implies that $x + y - k$ is positive so we can solve for n to get $n = \frac{xy}{x + y - k}$.

Now suppose $x + y > k + 1$ and consider the expression $\frac{(x - 1)y}{(x - 1) + y - k} - \frac{xy}{x + y - k}$, which

we can simplify as follows:

$$\begin{aligned}
\frac{(x-1)y}{(x-1)+y-k} - \frac{xy}{x+y-k} &= \frac{(x-1)y(x+y-k) - (xy)(x+y-1-k)}{(x+y-k)(x+y-1-k)} \\
&= \frac{x^2y + xy^2 - xyk - xy - y^2 + yk - x^2y - xy^2 + xy + xyk}{(x+y-k)(x+y-1-k)} \\
&= \frac{-y^2 + yk}{(x+y-k)(x+y-1-k)} \\
&= \frac{y(k-y)}{(x+y-k)(x+y-1-k)}
\end{aligned}$$

Carefully considering the components of the fraction above, we have that y is a positive integer with $y < k$, so the quantity $k - y$ must be positive, and since $x + y > k + 1$, the quantities $x + y - k$ and $x + y - k - 1$ are both positive. Therefore, the fraction at the end of the calculation above is positive, so we get that

$$\frac{(x-1)y}{(x-1)+y-k} - \frac{xy}{x+y-k} < 0$$

which implies that $\frac{(x-1)y}{(x-1)+y-k} < \frac{xy}{x+y-k}$.

We have shown that if $x + y > k + 1$, then the quantity $\frac{xy}{x+y-k}$ can be made *larger* by decreasing x by 1. A similar argument shows that if $x + y > k + 1$, then the quantity gets larger when y is decreased by 1. Since $n = \frac{xy}{x+y-k}$ and we seek the largest possible n , we conclude that n can always be made larger by decreasing $x + y$ by 1, provided $x + y$ is larger than $k + 1$. Therefore, the maximum value of n must occur when $x + y = k + 1$.

Note: It is possible (and essentially assured) that $\frac{xy}{x+y-k}$ is not always an integer. However, since it is maximized when $x + y = k + 1$, which makes the denominator equal to 1, the maximum possible value of $\frac{xy}{x+y-k}$ given the constraints on x and y happens to be an integer. Therefore, the maximum possible integer value of $\frac{xy}{x+y-k}$ is the maximum possible value, which occurs when $x + y = k + 1$.

Substituting $x + y = k + 1$ into $n = \frac{xy}{x+y-k}$, we get $n = xy$. Therefore, the task is to maximize xy subject to the conditions $1 < x \leq y < k$ and $x + y = k + 1$.

Substituting $y = k + 1 - x$ into xy , we get $xy = x(k + 1 - x) = (k + 1)x - x^2$. Since k is fixed, the expression $(k + 1)x - x^2$ is a quadratic in x with a negative coefficient of x^2 . Therefore, it has a maximum value and it occurs when $x = \frac{k+1}{2}$.

We are looking for the maximum among integer values of x , so if $\frac{k+1}{2}$ is an integer, the maximum occurs exactly when $x = \frac{k+1}{2}$. Since $x + y = k + 1$, this implies that $y = \frac{k+1}{2}$ as well.

Otherwise, the maximum will occur at the integer nearest $\frac{k+1}{2}$. There is a tie between $\frac{k+1}{2} - \frac{1}{2} = \frac{k}{2}$ and $\frac{k+1}{2} + \frac{1}{2} = \frac{k+2}{2}$. Notice that the sum of these two possible values of x is $\frac{k}{2} + \frac{k+2}{2} = k+1$, so one of the values must be x and the other must be y . Since $x \leq y$, we have that $x = \frac{k}{2}$ and $y = \frac{k+2}{2}$.

Now to summarize what we have so far: if we assume that $n^2 - kn$ is not in Row n , then n is no larger than $\left(\frac{k+1}{2}\right)^2$ when k is odd and n is no larger than $\frac{k}{2} \left(\frac{k+2}{2}\right) = \frac{k(k+2)}{4}$ when k is even. Note that when $k = 10$, we get $\frac{10(12)}{4} = 30$, which agrees with the answer to part (c) of the original problem.

To finish the argument, we need to verify that $\left[\left(\frac{k+1}{2}\right)^2\right]^2 - k \left(\frac{k+1}{2}\right)^2$ is not in Row $\left(\frac{k+1}{2}\right)^2$ when k is odd and that $\left(\frac{k(k+2)}{4}\right)^2 - k \left(\frac{k(k+2)}{4}\right)$ is not in Row $\frac{k(k+2)}{4}$ when k is even. We will only include the verification for when k is odd here. The verification when k is even is similar.

Expanding and simplifying, we have

$$\begin{aligned} \left[\left(\frac{k+1}{2}\right)^2\right]^2 - k \left(\frac{k+1}{2}\right)^2 &= \left(\frac{k+1}{2}\right)^2 \left[\left(\frac{k+1}{2}\right)^2 - k\right] \\ &= \left(\frac{k+1}{2}\right)^2 \left[\frac{k^2 + 2k + 1 - 4k}{4}\right] \\ &= \left(\frac{k+1}{2}\right)^2 \left[\frac{k^2 - 2k + 1}{4}\right] \\ &= \left(\frac{k^2 - 1}{4}\right)^2 \end{aligned}$$

Now set $n = \left(\frac{k+1}{2}\right)^2$, $m = n - k$, and $a = b = \frac{k^2 - 1}{4}$. It follows from the calculation above that $mn = ab$. If we can show that $m < a \leq b < n$, then we can apply Fact 2 to get that $mn = n^2 - kn$ is not in Row n . Note that $a = b$ by definition, so $a \leq b$ is automatically true. Expanding gives $n = \frac{k^2 + 2k + 1}{4}$ and $m = \frac{k^2 - 2k + 1}{4}$, and it is easily checked that

$$m = \frac{k^2 - 2k + 1}{4} < \frac{k^2 - 1}{4} < \frac{k^2 + 2k + 1}{4} = n$$

for all $k > 1$. This means we can apply Fact 2 as long as $k > 1$, which completes the verification that the maximum value of n is exactly $\left(\frac{k+1}{2}\right)^2$ when k is odd and $k > 1$.

Finally, if $k = 1$, then $n^2 - kn = n^2 - n$. By the original contest problem, $n^2 - n$ is in Row n for $n \geq 3$. Also, $n^2 - n$ is in Row n when $n = 2$, but when $n = 1$, $n^2 - n = 0$ is

not in Row n . Therefore, when $k = 1$, the maximum value of n for which $n^2 - kn$ is not in Row n is $n = 1$. Notice that $\left(\frac{k+1}{2}\right)^2$ evaluates to 1 when $k = 1$, so the formula works even in this case.

- (b) Suppose m is a positive integer with $m \leq p$ such that mp is not in Row p . By Fact 1, there exist positive integers a and b such that $mp = ab$ and $m < a \leq b < p$. The equation $mp = ab$ implies that ab is a multiple of p , and since p is a prime number, a or b must be a multiple of p . (This is by a fact often known as Euclid's Lemma.) We also have that a and b are positive and both less than p , so it is impossible for either of them to be a multiple of p because p is prime.

We conclude that there are no positive integers m with $m \leq p$ such that mp is not in Row p . In other words, mp is in Row p for every positive integer m with $m \leq p$.

Since 0 is not in any row and $0 = 0 \times p$ is a multiple of p , 0 is the largest integer m with the property that $m \leq p$ and mp is not in Row p . Thus, $f(p) = 0$ when p is a prime number.

- (c) We will show that $f(pq) = (p-1)(q-1)$. To do this, we need to establish two things:
- $(p-1)(q-1)pq$ is not in Row pq .
 - For every integer k with $(p-1)(q-1) < k \leq pq$, the integer kpq is in Row pq .

We will assume that $p \leq q$. If not, then $q \leq p$ and the proof is essentially identical.

To see that the first point is true, set $n = pq$, $m = (p-1)(q-1)$, $a = q(p-1)$, and $b = p(q-1)$ and apply Fact 2. Note that $p-1 < p$ so $(p-1)(q-1) < p(q-1)$ or $m < a$. We also have $p \leq q$, which can be rearranged to get $-q \leq -p$, and so $pq - q \leq pq - p$, which can be factored to get $q(p-1) \leq p(q-1)$ or $a \leq b$. Finally, $q-1 < q$, so $b = p(q-1) < pq = n$. Putting these inequalities together, we have $m < a \leq b < n$. Since $mn = (p-1)(q-1)pq = ab$, the conditions of Fact 2 are satisfied and we can conclude that $(p-1)(q-1)pq$ is not in Row pq .

Now suppose k is an integer such that $(p-1)(q-1) < k \leq pq$. We wish to show that kpq must be in Row pq . To do this, we will assume that it is not in Row pq and deduce a contradiction.

To that end, assume that kpq is not in Row pq . By Fact 1, there must be integers a and b with $k < a \leq b < pq$ such that $kpq = ab$. If a is a multiple of pq , then we would have $pq \leq a$, but $a < pq$ by assumption, so this is impossible. Therefore, a is not a multiple of pq . Similarly, b is not a multiple of pq .

Since ab is a multiple of pq and p and q are prime, either a is a multiple of p and b is a multiple of q , or a is a multiple of q and b is a multiple of p . We will assume that a is a multiple of p and b is a multiple of q . The other situation is similar. This implies the existence of integers x and y such that $px = a$ and $qy = b$. Observe that since $a < pq$, we must have $x < q$, and similarly $y < p$. Since x, y, p , and q are integers, we conclude that $x \leq q-1$ and $y \leq p-1$.

Since we are assuming that $kpq = ab$, we have

$$k = \frac{ab}{pq} = \frac{pxqy}{pq} = xy \leq (p-1)(q-1)$$

However, we are also assuming that $(p-1)(q-1) < k$, so we can deduce that

$$k \leq (p-1)(q-1) < k$$

which is impossible because it implies $k < k$. Therefore, our assumption that kpq is *not* in Row pq must be false, and we conclude that kpq is in Row pq .

Therefore, $f(pq) = (p-1)(q-1)$ as claimed earlier.

- (d) The value of $f(p^d)$ depends on whether d is even or odd. If $d = 2r$ for some integer r , then $f(p^d) = (p^r - 1)^2$. If $d = 2r + 1$ for some integer r , then $f(p^d) = (p^r - 1)(p^{r+1} - 1)$.

Here we include a proof in the case where d is even. The proof when d is odd is a slight modification of the proof given for when d is even.

Assume $d = 2r$ for some positive integer r . As with part (c), we need to show that

- $p^{2r} (p^r - 1)^2$ is not in Row p^{2r} ,
- and kp^{2r} is in Row p^{2r} for every k with $(p^r - 1)^2 < k \leq p^{2r}$.

If we set $m = (p^r - 1)^2$, $n = p^{2r}$, and $a = b = p^r (p^r - 1)$, then we will have $m < a \leq b < n$ and $mn = ab$, so $p^{2r} (p^r - 1)^2$ is not in Row p^{2r} by Fact 2. This establishes the first bullet point above.

For the second, suppose kp^{2r} is not in Row p^{2r} for some k with $(p^r - 1)^2 < k \leq p^{2r}$. By Fact 1, there must be integers a and b with $(p^r - 1)^2 < a \leq b < p^{2r}$ and $kp^{2r} = ab$.

The product ab must have at least $2r$ factors of the prime number p , meaning the two integers a and b have at least a total of $2r$ factors of p between them. Therefore, there are non-negative integers u and v such that $u + v = 2r$, a is a multiple of p^u , and b is a multiple of p^v . By definition, there are positive integers x and y such that $p^u x = a$ and $p^v y = b$.

We are also assuming that $a \leq b < p^{2r} = p^{u+v} = p^u p^v$, so $a = p^u x < p^u p^v$ and $b = p^v y < p^u p^v$, which implies $x < p^v$ and $y < p^u$. Since x , p , u , and v are non-negative integers, we have $x \leq p^v - 1$ and $y \leq p^u - 1$. Multiplying these inequalities, we obtain $xy \leq (p^u - 1)(p^v - 1)$. We are also assuming that $kp^{2r} = ab$, so we can substitute to get

$$kp^{2r} = xyp^u p^v = xyp^{u+v} = xyp^{2r}$$

Thus, $kp^{2r} = xyp^{2r}$, so $k = xy$, which implies $k \leq (p^u - 1)(p^v - 1)$.

We can now conclude that $(p^r - 1)^2 < k \leq (p^u - 1)(p^v - 1)$. To finish the argument, we will show that $(p^u - 1)(p^v - 1) \leq (p^r - 1)^2$, which would imply that $(p^r - 1)^2 < (p^r - 1)^2$, which is clearly false.

To show that $(p^u - 1)(p^v - 1) \leq (p^r - 1)^2$, we will show that if u and v are non-negative integers with $u + v = 2r$, then the quantity $(p^u - 1)(p^v - 1)$ is maximized when $u = v = r$.

Suppose u and v are non-negative integers with $u > v$ and $u + v = 2r$. Consider the quantity $(p^{u-1} - 1)(p^{v+1} - 1) - (p^u - 1)(p^v - 1)$. We can manipulate this difference as

follows.

$$\begin{aligned}
(p^{u-1} - 1)(p^{v+1} - 1) - (p^u - 1)(p^v - 1) &= p^{u+v} - p^{u-1} - p^{v+1} + 1 - p^{u+v} + p^u + p^v - 1 \\
&= p^u + p^v - p^{u-1} - p^{v+1} \\
&= p^v(p^{u-v} + 1 - p^{u-v-1} - p) \\
&= p^v(p^{u-v} - p) \left(1 - \frac{1}{p}\right)
\end{aligned}$$

We are assuming that $u > v$. Since $u + v$ is even, it must also be true that $u - v$ is even. Therefore, $u - v \geq 2$. As well, p is a prime number so $p \geq 2$, which implies $p^{u-v} - p > 0$ and $1 - \frac{1}{p} > 0$. The quantity p^v is positive, and so we conclude that $p^v(p^{u-v} - p) \left(1 - \frac{1}{p}\right) > 0$, which implies that $(p^{u-1} - 1)(p^{v+1} - 1) - (p^u - 1)(p^v - 1) > 0$. Rearranging this inequality, we get $(p^{u-1} - 1)(p^{v+1} - 1) > (p^u - 1)(p^v - 1)$.

What we have shown is that if u and v are non-negative integers with $u + v = 2r$ and $u > v$, we can increase the value of $(p^u - 1)(p^v - 1)$ by decreasing u by 1 and increasing v by 1. It follows that $(p^u - 1)(p^v - 1)$ is maximized when u and v are as close together as possible, which happens when $u = v = \frac{2r}{2} = r$.

- (e) In general, $f(n)$ can be computed as follows: find the unique positive integers x and y with $x \leq y$, $xy = n$, and $y - x$ as small as possible. Then $f(n) = (x - 1)(y - 1)$.

For example, if $n = p$ is a prime number, then the only positive factor pair is $p = 1 \times p$, so $x = 1$ and $y = p$. Indeed, from part (b), $f(p) = 0 = (1 - 1)(p - 1) = (x - 1)(y - 1)$. If n is a perfect square, then $n = z^2$ for some positive integer z . Here, we must have $x = y = z$ because this gives $y - x = 0$, which is the smallest that the difference between the factors in a factor pair could possibly be. Indeed, $f(z^2) = (z - 1)^2$ agrees with the case from part (d) where $n = p^{2r}$. As a final example, consider $n = 72$. Its positive factor pairs are $(1, 72)$, $(2, 36)$, $(3, 24)$, $(4, 18)$, $(6, 12)$, and $(8, 9)$. The pair with the smallest difference between the factors is $(8, 9)$, so we have $x = 8$ and $y = 9$, giving $(x - 1)(y - 1) = (8 - 1)(9 - 1) = 56$. Indeed, $56 \times 72 = 63 \times 64$, so 56×72 is not in Row 72. As well, you can check that $72m$ is in Row 72 for each m from 57 through 72, which shows that $f(72) = 56$.

We will now justify that this “formula” always works.

Fix an integer n and suppose $x, y, u,$ and v are positive integers with $x \leq y$, $u \leq v$, and $xy = uv = n$.

If $y - x \leq v - u$, then since $x \leq y$, we have $0 \leq y - x \leq v - u$, so we can conclude that $(y - x)^2 \leq (v - u)^2$. Expanding these expressions and adding $4n$ to both sides, we have

$$\begin{aligned}
y^2 - 2xy + x^2 &\leq v^2 - 2uv + u^2 \\
y^2 - 2xy + x^2 + 4n &\leq v^2 - 2uv + u^2 + 4n \\
y^2 - 2xy + x^2 + 4xy &\leq v^2 - 2uv + u^2 + 4uv \\
y^2 + 2xy + x^2 &\leq v^2 + 2uv + u^2 \\
(y + x)^2 &\leq (v + u)^2
\end{aligned}$$

and since $y + x$ and $v + u$ are both positive, we must have $y + x \leq v + u$. Essentially the same argument in reverse shows that if $y + x \leq v + u$, then $y - x \leq v - u$.

We have shown that $y - x \leq v - u$ exactly when $y + x \leq v + u$. This implies that if we consider all positive factor pairs of n , the one that has the smallest difference is also the one that has the smallest sum. Therefore, we can restate our “formula” as follows: Given a positive integer n , let x and y be positive integers with $xy = n$ and $x + y$ as small as possible. Then $f(n) = (x - 1)(y - 1)$. We will assume that the integers x and y satisfy $x \leq y$. If not, then they can be relabelled.

The overall structure of the argument will resemble the arguments for parts (c) and (d). That is, we will show that $xy(x - 1)(y - 1)$ is not in Row n , but if $(x - 1)(y - 1) < m \leq n$, then mxy is in Row n . The solution presented will use, without proof, a few well-known facts about divisibility and greatest common divisors.

To see that $xy(x - 1)(y - 1)$ is not in Row n , take $m = (x - 1)(y - 1)$, $a = y(x - 1)$, and $b = x(y - 1)$ and apply Fact 2. That these choices of m , a , b , and c satisfy the conditions of Fact 2 can be verified in essentially the same way as they were in the solutions to parts (c) and (d).

Now suppose m is an integer with $(x - 1)(y - 1) < m \leq n = xy$ such that mn is not in Row n . By Fact 1, there are integers a and b such that $m < a \leq b < n$ and $ab = mn$. Let $d = \gcd(n, b)$ and then define $e = \frac{n}{d}$ and $f = \frac{b}{d}$. Observe that the integers e and f must satisfy $\gcd(e, f) = 1$ since if they had a common divisor larger than 1, then the greatest common divisor of n and b would need to be larger than d .

Dividing both sides of $ab = mn$ by d , we get $af = me$, which shows that af is a multiple of e . We have already noted that e and f have no common divisors larger than 1, so we are forced to conclude that a is a multiple of e . That is, there must be some integer g such that $a = eg$.

To summarize, so far we have that $d = \gcd(b, n)$, $de = n$, $df = b$, and $eg = a$. Consider the integers e and f and suppose $e \leq f$. Multiplying both sides by d gives $de \leq df$ or $n \leq b$, but this contradicts our assumption that $b < n$, so we cannot have $e \leq f$. Therefore, $f < e$, and since both e and f are integers, we get $f \leq e - 1$. Similarly, if $d \leq g$, then $de \leq eg$ or $n \leq a$, which contradicts our assumption, so we conclude that $g < d$ and so $g \leq d - 1$.

We are assuming that $(x - 1)(y - 1) < m$, so we have the following

$$n(x - 1)(y - 1) < mn = ab = egdf \leq de(e - 1)(d - 1)$$

which implies $n(x - 1)(y - 1) < de(d - 1)(e - 1)$. Using that $n = de$, we cancel to get $(x - 1)(y - 1) < (d - 1)(e - 1)$ which can be expanded to get $xy - x - y + 1 < de - d - e + 1$, and since $de = xy = n$, this simplifies to $d + e < x + y$. However, $xy = n$ is the factorization of n into a product of positive integers that minimizes the sum of the factors. Since $de = n$, the inequality $d + e < x + y$ is a contradiction of this minimality.

We conclude that our assumption that mn is not in Row n must be false, which means mn is in Row n . This completes the proof.



Problem of the Month

Problem 4: January 2024

In this problem, we will enclose a list of positive integers in square brackets. For example, $[1, 2, 4, 7, 9]$ is a list of positive integers of length five. All lists in this problem will be expressed in *non-decreasing* order. To denote the list of consecutive integers from a to b inclusive, we will use the notation $[a : b]$. For example, $[4 : 7]$ denotes the list $[4, 5, 6, 7]$.

Given a list, A , of positive integers, we will define $f(A)$ to be the increasing list of distinct positive integers that are expressible as the sum of some, but not all, of the items in A . A sum is allowed to consist of just one item in A , and each item in A can be used at most once in a particular sum.

For example, if $A = [1, 1, 3, 7]$, then the sums of at least one but not all of the items in A are shown below.

$1 = 1$	$1 + 1 = 2$	$1 + 1 + 3 = 5$
$1 = 1$	$1 + 3 = 4$	$1 + 1 + 7 = 9$
$3 = 3$	$1 + 7 = 8$	$1 + 3 + 7 = 11$
$7 = 7$	$1 + 3 = 4$	$1 + 3 + 7 = 11$
	$1 + 7 = 8$	
	$3 + 7 = 10$	

Therefore, $f(A) = [1, 2, 3, 4, 5, 7, 8, 9, 10, 11]$.

A list, D , of positive integers is called *compressible* if there exists some list A with $f(A) = D$. In this situation, we say that D is compressed by A or that A compresses D . From the example above, we have that $[1, 2, 3, 4, 5, 7, 8, 9, 10, 11]$ is compressible and is compressed by $[1, 1, 3, 7]$.

- Find all lists of length four that compress $[1 : 9]$ and explain why no list of length three or less can compress $[1 : 9]$.
 - Find a list, A , that compresses $[1 : 100]$ and is as short as possible.
 - Show that for all positive integers n , the list $[1 : n]$ is compressible. For each n , determine the smallest possible length of a list that compresses $[1 : n]$.
 - Show that for all positive integers $k \geq 3$ there are only finitely many compressible lists of k consecutive positive integers. That is, for each positive integer $k \geq 3$, show that there are only finitely many positive integers m for which $[m : m + k - 1]$ is compressible.
 - Find the largest positive integer k such that $[5 : k]$ is not compressible. (A full solution will not be provided for this part.)
-
-



Problem of the Month

Problem 4: January 2024

Hint

- (a) Any list that compresses $[1 : 9]$ must contain 1. Think about the largest possible number of integers in $f(A)$ when A is a list of length k .
 - (b) First, try to find a list that compresses $[1 : 63]$ that is as short as possible. It might help to read about the *binary representation* of positive integers.
 - (c) Work out a few more examples like the one in (b). It is possible to compress $[1 : n]$ using a list A that consists entirely or almost entirely of powers of 2.
 - (d) For $k \geq 3$ and $m \geq 2$, if A compresses $[m : m + k - 1]$, then A must contain m and $m + 1$.
 - (e) The answer is 39. Do not worry about trying to compress $[5 : k]$ using as short a list as possible. As well, *inductive* thinking could be useful here. Suppose you can show that there is some k with the property that $[5 : k]$, $[5 : k + 1]$, $[5 : k + 2]$, $[5 : k + 3]$, and $[5 : k + 4]$ are all compressible. Can you deduce that $[5 : n]$ is compressible for all $n \geq k$?
-
-

Problem of the Month

Solution to Problem 4: January 2023

- (a) The positive integer 1 cannot be expressed as the sum of more than 1 positive integer, so if A compresses $[1 : 9]$, then A must contain the integer 1. We want A to be a list of positive integers of length four, and 1 is the smallest positive integer, so we can assume that $A = [1, a, b, c]$ where $a, b,$ and c are integers with $1 \leq a \leq b \leq c$.

The largest sum in $f(A)$ must be the sum of the three largest items in A , which is $a + b + c$, and since $f(A) = [1 : 9]$, we have $a + b + c = 9$.

Suppose that $a = 1$ and $b = 1$, then $a + b + c = 9$ implies that $c = 7$, but then $A = [1, 1, 1, 7]$, which does not satisfy $f(A) = [1 : 9]$ since, for example, 4 is not in $f(A)$. Therefore, a and b are not both 1.

Suppose that $a = 1$ and $b = 2$. Then $a + b + c = 9$ implies $c = 6$, so $A = [1, 1, 2, 6]$, but then $f(A)$ does not contain 5 so $f(A) \neq [1 : 9]$. Therefore, we cannot have $a = 1$ and $b = 2$.

Suppose that $a = 1$ and $b = 3$. Then $c = 5$, so $A = [1, 1, 3, 5]$. It can be verified that $f([1, 1, 3, 5]) = [1 : 9]$, so this gives one possible list A .

Suppose $a = 1$ and $b \geq 4$. Then $c \geq 4$ as well since $b \leq c$, but if $A = [1, 1, b, c]$ where $b \geq 4$ and $c \geq 4$, then $f(A)$ does not contain 3 which means $f(A) \neq [1 : 9]$.

So far, we have shown that $A = [1, 1, 3, 5]$ is the only list of the form we seek with $a = 1$ and $f(A) = [1 : 9]$.

We will now similarly examine the possibilities when $a = 2$.

If $a = 2$ and $b = 2$, then $c = 5$, so $A = [1, 2, 2, 5]$. It can be checked that $f(A) = [1 : 9]$ in this case.

If $a = 2$ and $b = 3$, then $c = 4$ and $A = [1, 2, 3, 4]$. It can be checked that $f(A) = [1 : 9]$ in this case.

If $a = 2$ and $b \geq 4$, then $a + b + c = 9$ implies that $c < 4$, but we are assuming that $b \leq c$, so this is impossible.

Therefore, the only possibilities with $a = 2$ are $A = [1, 2, 2, 5]$ and $A = [1, 2, 3, 4]$.

If $a \geq 3$, then $A = [1, a, b, c]$ has $b \geq 3$ and $c \geq 3$ as well, but this would prevent 2 from being in $f(A)$, so A cannot compress $[1 : 9]$ in this case.

Therefore, the only lists of length four that compress $[1 : 9]$ are $[1, 1, 3, 5]$, $[1, 2, 2, 5]$, and $[1, 2, 3, 4]$.

To see why no shorter list can compress $[1 : 9]$, we use the general observation that if A is a list of length k , then there are at most $2^k - 2$ integers in $f(A)$. This is because for each sum computed for $f(A)$, each of the k items in A is either included in the sum or it is not. This gives 2^k possible ways of computing a sum of some of the items in A . However, this count includes the sum of none of the items in A (all are excluded from the sum) and the sum of all of the items in A . Both of these sums are excluded from $f(A)$, so there are $2^k - 2$ ways to compute a sum to go in $f(A)$. We note that there could be multiple ways

to express the same integer in $f(A)$ as a sum of items in A , which is why we can only say there are *at most* $2^k - 2$ integers in $f(A)$ – in practice, there could be fewer than $2^k - 2$.

If $k \leq 3$, then $2^k - 2 \leq 2^3 - 2 = 6$, and so if A is a list of length at most three, then $f(A)$ has at most 6 integers. Therefore, $[1 : 9]$, which has nine integers, cannot be compressed by a list of length less than four.

- (b) At the end of the solution to part (a), we argued that if A is a list of length k , then there are at most $2^k - 2$ integers in $f(A)$. Since $2^6 - 2 = 62$ and $2^7 - 2 = 126$, we need $k \geq 7$ to achieve $2^k - 2 \geq 100$. Therefore, a list that compresses $[1 : 100]$ must have length at least seven. Now, we will provide a list of minimal length (seven) that compresses $[1 : 100]$.

Consider $A = [1, 2, 4, 8, 16, 32, 38]$. Note that 1 is in $f(A)$ and that the sum of all items in A is $1 + 2 + 4 + 8 + 16 + 32 + 38 = 101$. Since $f(A)$ has sums of some but not all of the items in A , the integers in $f(A)$ are at most 100. Rather than computing all of the sums, we will explain why $f(A) = [1 : 100]$ in a way that will give some insight for part (c).

We first consider the sums that are achievable without using the integer 38. The other integers in A are $1 = 2^0$, $2 = 2^1$, $4 = 2^2$, $8 = 2^3$, $16 = 2^4$, and $32 = 2^5$. The sums that can be obtained by adding some or all of these integers are exactly the integers from 1 through $63 = 2^6 - 1$. To get an idea of how this works, read about “binary expansions” or “binary representations” of integers. As an example, to represent 53 as a sum of powers of 2, first, find the largest power of 2 that is no larger than 53, which is 32. Then compute $53 - 32 = 21$ to get $53 = 32 + 21$. Now find the largest power of 2 that is no larger than 21, which is 16. Subtract to get $21 - 16 = 5$ or $21 = 16 + 5$. Now substitute to get $53 = 32 + 16 + 5$. Repeating this process, find the largest power of 2 that is no larger than 5, which is 4. Subtracting, $5 - 4 = 1$ so $5 = 4 + 1$, but what remains, 1, is a power of 2, so we get $53 = 32 + 16 + 4 + 1 = 2^5 + 2^4 + 2^2 + 2^0$.

The integers from 1 through 63 are all in $f(A)$. To write the integers from 64 through 100 as a sum of integers in A , notice that $100 - 38 = 62$, and so if $64 \leq m \leq 100$, then $m - 38 \leq 62$. To write such m as a sum of integers from A , compute $r = m - 38 \leq 62 < 63$, write r as a sum of the integers from A other than 38, then include 38 in the sum. For example, to see that 91 is in $f(A)$, compute $r = 91 - 38 = 53$, then use that $53 = 32 + 16 + 4 + 1$ to get that $91 = 1 + 4 + 16 + 32 + 38$.

The only thing left to check is that none of the sums described above require using all seven items in A . To see why this is not a concern, recall that the sum of all items in A is 101, so if we express an integer that is no larger than 100 as a sum of items from A , then it cannot possibly use every item in A .

- (c) The result of part (b) generalizes as follows. For every positive integer n , the minimum length of a list A that compresses $[1 : n]$ is $\lceil \log_2(n + 2) \rceil$. Notice that when $n = 100$, since $100 + 2 = 102$ is strictly between $2^6 = 64$ and $2^7 = 128$, we have $\lceil \log_2(100 + 2) \rceil = 7$, which agrees with the result from part (b).

We will prove that $\lceil \log_2(n + 2) \rceil$ is the minimum length of a list that compresses $[1 : n]$ and, in the process, prove that $[1 : n]$ is always compressible.

To see that $\lceil \log_2(n + 2) \rceil$ is the minimum length of a list that compresses $[1 : n]$ in general, we will show that a list of length less than $\lceil \log_2(n + 2) \rceil$ cannot possibly compress $[1 : n]$, and then we will construct a list of length exactly $\lceil \log_2(n + 2) \rceil$ that does compress $[1 : n]$.

Suppose A has length $k < \lceil \log_2(n+2) \rceil$. Since k and $\lceil \log_2(n+2) \rceil$ are both integers, this implies $k \leq \lceil \log_2(n+2) \rceil - 1$. For every integer x , it is true that $\lceil x \rceil - 1 < x \leq \lceil x \rceil$, so we conclude that $k \leq \lceil \log_2(n+2) \rceil - 1 < \log_2(n+2)$.

From $k < \log_2(n+2)$, we get $2^k < n+2$ or $2^k - 2 < n$. As argued earlier, a list A of length k has the property that there are at most $2^k - 2$ distinct integers in $f(A)$. Since $2^k - 2 < n$, we cannot have $f(A) = [1 : n]$ when A is a list of length $k < \lceil \log_2(n+2) \rceil$ since $[1 : n]$ contains n integers.

We have shown that if A compresses $[1 : n]$, then it must have length at least $\lceil \log_2(n+2) \rceil$. We will now produce a list of length $\lceil \log_2(n+2) \rceil$ that compresses $[1 : n]$. This requires explaining how to produce the list, then showing that the list has the correct length.

Suppose n is a positive integer. Define k to be the largest non-negative integer with the property that $2^k \leq n+1$ and define $m = n+2 - 2^k$. The list A consisting of the powers of 2 from 1 through 2^{k-1} together with m will compress $[1 : n]$ and have length exactly $\lceil \log_2(n+2) \rceil$. Notice that it is possible for m to be equal to one of the powers of 2 from 1 through 2^{k-1} . In this situation, the list A will include two copies of that power of 2.

Before verifying that the list described above does what is required, we will work through a couple of examples.

- When $n = 1$, we observe that $2^0 = 1$ and $2^1 = 2$ are the powers of 2 that are no larger than $n+1 = 2$, and so $k = 1$. Thus, $2^{k-1} = 1$ and $m = n+2 - 2^k = 1+2 - 2 = 1$, so the list is $A = [1, 1]$. Indeed $f([1, 1]) = [1]$.
- When $n = 100$, we get $k = 6$, so $2^{k-1} = 32$ and $m = n+2 - 2^k = 100+2 - 64 = 38$, so the list is $A = [1, 2, 4, 8, 16, 32, 38]$, which is exactly the list from part (b).

We will now show that list A has length $\lceil \log_2(n+2) \rceil$. The integer k is the largest non-negative integer with the property that $2^k \leq n+1$, and A contains $2^0, 2^1$, and so on up to 2^{k-1} , along with the integer m . This gives a total of $k+1$ items in A . Therefore, it suffices to show that with k chosen as described above, we have $\lceil \log_2(n+2) \rceil = k+1$.

The function \log_2 is increasing, meaning that if x and y are positive real numbers with $x < y$, then $\log_2(x) < \log_2(y)$. Using this fact along with the fact that $2^k \leq n+1$, we get that $k \leq \log_2(n+1) < \log_2(n+2)$. As well, k is the largest non-negative integer with the property that $2^k \leq n+1$, which means $n+1 < 2^{k+1}$.

Suppose $k+1 < \log_2(n+2)$. Then $2^{k+1} < n+2$. From above, we also have $n+1 < 2^{k+1}$, so we conclude that $n+1 < 2^{k+1} < n+2$. The quantities $n+1$ and $n+2$ are consecutive integers, so the integer 2^{k+1} cannot lie strictly between them. Therefore, it is impossible for $k+1 < \log_2(n+2)$, which means we must have $\log_2(n+2) \leq k+1$. Combining this inequality with $k < \log_2(n+2)$, we have $k < \log_2(n+2) \leq k+1$, and so we conclude that $\lceil \log_2(n+2) \rceil = k+1$.

It remains to show that A compresses $[1 : n]$. As discussed earlier, since list A contains the items $2^0, 2^1$, and so on up to 2^{k-1} , as well as at least one other item, m , $f(A)$ contains all of the integers from 1 through $2^{k-1+1} - 1 = 2^k - 1$. This is because these integers can be expressed using the sum of some or all of the powers of 2 from 1 through 2^{k-1} . These are exactly the integers that can be expressed without using m .

If we do use m , then we can express every integer from m through $m + 2^k - 1$ as a sum of

items in A . Since $m + 2^k - 1$ is the sum of all items in A , it is excluded from $f(A)$ and so we have that $f(A)$ contains all of the integers from m to $m + 2^k - 2$, and nothing larger. By definition, $m = n + 2 - 2^k$, so $m + 2^k - 2 = (n + 2 - 2^k) + 2^k - 2 = n$.

We have shown that $f(A)$ is the list consisting of the integers that are in $[1 : 2^k - 1]$ or $[m : n]$. To see that $f(A)$ is exactly $[1 : n]$, we need to show that $m \leq 2^k$. There might be overlap corresponding to multiple ways to express some integers in $[1 : n]$ in $f(A)$, but this is allowed.

Suppose $m > 2^k$. Since these quantities are integers, we must have $m \geq 2^k + 1$. By definition, $m = n + 2 - 2^k$, and so we get that $n + 2 - 2^k \geq 2^k + 1$, which can be rearranged to get $n + 1 \geq 2^k + 2^k = 2^{k+1}$. Therefore, we have $2^{k+1} \leq n + 1$, but k was chosen to be the largest integer with $2^k \leq n + 1$, so it is not possible that $2^{k+1} \leq n + 1$. Therefore, our assumption that $m > 2^k$ must be wrong, so we conclude that $m \leq 2^k$, as desired.

- (d) Fix a positive integer $k \geq 3$. We will show that for every integer $m \geq k$, $[m : m + k - 1]$ is not compressible.

To see this, suppose A is a list that compresses $[m : m + k - 1]$. Since $k \geq 3$, there are at least three items in $[m : m + k - 1]$. If A has only two items, then $f(A)$ has at most two integers since the allowable sums are just the two “singleton sums”. Therefore, A also contains at least three items. Note that since m is the smallest integer in $f(A)$, m must also be the smallest integer in A . (If r is in A , then the “singleton sum” r must also be in $f(A)$, so all integers in A must be at least m . Also, since there is nothing smaller than m in A , the only way to produce m in $f(A)$ is by the “singleton sum” m .)

Since $k \geq 3$, $m + 1$ is also in $f(A)$. If $m + 1$ is the sum of at least two items in A , then they must all be smaller than $m + 1$, but the only integer in A that is smaller than $m + 1$ is m . This would mean 1 must be in A , but this is impossible since $1 < k \leq m$ and m is the smallest element in A . Therefore, we must also have $m + 1$ in A , and so both m and $m + 1$ are in A .

Now recall that A has at least three items, so there is at least one element other than m and $m + 1$, and so $m + m + 1 = 2m + 1$ is in $f(A)$. Since $f(A) = [m : m + k - 1]$, we must then have $2m + 1 \leq m + k - 1$, which can be rearranged to get $m \leq k - 2$, which is impossible since $m \geq k$.

Therefore, it is not possible to compress $[m : m + k - 1]$ when $m \geq k$. This means that there are only finitely many m for which $[m : m + k - 1]$ is compressible since all such m must be at most k .

- (e) As mentioned in the hint, the answer is 39. This means $[5 : 39]$ is not compressible, but $[5 : k]$ is compressible for all $k \geq 40$. We will include a sketch of the proof here.

One can check that the lists $A = [5, 6, 7, 8, 9, 10]$, $B = [5, 5, 6, 6, 7, 8, 9]$, $C = [5, 5, 6, 7, 7, 8, 9]$, $D = [5, 5, 6, 7, 8, 8, 9]$, and $E = [5, 5, 6, 7, 8, 9, 9]$ compress the lists $[5 : 40]$, $[5 : 41]$, $[5 : 42]$, $[5 : 43]$, and $[5 : 44]$, respectively.

Now suppose a list A compresses $[5 : k]$ and suppose B is the list A with a 5 added to it. It can be shown that B compresses the list $[5 : k + 5]$. Since $[5 : 40]$ is compressible, this shows that $[5 : 45]$ is compressible. Since $[5 : 41]$ is compressible, $[5 : 46]$ is compressible. Since we have five consecutive values of k for which $[5 : k]$ is compressible ($k = 40$ through

$k = 44$), this reasoning can be used to show that $[5 : k]$ is compressible for all $k \geq 40$. Note that the lists to compress $[5 : k]$ given by this inductive process will not, in general, be as short as possible.

Now suppose a list A compresses $[5 : 39]$. It can be argued using reasoning similar to that from earlier parts that 5 must be in A , 5 is the smallest integer in A , and that the integers 6, 7, 8, and 9 also appear in A . Since the smallest integer in A is 5, the largest sum in $f(A)$ is 5 less than the sum of all items in A . Therefore, the sum of all items in A must be $39 + 5 = 44$. We already have 5, 6, 7, 8, and 9 in A which have a sum of $5 + 6 + 7 + 8 + 9 = 35$, and so the remaining items in A have a sum of $44 - 35 = 9$.

Next, note that using only the five items 5, 6, 7, 8, and 9, a sum of 10 is impossible. The list A cannot contain the integer 10 itself since we already determined that the remaining items in A have a sum of 9. It also cannot include the integers 1, 2, 3, or 4 since 5 is the smallest integer in A . Therefore, the 10 in $f(A)$ must come from the sum of two 5s, which means A must include a second 5.

We now have that A contains (at least) the items 5, 5, 6, 7, 8, and 9, which have a total of 40. Since the sum of all items in A is 44, there must be an additional item in A that is no larger than $44 - 40 = 4$. This is impossible, so we conclude that $[5 : 39]$ is not compressible.



Problem of the Month

Problem 5: February 2024

This problem explores a counting technique that uses *binomial coefficients*. **The main problems for this month are on the next page.** For anyone who is not already familiar with the notation of binomial coefficients, the first page includes an introduction with a few standard exercises.

For non-negative integers n and k , the expression $\binom{n}{k}$ denotes the number of ways to choose k objects from n distinct objects. For this reason, the quantity $\binom{n}{k}$ is pronounced “ n choose k ”.

For example, consider the four letters A , B , C , and D . There are 4 ways to select three of them. They are

$$A, B, C \quad A, B, D \quad A, C, D \quad B, C, D$$

Since there are 4 ways to choose 3 of the objects (order does not matter – it only matters which objects are chosen), we have that $\binom{4}{3} = 4$.

Try the following exercises. Hints will be provided, but solutions will not. These exercises are intended as a way to get familiar with the notation, so they may not be explicitly useful in the problems on the next page.

- (i) Show that $\binom{5}{2} = 10$.
- (ii) What are $\binom{n}{0}$ and $\binom{n}{1}$?
- (iii) Convince yourself that $\binom{n}{k} = \binom{n}{n-k}$ for $0 \leq k \leq n$.
- (iv) Convince yourself that $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$ when $0 \leq k < n$.
- (v) Convince yourself that $\binom{n}{k}$ is equal to the coefficient of x^k in the expanded form of $(1+x)^n$.
This is why $\binom{n}{k}$ is called a “binomial coefficient”.

It turns out that there is a convenient way to compute binomial coefficients using *factorial* notation. If $n \geq k \geq 0$ are integers, then $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. Here, $m!$, pronounced “ m factorial”, is defined so that $m! = 1$ when $m = 0$ and $m = 1$, and $m! = 1 \times 2 \times 3 \times \cdots \times (m-1) \times m$ for integers $m \geq 2$. It is also a useful exercise to think about why this formula for $\binom{n}{k}$ works.



(a) For a positive integer n , show that $\binom{n+1}{2}$ is the answer to each of these three questions. Think about why the answers to all three questions are the same.

(i) What is $1 + 2 + 3 + \cdots + n$?

(ii) How many pairs (x, y) of non-negative integers satisfy $x + y \leq n - 1$?

(iii) How many triples (x, y, z) of non-negative integers satisfy $x + y + z = n - 1$?

(b) Let $n \geq 0$ and $r \geq 1$ be integers. Show that the following two questions have the same answer and express the common answer as a single binomial coefficient.

(i) How many ways are there to place n indistinguishable balls in r distinguishable cups?

(ii) How many non-negative integer solutions are there to the equation $x_1 + x_2 + \cdots + x_r = n$?

(c) Let $n \geq 0$ and $r \geq 1$ be integers. By counting r -tuples of non-negative integers (x_1, x_2, \dots, x_r) that satisfy $x_1 + x_2 + x_3 + \cdots + x_r \leq n$, prove the identity

$$\binom{r-1}{r-1} + \binom{r}{r-1} + \binom{r+1}{r-1} + \binom{r+2}{r-1} + \cdots + \binom{r-1+n}{r-1} = \binom{n+r}{r}$$

Clarification: The quantity on the left in the equation above is the sum of the binomial coefficients $\binom{r-1+m}{r-1}$ where m ranges from 0 to n .

(d) Determine the number of non-negative integers x with $x < 10^{10}$ that have a digit sum of 21

(e) For a positive integer $x = a_n a_{n-1} a_{n-2} \cdots a_1 a_0$ where the a_i are the digits of x , the *alternating sum* of x is the expression $a_0 - a_1 + a_2 - a_3 + \cdots + (-1)^n a_n$. For example, the alternating sum of 744923 is $3 - 2 + 9 - 4 + 4 - 7 = 3$. Determine the number of non-negative integers with at most 8 digits that have an alternating sum equal to 0.

(f) How many ways are there to choose integers a , b , and c such that $1 \leq a < b < c \leq 2024$ and $a + b + c = 2027$?



Problem of the Month

Problem 5: February 2024

Hint

First, some hints on the exercises.

- (i) List the two-element subsets of $\{1, 2, 3, 4, 5\}$.
- (ii) The answers are 1 and n . It might take a moment of reflection to convince yourself that $\binom{n}{0} = 1$ makes sense.
- (iii) If k objects are chosen from a set of n objects, then how many objects are not chosen?
- (iv) If you are to choose $k + 1$ integers from the set $\{1, 2, 3, \dots, n, n + 1\}$, then either $n + 1$ is chosen or it is not.
- (v) The quantity $(1 + x)^n$ is equal to the product of n copies of $(1 + x)$. Try expanding $(1 + x)^n$ for a few small values of n without collecting like terms. As an example, think about how an x^3 term could arise during the expansion of $(1 + x)(1 + x)(1 + x)(1 + x)(1 + x)$.

Below are the hints for the main problems.

- (a) If you have never seen a proof of (a)(i), try writing the sum $1 + 2 + 3 + \dots + n$ in reverse order directly under the sum $1 + 2 + 3 + \dots + n$. Now add each column. For (a)(ii), consider the possible values of x . For (a)(iii), consider the equation $x + y + z = n - 1$ for a fixed pair (x, y) .
 - (b) Imagine arranging the n identical balls in a row and placing $r - 1$ sticks between them. By doing this, you have partitioned the n balls into r smaller groups.
 - (c) Introduce a new variable, x_0 , and consider the equation $x_0 + x_1 + \dots + x_r = n$.
 - (d) The non-negative integers x with $x < 10^{10}$ are exactly the integers that have at most 10 digits. Consider the equation $x_1 + x_2 + \dots + x_{10} = 21$ where $0 \leq x_i \leq 9$.
There was an omission in part (d). The original question said “integers” where it should have said “non-negative integers”.
 - (e) This question can be analyzed by examining the equation $x_1 - x_2 + x_3 - x_4 + x_5 - x_6 + x_7 - x_8 = 0$. Rearrange this equation and use the ideas from (b) and (d).
 - (f) Let $x = a - 1$, $y = b - 1$, and $z = c - 1$. Find the number of non-negative integer solutions to $x + y + z = 2024$ with x , y , and z distinct.
-
-



Problem of the Month

Solution to Problem 5: February 2024

- (a) The binomial coefficient $\binom{n+1}{2}$ is equal to $\frac{(n+1)!}{2!(n-1)!}$, and since $\frac{(n+1)!}{(n-1)!} = (n+1)n$, we get that $\binom{n+1}{2} = \frac{n(n+1)}{2}$. By the well-known formula for the sum of the first n positive integers, we have $1 + 2 + 3 + \cdots + (n-1) + n = \frac{n(n+1)}{2}$, and so the answer to (i) is $\binom{n+1}{2}$.

For (ii), we will consider the possible values of x . Note that since x and y are non-negative and $x + y \leq n - 1$, we must have that $0 \leq x \leq n - 1$.

If $x = 0$ and $x + y \leq n - 1$, then y can be any of the integers from 0 through $n - 1$, for a total of n possibilities.

If $x = 1$, then y can be any of the integers from 0 through $n - 1 - 1 = n - 2$, for a total of $n - 1$ possibilities.

In general, if $x = k$ for some k with $0 \leq k \leq n - 1$, then y can be any integer from 0 through $n - 1 - k$. There are a total of $n - k$ integers from 0 through $n - 1 - k$.

Therefore, the number of pairs is $n + (n - 1) + (n - 2) + \cdots + [n - (n - 1)]$, but this is just the sum $1 + 2 + 3 + \cdots + n$ written in reverse order. By part (i), we conclude that the number of non-negative integer pairs (x, y) with $x + y \leq n - 1$ is $\binom{n+1}{2}$.

For (iii), observe that if x and y are non-negative integers with $x + y \leq n - 1$, then there is exactly one non-negative integer z for which $x + y + z = n - 1$. Therefore, the number of non-negative integer pairs (x, y) with $x + y \leq n - 1$ is exactly the same as the number of non-negative integer triples (x, y, z) with $x + y + z = n - 1$.

- (b) Suppose the r distinguishable cups are labelled Cup 1, Cup 2, and so on, up to Cup r . Suppose the n balls are placed in the r cups. If we let x_k be equal to the number of balls in Cup k (noting that it is possible that $x_k = 0$), then we will have $x_1 + x_2 + \cdots + x_r = n$ since there are n balls in total. On the other hand, if we have a non-negative integer solution (x_1, x_2, \dots, x_r) to $x_1 + x_2 + \cdots + x_r = n$, then if we place x_k balls in Cup k for each k , then we will have distributed exactly n balls among the r cups. Finally, observe that every placement of n balls in the r cups leads to a distinct solution to $x_1 + x_2 + \cdots + x_r = n$, and every solution to this equation gives a distinct way to distribute n balls in the r cups. This shows that the answers to questions (i) and (ii) are equal.

We will now show that the answer to question (ii) (and hence, the answer to question (i)), is $\binom{n+r-1}{r-1}$.

We need to reimagine the distribution of balls into cups as a choice of some objects. To give an idea of how it works, we will focus on the special case with $n = 10$ and $r = 4$. That is, we want to count the number of ways to place 10 indistinguishable balls in 4 distinguishable cups.

Imagine laying the 10 balls in a row. We want to place them among 4 cups, which means we want to break the 10 balls into 4 groups (where some of the groups could be empty). We can do this by placing three partitions somewhere among the 10 balls that are now

lined up in a row. For example, the partitions could be placed as shown in the diagram below. Each ball is represented by a circle and each partition is represented by a vertical line.



This corresponds to placing 2 balls in Cup 1, 3 balls in Cup 2, 1 ball in Cup 3, and 4 balls in Cup 4. For a different example, we could place the partitions as follows.



In this case, two of the partitions are next to each other with no balls between them. This corresponds to the third cup having no balls in it.

In fact, any arrangement of 10 indistinguishable balls and 3 indistinguishable partitions will correspond to an ordered list of four non-negative integers that have a sum of 10.

Therefore, there are $10 + 3 = 13$ positions where the balls and partitions are to be placed, and every way to choose three of the positions to be partitions corresponds to an ordering of 10 balls and 3 partitions. Therefore, the answer to the question in this case is $\binom{13}{3}$.

In general, if there are r cups, then we need to use $r - 1$ partitions. This means there will be n balls and $r - 1$ partitions for a total of $n + r - 1$ objects. There are $r - 1$ positions that need to be chosen for the partitions. Thus, the number of ways to place n indistinguishable balls in r distinguishable cups is $\binom{n+r-1}{r-1}$.

- (c) First note that non-negative integers x_1, x_2, \dots, x_r satisfy $x_1 + x_2 + \dots + x_r \leq n$ if and only if $x_1 + x_2 + \dots + x_r = m$ for some integer m with $0 \leq m \leq n$. From part (b), the number of non-negative integer solutions to $x_1 + x_2 + \dots + x_r = m$ is $\binom{r-1+m}{r-1}$. Therefore, the number of non-negative integer solutions to $x_1 + x_2 + \dots + x_r \leq n$ is

$$\begin{aligned} & \binom{r-1+0}{r-1} + \binom{r-1+1}{r-1} + \binom{r-1+2}{r-1} + \dots + \binom{r-1+n}{r-1} \\ &= \binom{r-1}{r-1} + \binom{r}{r-1} + \binom{r+1}{r-1} + \dots + \binom{r+n-1}{r-1} \end{aligned}$$

Now suppose x_0, x_1, \dots, x_r are non-negative integers with $x_0 + x_1 + \dots + x_r = n$. Then we must have $x_1 + x_2 + \dots + x_r \leq n$. On the other hand, if $x_1 + x_2 + \dots + x_r \leq n$, then there is a unique non-negative integer x_0 that satisfies $x_0 + x_1 + \dots + x_r = n$. This shows that the number of non-negative integer solutions to $x_0 + x_1 + \dots + x_r = n$ is the same as the number of non-negative integer solutions to $x_1 + x_2 + \dots + x_r \leq n$. From part (b), the number of solutions to $x_0 + x_1 + \dots + x_r = n$ is $\binom{n+(r+1)-1}{(r+1)-1} = \binom{n+r}{r}$.

We have shown that the number of non-negative integer solutions to the inequality $x_1 + x_2 + \dots + x_r \leq n$ is equal to

$$\binom{r-1}{r-1} + \binom{r}{r-1} + \binom{r+1}{r-1} + \dots + \binom{r+n-1}{r-1}$$

and it is also equal to $\binom{n+r}{r}$, so we conclude that

$$\binom{r-1}{r-1} + \binom{r}{r-1} + \binom{r+1}{r-1} + \cdots + \binom{r+n-1}{r-1} = \binom{n+r}{r}$$

This identity is sometime called the *Hockey Stick Identity*. If you are interested in knowing why it has such a strange name, read about Pascal's Triangle and try to interpret this result using Pascal's Triangle.

- (d) The integer k having the property that $k < 10^{10}$ is the same as the integer k having at most 10 decimal digits. Thus, every such integer takes the form $k = a_9 a_8 a_7 \cdots a_1 a_0$ where $a_9, a_8, a_7, \dots, a_1, a_0$ are the digits of k , but some of the leading digits may be equal to 0. For example, to represent the integer 19723 in this way, we would write it as 0000019723 so $a_9 = a_8 = a_7 = a_6 = a_5 = 0$, $a_4 = 1$, $a_3 = 9$, $a_2 = 7$, $a_1 = 2$, and $a_0 = 3$.

Thus, to count the non-negative integers with a digit sum of 21, we need to find all non-negative integer solutions to the equation $a_9 + a_8 + a_7 + \cdots + a_1 + a_0 = 21$, but we have an additional restriction that each a_i is a digit, meaning $0 \leq a_i \leq 9$.

Switching notation slightly to match earlier parts, the answer to the question is equal to the number of solutions to the equation $x_1 + x_2 + \cdots + x_{10} = 21$ where each x_i is an integer with $0 \leq x_i \leq 9$.

By part (b), the number of non-negative integer solutions to $x_1 + x_2 + \cdots + x_{10} = 21$ without the restriction that $x_i \leq 9$ is $\binom{21+10-1}{10-1} = \binom{30}{9}$. The total of $\binom{30}{9}$ includes some solutions that violate $x_i \leq 9$ for some i . We will now count the solutions that have $x_i \geq 10$ for at least one i so we can remove these from the count.

First, we claim that for each index i , there are $\binom{20}{9}$ solutions that have $x_i \geq 10$. For example, consider a non-negative integer solution to $x_1 + x_2 + \cdots + x_{10} = 21$ with $x_1 \geq 10$. If we set $y_1 = x_1 - 10$ then y_1 is a non-negative integer and

$$y_1 + x_2 + \cdots + x_{10} = (x_1 - 10) + x_2 + \cdots + x_{10} = (x_1 + x_2 + \cdots + x_{10}) - 10 = 21 - 10 = 11$$

Also, if we consider a non-negative integer solution to $y_1 + x_2 + \cdots + x_{10} = 11$, and set $x_1 = y_1 + 10$, then we obtain a non-negative integer solution to $x_1 + x_2 + \cdots + x_{10} = 21$ with $x_1 \geq 10$. By part (b), there are $\binom{20}{9}$ non-negative integer solutions to $y_1 + x_2 + \cdots + x_{10} = 11$ and so, equivalently, there are $\binom{20}{9}$ non-negative integer solutions to $x_1 + x_2 + \cdots + x_{10} = 21$ with $x_1 \geq 10$.

Similarly, for each i from 1 through 10, there are exactly $\binom{20}{9}$ non-negative integer solutions to $x_1 + x_2 + \cdots + x_{10} = 21$ with $x_i \geq 10$. Note that these different sets of solutions have overlap and so the number of solutions with $x_i \geq 10$ for at least one i is not $10\binom{20}{9}$. This is an *overcount*. For example, the solution $x_1 = 10$, $x_2 = 10$, $x_3 = 1$, and $x_4 = x_5 = x_6 = x_7 = x_8 = x_9 = x_{10} = 0$ is double counted: once as a solution with $x_1 \geq 10$ and once as a solution with $x_2 \geq 10$.

Next, we claim that for every pair of indices i and j , with $i \neq j$, there are $\binom{10}{9}$ non-negative integer solutions to $x_1 + x_2 + \cdots + x_{10} = 21$ with $x_i \geq 10$ **and** $x_j \geq 10$. Each of these solutions will be *double counted* in the count $10\binom{20}{9}$. Note that there are no solutions to $x_1 + x_2 + \cdots + x_{10} = 21$ that have more than two variables exceeding 9 and so this will give us the complete picture.

For an example, consider a non-negative integer solution to $x_1 + x_2 + x_3 + \cdots + x_{10} = 21$ with $x_1 \geq 10$ and $x_2 \geq 10$. If we set $y_1 = x_1 - 10$ and $y_2 = x_2 - 10$ then y_1 and y_2 are non-negative integers satisfying $y_1 + y_2 + x_3 + \cdots + x_{10} = 1$. Similarly, a non-negative integer solution to $y_1 + y_2 + x_3 + \cdots + x_{10} = 1$ corresponds to a non-negative integer solution to $x_1 + x_2 + x_3 + \cdots + x_{10} = 21$ with $x_1, x_2 \geq 10$. By part (b), there are $\binom{10}{9}$ non-negative integer solutions to $y_1 + y_2 + x_3 + \cdots + x_{10} = 1$ and so, equivalently, there are $\binom{10}{9}$ non-negative integer solutions to $x_1 + x_2 + \cdots + x_{10} = 21$ with $x_1 \geq 10$ and $x_2 \geq 10$.

Similarly, for each pair i and j , with $i \neq j$, there are $\binom{10}{9}$ solutions to $x_1 + x_2 + \cdots + x_{10} = 21$ with $x_i \geq 10$ and $x_j \geq 10$.

Since there are $\binom{10}{2}$ ways to choose two indices, there are $\binom{10}{2} \binom{10}{9} = 45 \binom{10}{9}$ non-negative integer solutions that have two variables exceeding 9. From this, we conclude that the number of non-negative integer solutions to $x_1 + x_2 + \cdots + x_{10} = 21$ that have $x_i \geq 10$ for at least one i is given by $10 \binom{20}{9} - 45 \binom{10}{9}$.

Therefore, the total number of non-negative integer solutions to $x_1 + x_2 + \cdots + x_{10} = 21$ with $x_i \leq 9$ for all i can be calculated as follows:

$$\begin{aligned} \binom{30}{9} - 10 \binom{20}{9} + 45 \binom{10}{9} &= 14\,307\,150 - 10(167\,960) + 45(10) \\ &= 14\,307\,150 - 1\,679\,600 + 450 \\ &= 12\,628\,000 \end{aligned}$$

- (e) By the same convention as the previous problem, we can recognize every positive integer less than 10^8 as an 8-digit integer by possibly prepending some 0s. Thus, we wish to count the number of solutions to the equation $x_1 - x_2 + x_3 - x_4 + x_5 - x_6 + x_7 - x_8 = 0$ where the x_i are integers with $0 \leq x_i \leq 9$ for each i .

We can rearrange this equation to get $x_1 + x_3 + x_5 + x_7 = x_2 + x_4 + x_6 + x_8$, so we want to count non-negative integer solutions to this equation where $x_i \leq 9$ for each i .

Since $0 \leq x_i \leq 9$, we must have that $0 \leq x_1 + x_3 + x_5 + x_7 \leq 36$ and $0 \leq x_2 + x_4 + x_6 + x_8 \leq 36$. For each n from 0 through 36, let A_n denote the number of non-negative integer solutions to $x_1 + x_3 + x_5 + x_7 = n$ with $x_i \leq 9$ for $i = 1, 3, 5, 7$. Since all of the 8 variables have the exact same restrictions, we also have that A_n is the number of non-negative integer solutions to $x_2 + x_4 + x_6 + x_8 = n$ where each variable is no larger than 9. Thus, for each n from 0 through 36, there are A_n^2 solutions to the equation with each side equal to n . Therefore, the answer to the question is

$$A_0^2 + A_1^2 + A_2^2 + \cdots + A_{36}^2$$

Now suppose that $0 \leq n \leq 36$ and $x_1 + x_3 + x_5 + x_7 = n$ with $0 \leq x_i \leq 9$ for each i . Then $(9 - x_1) + (9 - x_3) + (9 - x_5) + (9 - x_7) = 36 - n$, and observe that since $0 \leq x_i \leq 9$, we have $0 \leq 9 - x_i \leq 9$, and since $0 \leq n \leq 36$, we have $0 \leq 36 - n \leq 36$. You should convince yourself that this shows that the number of solutions to $x_1 + x_3 + x_5 + x_7 = n$ is the same as the number of solutions to $x_1 + x_3 + x_5 + x_7 = 36 - n$. In other words, $A_n = A_{36-n}$. When $n = 18$, we also have $36 - n = 18$, so the total we seek is equal to

$$2(A_0^2 + A_1^2 + A_2^2 + \cdots + A_{17}^2) + A_{18}^2$$

For $n = 0$ through $n = 9$, if $x_1 + x_3 + x_5 + x_7 = n$, then $x_i \leq 9$ for each i since the total is at most 9. This means the number of solutions to $x_1 + x_3 + x_5 + x_7 = n$ where $x_i \leq 9$

for each i is equal to the number of non-negative integer solutions without restriction. By part (b), if $0 \leq n \leq 9$, then $A_n = \binom{n+3}{3}$.

For $n = 10$ through $n = 18$, there are still $\binom{n+3}{3}$ unrestricted solutions, but here the count includes solutions with $x_i \geq 10$ for some i . Note that since the total is at most 18, it is not possible for more than one variable to exceed 9. Similar to the argument in part (d), for each i , there are $\binom{n+3-10}{3}$ solutions with $x_i \geq 10$. Since there are 4 variables, there are $4\binom{n+3-10}{3}$ solutions with a variable exceeding 9. Therefore, $A_n = \binom{n+3}{3} - 4\binom{n+3-10}{3}$.

We now just need to compute the total.

$$\begin{aligned}
& 2(A_0^2 + \cdots + A_9^2) + 2(A_{10}^2 + \cdots + A_{17}^2) + A_{18}^2 \\
&= 2\left(\binom{3}{3}^2 + \cdots + \binom{12}{3}^2\right) + 2\left(\left(\binom{13}{3} - 4\binom{3}{3}\right)^2 + \cdots + \left(\binom{20}{3} - 4\binom{10}{3}\right)^2\right) + \left(\binom{21}{3} - 4\binom{11}{3}\right)^2 \\
&= 2(1^2 + 4^2 + 10^2 + 20^2 + 35^2 + 56^2 + 84^2 + 120^2 + 165^2 + 220^2) \\
&\quad + 2(79524 + 121104 + 172225 + 230400 + 291600 + 350464 + 400689 + 435600) \\
&\quad + 448900 \\
&= 4816030
\end{aligned}$$

- (f) Let $x = a - 1$, $y = b - 1$, and $z = c - 1$. Then we have $0 \leq x < y < z \leq 2023$ and $x + y + z = 2024$. Thus, we want to count the number of non-negative integer solutions to $x + y + z = 2024$ where $x < y < z \leq 2023$. Suppose T is the number of non-negative integer solutions to $x + y + z = 2024$ where x , y , and z are all distinct and no larger than 2023. Since there are six different arrangements of three distinct integers, exactly one of which puts them in increasing order, the answer to the given question is $\frac{1}{6}T$.

To compute T , we note that by part (b) there are $\binom{2026}{2}$ non-negative integer solutions to $x + y + z = 2024$ where there are no restrictions on the non-negative integers x , y , and z . We now need to remove those solutions where at least two of the variables are equal as well as any that have at least one of x , y , and z greater than 2023.

If one of x , y , and z is greater than 2023, then the condition $x + y + z = 2024$ and the fact that x , y , and z are non-negative integers implies that one of the variables equals 2024 and the other two equal 0. Thus, any solutions with a variable greater than 2023 will also have two variables equal to each other, so we can just count the number of solutions with at least two variables equal and subtract this total from $\binom{2026}{2}$. Since 2024 is not a multiple of 3, it is impossible that $x = y = z$. Therefore, we need to count the number of solutions where exactly two variables are equal.

To count the number of solutions to $x + y + z = 2024$ with two variables equal, we will count the number of solutions with $x = y$ and triple the result since there are three choices of which two variables are equal.

We now want to count non-negative integer solutions to $2x + z = 2024$. This equation implies that z is even, or that $z = 2w$ for some non-negative integer w . Thus, we need to count the number of solutions to $2x + 2w = 2024$ or $x + w = 1012$. By part (b), this is 1013, so the number of solutions to the equation with two variables equal to each other is $3 \times 1013 = 3039$.

We now have that $T = \binom{2026}{2} - 3039$, and so the number of solutions is

$$\frac{1}{6} \left(\binom{2026}{2} - 3039 \right) = \frac{1}{6} (1013 \times 2025 - 3039) = 1013(337) = 341381$$



Problem of the Month

Problem 6: March 2024

This month's problem is an introduction to some of the basic ideas that come up when studying polynomials. It is presented with many exercises interspersed among some definitions. Some of the exercises are computational and some are theoretical.

Complex Numbers: The polynomial $x^2 + 1$ does not have any real roots. Because of this, we “invent” a root, and it is traditionally called i for “imaginary”. That is, we declare that i is a number such that $i^2 + 1 = 0$ or $i^2 = -1$. From a desire to do arithmetic with numbers “like” i , we declare that a *complex number* is a number of the form $a + bi$ where a and b are real numbers. For example, $1 + i$ and $-3 + 2i$ are complex numbers. Remembering that $i^2 = -1$ and using expected arithmetic rules, we can add and multiply complex numbers. For example,

$$\begin{aligned}(1 + i) + (-3 + 2i) &= 1 + i - 3 + 2i = -2 + 3i \\(1 + i)(-3 + 2i) &= (1)(-3) + (1)(2i) + (i)(-3) + (i)(2i) \\ &= -3 + 2i - 3i + 2i^2 \\ &= -3 + 2i - 3i + 2(-1) \\ &= -5 - i\end{aligned}$$

Using the quadratic formula, complex numbers give roots to all real quadratics. For example, the quadratic $f(x) = x^2 - 4x + 13$ has a discriminant of $(-4)^2 - 4(13) = -36$, and so it has no real roots. However, if we accept that $\sqrt{-36}$ means “a number that squares to -36 ”, we can infer that $6i$ or $-6i$ could reasonably be considered as $\sqrt{-36}$. Indeed, using the quadratic formula, we get

$$\frac{4 \pm \sqrt{-36}}{2} = \frac{4 \pm 6i}{2} = 2 \pm 3i$$

If you are new to complex numbers, or you just want to have some fun, you might want to check that $f(2 + 3i) = 0$ and $f(2 - 3i) = 0$ using the methods for adding and multiplying complex numbers demonstrated above.

For the next three exercises, it will be useful to remember that if r is a root of a polynomial $p(x)$, then $x - r$ is a factor of $p(x)$.

- Find all roots of the polynomial $f(x) = x^4 - 2x^3 - 2x^2 + 8$.
- Find all roots of the polynomial $g(x) = x^4 + 2x^2 + 1$.
- Find all roots of the polynomial $h(x) = x^4 - 4$.

Definition: A polynomial $p(x)$ is called a *rational polynomial* if all of its coefficients are rational numbers. The set of rational polynomials is denoted by $\mathbb{Q}[x]$.

Definition: Suppose $p(x) \in \mathbb{Q}[x]$ has degree at least 1. We say that $p(x)$ is *reducible* in $\mathbb{Q}[x]$ (or *reducible over* \mathbb{Q}) if there are rational polynomials $u(x)$ and $v(x)$, both of degree at least 1, such that $p(x) = u(x)v(x)$. We say that $p(x)$ is *irreducible over* \mathbb{Q} if it is not reducible over \mathbb{Q} . In other



words, $p(x)$ is irreducible over \mathbb{Q} if it cannot be factored as a product of rational polynomials of degree lower than that of $p(x)$.

(d) Determine whether each given polynomial is reducible or irreducible over \mathbb{Q} .

(i) $x^2 - 2$

(iii) $x^3 + x + 1$

(v) $x^4 + 3x^2 + 2$

(ii) $x^3 - 6x^2 + 11x - 6$

(iv) $2x - 5$

(vi) $x^4 + 1$

Definition: A complex number α is called *algebraic* if it is a root of some non-constant rational polynomial. That is, α is algebraic if there is a polynomial $p(x) \in \mathbb{Q}[x]$ of degree at least 1 such that $p(\alpha) = 0$. The *degree* of the algebraic number α is the smallest positive integer d such that there is a rational polynomial $p(x)$ of degree d with $p(\alpha) = 0$. [The word “positive” is important because every number α is a root of the constant 0 polynomial.]

(e) Show that $\sqrt{2}$, $1 + \sqrt[3]{2}$, $1 + 2i$, and $\sqrt{2} + \sqrt{3}$ are all algebraic numbers and find their degrees. It might be easier to find the degrees after thinking about some of the later questions.

(f) Suppose α is an algebraic number of degree d . Prove that there is a unique irreducible polynomial $m(x)$ of degree d with leading coefficient equal to 1 such that $m(\alpha) = 0$.

Note: Numbers that are not algebraic are called *transcendental* numbers. Two famous examples of transcendental numbers are π and e .

Definition: If $f(x)$ is a polynomial and α is a root of $f(x)$, then α is a *repeated root* of $f(x)$ if there is a polynomial $g(x)$ such that $f(x) = g(x)(x - \alpha)^2$. Note that α might be a complex number, which means $g(x)$ could have complex coefficients.

(g) Let $f(x) = x^7 - 3x^6 + 2x^5 - 2x^4 + x^3 + 5x^2 + 4$. Find complex numbers r_1, r_2, \dots, r_7 such that $f(x) = (x - r_1)(x - r_2) \cdots (x - r_7)$. In other words, find all roots of $f(x)$.

Hint: Some of the roots are small integers and every root corresponds a linear factor.

(h) Given a polynomial $f(x)$ of degree n , if we expand the expression $f(x + y)$ (here, y is another variable), we can write $f(x + y)$ as

$$f(x + y) = f_0(x) + yf_1(x) + y^2f_2(x) + \cdots + y^n f_n(x)$$

for unique polynomials $f_0(x), f_1(x), \dots, f_n(x)$.

(i) For the polynomial $f(x)$ from part (g), determine the polynomials $f_0(x)$ and $f_1(x)$ described above and find all roots of $f_0(x)$ and $f_1(x)$.

(ii) Suppose that $f(x)$ is a polynomial and that r is a root of $f(x)$. Prove that r is a repeated root of $f(x)$ if and only if r is a root of $f_1(x)$.

(j) Prove that if $p(x)$ and $q(x)$ are irreducible rational polynomials with a root in common, then there is a rational number c such that $p(x) = cq(x)$.

(k) Prove that an irreducible rational polynomial cannot have a repeated root.

Note: The *division algorithm for polynomials* might be useful in some of the later problems. It will be explained briefly in the hint.



Problem of the Month

Problem 6: March 2024

Hint

- (a) The *Rational Root Theorem* could come in handy: If a polynomial $p(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ has integer coefficients and r is a rational root of $p(x)$, then it must be of the form $r = \frac{u}{v}$ where u and v are integers, u is a divisor of a_0 , and v is a divisor of a_n .

For the polynomial in (a), this means the only possible rational roots are ± 1 , ± 2 , ± 4 , and ± 8 , the divisors of 8 (the leading coefficient is 1).

- (b) This polynomial is a perfect square.
- (c) This polynomial has no rational roots, but it does have a real root that is easy to find.
- (d) A polynomial has a rational root if and only if it has a rational linear factor. If $p(x)$ and $q(x)$ are polynomials of degree m and n , then what is the degree of $p(x)q(x)$?
- (e) To show that a number is algebraic, you need to find a rational polynomial with that number as a root. For $1 + \sqrt[3]{2}$, let $r = 1 + \sqrt[3]{2}$ so that $r - 1 = \sqrt[3]{2}$. Now cube both sides.
- (f) If $p(x)$ has degree d and factors as the product of two polynomials of degree at least 1, then both of these polynomials have degree less than d . Read the definition of “degree” carefully.

For uniqueness, suppose that two polynomials, $p(x)$ and $q(x)$, have the described properties. What can be said about the degree of their difference?

- (h) (i) Warm up by trying this with a polynomial of lower degree. It turns out that the polynomial $f_1(x)$ is the *derivative* of $f(x)$. This is not important for the problem, but it is interesting.
- (ii) If you know some calculus, then there is a nice proof of this involving the product rule. Otherwise, if $f(x) = (x - r)^2p(x)$, then $f(x + y) = [(x - r) + y]^2p(x + y)$. Expand $p(x + y)$ and $f(x + y)$ as described in part (i) and compare “coefficients” of y .
- (j) By definition, the shared root is algebraic and so has a minimal polynomial, $m(x)$. Show that each of $p(x)$ and $q(x)$ is a scalar multiple of $m(x)$.

Division Algorithm for Polynomials: For polynomials $f(x)$ and $g(x)$ with $g(x)$ not the zero polynomial, there exist unique polynomials $h(x)$ and $r(x)$ such that

$$f(x) = h(x)g(x) + r(x)$$

where the degree of $r(x)$ is less than the degree of $g(x)$. This is essentially the result of doing polynomial or synthetic division of $f(x)$ by $g(x)$.

- (k) Convince yourself that every polynomial can be expressed as a product of irreducible polynomials. As well, as long as $f(x)$ has degree at least 1, the polynomial $f_1(x)$ is not the zero polynomial and has degree less than that of $f(x)$ (in fact, the degree is one less).



Problem of the Month

Solution to Problem 6: March 2024

- (a) Notice that $f(2) = 2^4 - 2(2)^3 - 2(2)^2 + 8 = 16 - 16 - 8 + 8 = 0$, so 2 is a root of $f(x)$. This means $(x - 2)$ is a factor of $f(x)$. Factoring gives $f(x) = (x - 2)(x^3 - 2x - 4)$. Now evaluating $2^3 - 2(2) - 4 = 8 - 4 - 4 = 0$, we get that 2 is a root of $x^3 - 2x - 4$, so $(x - 2)$ is a factor of $x^3 - 2x - 4$. Factoring gives $x^3 - 2x - 4 = (x - 2)(x^2 + 2x + 2)$.

Applying the quadratic formula to $x^2 + 2x + 2$ gives

$$\frac{-2 \pm \sqrt{2^2 - 4(2)}}{2} = \frac{-2 \pm \sqrt{-4}}{2} = \frac{-2 \pm 2i}{2} = -1 \pm i$$

So the roots of $f(x)$ are 2, $-1 + i$, and $-1 - i$ with 2 appearing as a repeated root.

- (b) This polynomial has no real roots, but notice that $g(x) = x^4 + 2x^2 + 1 = (x^2 + 1)^2$. The roots of $x^2 + 1$ are $\pm i$, so the only roots of $g(x)$ are i and $-i$. In fact, $g(x)$ can be factored as $g(x) = (x - i)^2(x + i)^2$ which shows that each of these two roots are repeated roots.
- (c) We can use a difference of square to get that $h(x) = (x^2 + 2)(x^2 - 2)$. The roots of $x^2 - 2$ are $\sqrt{2}$ and $-\sqrt{2}$. The roots of $x^2 + 2$ are $i\sqrt{2}$ and $-i\sqrt{2}$. Thus, $h(x)$ has four distinct roots (no repeated roots), two of which are real and two of which are not real.

Notation: In several of the remaining solutions, we will use the word *monic* to describe a polynomial with a leading coefficient of 1. One of the main observations about monic polynomials is that if $p(x)$ is a polynomial with leading coefficient $a \neq 0$, then $\frac{1}{a}p(x)$ is a monic polynomial of the same degree with exactly the same roots as $p(x)$. We will leave the following fact as an exercise: If $p(x)$ is a reducible monic polynomial, then $p(x)$ factors as the product of two monic polynomials of degree at least 1.

- (d) (i) **The polynomial is irreducible.** Suppose $x^2 - 2$ is reducible. Since the polynomial is monic, there must be monic rational polynomials $p(x)$ and $q(x)$ of degree at least 1 such that $p(x)q(x) = x^2 - 2$. Since $p(x)$ and $q(x)$ both have degree at least 1 and their product has degree 2, they must both have degree exactly 1.

Therefore, there are rational numbers a and b such that $p(x) = x + a$ and $q(x) = x + b$, so $x^2 - 2 = (x + a)(x + b)$. Expanding, we get $x^2 - 2 = x^2 + (a + b)x + ab$. Comparing coefficients, $a + b = 0$ or $a = -b$, and $ab = -2$. Substituting $a = -b$ into $ab = -2$ gives $-b^2 = -2$ or $b^2 = 2$. It is well known that no rational number b has the property that $b^2 = 2$, and so there is a problem. We conclude that $x^2 - 2$ cannot be factored into the product of two rational polynomials both with positive degree.

Observation: It is important to observe that we have specifically shown that $x^2 - 2$ does not factor over \mathbb{Q} . If we allow for any real coefficients, we easily get that $x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$ which is a perfectly good factorization into a product of polynomials with real coefficients. Extending the language defined in the problem, we would say that while $x^2 - 2$ is irreducible over \mathbb{Q} , it is reducible over \mathbb{R} .

- (ii) **The polynomial is reducible.** Checking for rational roots and then factoring, one finds that $x^3 - 6x^2 + 11x - 6 = (x - 1)(x - 2)(x - 3)$.

- (iii) **The polynomial is irreducible.** If a cubic polynomial is equal to the product of polynomials $p(x)$ and $q(x)$ each with degree at least 1, then one of them must be linear and the other must be quadratic. This is because $1 + 2$ is the only way to express 3 (the degree) as the sum of positive integers. Since $x^3 + x + 1$ is monic, there must be rational numbers a , b , and c so that $x^3 + x + 1 = (x + a)(x^2 + bx + c)$. Of course, this shows that $-a$ is a rational root, so we get the following useful fact that is special to cubics (and quadratics): A cubic polynomial is reducible over \mathbb{Q} if and only if it has a rational root.

By the rational root theorem, 1 and -1 are the only candidates for a rational root of $x^3 + x + 1$. Neither is a root, so the polynomial has no rational roots. Therefore, by the argument above, the polynomial is irreducible.

- (iv) **The polynomial is irreducible.** Every linear polynomial is irreducible. This is because the product of two polynomials of positive degree must have degree at least 2, so a polynomial of degree less than 2 cannot possibly be expressed as the product of two polynomials of positive degree.
- (v) **The polynomial is reducible.** Observe that $x^4 + 3x^2 + 2 = (x^2 + 1)(x^2 + 2)$, and so $x^4 + 3x^2 + 2$ can be expressed as the product of two rational polynomials of positive degree.

The roots of the polynomial are i , $-i$, $i\sqrt{2}$ and $-i\sqrt{2}$, none of which are real. In part (iii) above, it was noted that a rational cubic is irreducible over \mathbb{Q} if and only if it has a rational root. This example shows that this is special property of polynomials of low degree since $x^4 + 3x^2 + 2$ is reducible, but it does not have any rational roots.

- (vi) **The polynomial is irreducible.** If $x^4 + 1 = 0$, then $(x^2)^2 = -1$, which is impossible for a real number x . Therefore, $x^4 + 1$ has no rational roots (since it has no real roots).

If $x^4 + 1$ factors as the product of two rational polynomials of positive degree, then it must be the product of a linear with a cubic, or the product of two quadratics. The polynomial has no rational root, so it has no linear factor, which means the only remaining possibility is that $x^4 + 1$ is the product of two rational quadratics. As mentioned earlier, if a monic polynomial factors, then it factors as the product monic polynomials.

We will assume that a , b , c , and d are real numbers such that $x^4 + 1 = (x^2 + ax + b)(x^2 + cx + d)$ and deduce that these numbers cannot all be rational.

Expanding, we have

$$x^4 + 1 = x^4 + (a + c)x^3 + (ac + b + d)x^2 + (ad + bc)x + bd$$

and by comparing coefficients, we get

$$a + c = 0 \tag{1}$$

$$ac + b + d = 0 \tag{2}$$

$$ad + bc = 0 \tag{3}$$

$$bd = 1 \tag{4}$$

From Equation (1), we get $a = -c$ and so we can substitute into Equations (2) and (3) above to get

$$-c^2 + b + d = 0 \tag{2'}$$

$$-cd + bc = 0 \tag{3'}$$

If $c = 0$, then Equation (2') implies $b = -d$, and substituting into Equation (4) gives $-d^2 = 1$ or $d^2 = -1$. This means $d = i$ or $d = -i$, neither of which is rational.

If $c \neq 0$, then we can divide through by c in Equation (3') to get $-d + b = 0$ or $b = d$. Equation (4) now implies that $b^2 = d^2 = 1$, so either $b = d = 1$ or $b = d = -1$.

If $b = d = 1$, then Equation (2') gives $-c^2 + 2 = 0$ or $c^2 = 2$, so $c = \pm\sqrt{2}$. Either way, c is irrational.

If $b = d = -1$, then Equation (2') gives $c^2 = -2$ and so $c = \pm i\sqrt{2}$, both of which are irrational.

We have exhausted all possibilities and deduced that at least one of a , b , c , and d must be irrational in all cases, so we have shown that $x^4 + 1$ cannot possibly factor as the product of two rational quadratic polynomials.

If you look a bit more closely at the case work above, it actually shows that $x^4 + 1$ has the following three different factorizations:

$$x^4 + 1 = (x^2 + i)(x^2 - i) = (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1) = (x^2 - i\sqrt{2}x - 1)(x^2 + i\sqrt{2}x - 1)$$

but none of them are factorizations into rational polynomials.

- (e) The real number $\sqrt[3]{2}$ is algebraic because it is a root of the rational polynomial $x^3 - 2$.

To find a rational polynomial to which $1 + \sqrt[3]{2}$ is a root, we write $r = 1 + \sqrt[3]{2}$ and rearrange to get $r - 1 = \sqrt[3]{2}$. Cubing both sides, we get $r^3 - 3r^2 + 3r - 1 = 2$. We can rearrange this to see that $r^3 - 3r^2 + 3r - 3 = 0$, which shows that $r = 1 + \sqrt[3]{2}$ is a root of the polynomial $x^3 - 3x^2 + 3x - 3$, which is a rational cubic.

Let $\alpha = 1 + 2i$ and $\beta = 1 - 2i$. Notice that $\alpha + \beta = 2$ and $\alpha\beta = (1 + 2i)(1 - 2i) = 1^2 - (2i)^2 = 5$. The polynomial $(x - \alpha)(x - \beta) = x^2 - (\alpha + \beta)x + \alpha\beta = x^2 - 2x + 5$, a rational quadratic, has $\alpha = 1 + 2i$ as a root by construction.

For $\sqrt{2} + \sqrt{3}$, we will use a similar trick to that which was used for $1 + \sqrt[3]{2}$. Set $r = \sqrt{2} + \sqrt{3}$ and rearrange to get $r - \sqrt{2} = \sqrt{3}$. Squaring both sides gives $r^2 - 2\sqrt{2}r + 2 = 3$ which can be rearranged to get $r^2 - 1 = 2\sqrt{2}r$. Squaring both sides again gives $r^4 - 2r^2 + 1 = 8r^2$, which can be rearranged to $r^4 - 10r^2 + 1 = 0$. Therefore, $\sqrt{2} + \sqrt{3}$ is a root of the rational polynomial $x^4 - 10x^2 + 1$.

The degrees of $\sqrt{2}$, $1 + \sqrt[3]{2}$, $1 + 2i$, and $\sqrt{2} + \sqrt{3}$ are 2, 3, 2, and 4, respectively. We will justify this at the end of the solution to part (j).

- (f) Suppose α is an algebraic number of degree d . It follows from the remark before the solution to part (d) that there is a rational monic polynomial $p(x)$ of degree d such that $p(\alpha) = 0$.

Suppose $p(x)$ is reducible over \mathbb{Q} . Then there are rational polynomials $f(x)$ and $g(x)$ such that $p(x) = f(x)g(x)$ and both $f(x)$ and $g(x)$ have degree at least 1. Since the sum of the degrees of $f(x)$ and $g(x)$ is d , it follows that each of them has degree less than d .

Since $p(\alpha) = 0$, we get $0 = f(\alpha)g(\alpha)$, and so either $f(\alpha) = 0$ or $g(\alpha) = 0$. This is impossible since d is the smallest positive degree of a rational polynomial having α as a root.

This shows $p(x)$ is irreducible, so we have shown that there exists a monic irreducible rational polynomial of degree d with $p(\alpha) = 0$.

Now we want to show that that is only one such polynomial. To do this, suppose $p(x)$ and $q(x)$ are monic irreducible polynomials of degree d such that $p(\alpha) = q(\alpha) = 0$. Let $h(x) = p(x) - q(x)$. Since $p(x)$ and $q(x)$ have the same leading term, x^d , the degree of $h(x)$ must be less than d . As well, $h(\alpha) = p(\alpha) - q(\alpha) = 0$, so α is a root of a polynomial with degree less than d . By the definition of d , $h(x)$ cannot have positive degree, so it must be constant. The only constant polynomial with roots is the constant zero polynomial, so $h(x) = p(x) - q(x) = 0$ for all x , from which it follows that $p(x) = q(x)$.

We have assumed that two monic irreducible polynomials of degree d have α as a root and deduced that they are the same polynomial. We conclude that there is a *unique* monic irreducible polynomial $m(x)$ of degree d such that $m(\alpha) = 0$.

(g) By looking for rational roots and removing corresponding factors, we arrive at

$$f(x) = (x + 1)(x - 2)^2(x^4 + 2x^2 + 1)$$

In part (b), it was observed that $x^4 + 2x^2 + 1 = (x^2 + 1)^2$, so $f(x)$ factors completely as

$$f(x) = (x + 1)(x - 2)^2(x - i)^2(x + i)^2$$

and so the values of $r_1, r_2, r_3, r_4, r_5, r_6, r_7$ are $-1, 2, 2, i, i, -i, -i$.

(h) (i) Expanding $(x + y)^n$ for each n from 2 through 7, we have

$$\begin{aligned} (x + y)^7 &= x^7 + 7x^6y + 21x^5y^2 + 35x^4y^3 + 35x^3y^4 + 21x^2y^5 + 7xy^6 + y^7 \\ (x + y)^6 &= x^6 + 6x^5y + 15x^4y^2 + 20x^3y^3 + 15x^2y^4 + 6xy^5 + y^6 \\ (x + y)^5 &= x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5 \\ (x + y)^4 &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4 \\ (x + y)^3 &= x^3 + 3x^2y + 3xy^2 + y^3 \\ (x + y)^2 &= x^2 + 2xy + y^2 \end{aligned}$$

Substituting $x + y$ for x in $f(x)$, we get

$$f(x + y) = (x + y)^7 - 3(x + y)^6 + 2(x + y)^5 - 2(x + y)^4 + (x + y)^3 + 5(x + y)^2 + 4$$

Without actually substituting the expressions for $(x + y)^2$ through $(x + y)^7$ from above, we can imagine what will happen if we do. The polynomial $f_0(x)$ is equal to the sum of the terms we would get that have no factor of y . The only term in $(x + y)^n$ that does not have a factor of y is x^n . Therefore, we can conclude that

$$f_0(x) = x^7 - 3x^6 + 2x^5 - 2x^4 + x^3 + 5x^2 + 4$$

which is $f(x)$, the roots of which were found in the previous part.

The polynomial $f_1(x)$ has the property that $yf_1(x)$ is the sum of all terms that have a factor of y and the exponent of y is exactly 1. In $(x + y)^n$, this term is always of the form $nx^{n-1}y$.

Therefore, we can collect all terms that have exactly one factor of y in $f(x + y)$ to get

$$\begin{aligned} yf_1(x) &= 7x^6y - 3(6x^5y) + 2(5x^4y) - 2(4x^3y) + (3x^2y) + 5(2xy) \\ &= y(7x^6 - 18x^5 + 10x^4 - 8x^3 + 3x^2 + 10x) \end{aligned}$$

So we conclude that $f_1(x) = 7x^6 - 18x^5 + 10x^4 - 8x^3 + 3x^2 + 10x$.

After checking for rational roots, one finds that $f_1(x) = x(x - 2)(7x^4 - 4x^3 + 2x^2 - 4x - 5)$ and that $h(x) = 7x^4 - 4x^3 + 2x^2 - 4x - 5$ has no rational roots. However, a bit of experimentation or observation leads to

$$\begin{aligned} h(i) &= 7(i)^4 - 4i^3 + 2i^2 - 4i - 5 \\ &= 7(-1)^2 - 4(-1)i + 2(-1) - 4i - 5 \\ &= 7 + 4i - 2 - 4i - 5 \\ &= 0 \end{aligned}$$

and one can check that $h(-i) = 0$ as well. Thus, we should be able to factor $(x - i)$ and $(x + i)$ out of $h(x)$, but $(x - i)(x + i) = x^2 + 1$, so we can factor $x^2 + 1$ out of $h(x)$ and avoid arithmetic with complex numbers. After doing this, we find $h(x) = (x^2 + 1)(7x^2 - 4x - 5)$. Using the quadratic formula on $7x^2 - 4x - 5$ gives $x = \frac{2 \pm \sqrt{39}}{7}$.

We have now found all roots of $f_1(x)$ and they are $0, 2, i, -i, \frac{2 + \sqrt{39}}{7},$ and $\frac{2 - \sqrt{39}}{7}$. Observe that $2, i,$ and $-i$ were all repeated roots of $f(x)$ (from part (g)) and they all appear as roots of $f_1(x)$.

(ii) In general, $f(x + y)$ can be expressed as

$$f(x + y) = f(x) + f_1(x)y + X$$

where X is some expression that is the sum of products of scalars, powers of x , and powers of y . However, since $f(x)$ and $f_1(x)y$ are the collections of terms that have no factor of y and exactly one factor of y , we can conclude that X has a factor of y^2 . Therefore, we can refine this observation to get that $f(x + y) = f(x) + f_1(x)y + y^2\bar{f}(x, y)$ where $\bar{f}(x, y)$ is the sum of terms that are products of scalars, powers of x , and powers of y . Note that if $f(x)$ is constant, then $f_1(x)$ and $\bar{f}(x, y)$ will be 0, and if $f(x)$ has degree 1, then $\bar{f}(x, y)$ will be 0.

Assume that r is a root of $f(x)$.

We will first show that if r is a repeated root of $f(x)$, then r must be a root of $f_1(x)$. To do that, we assume that r is a repeated root of $f(x)$. By definition, this means there is a polynomial $p(x)$ such that $f(x) = (x - r)^2p(x)$.

Expanding $p(x + y)$, we get

$$p(x + y) = p(x) + p_1(x)y + y^2\bar{p}(x, y)$$

Then we have

$$\begin{aligned}
f(x+y) &= (x+y-r)^2 p(x+y) \\
&= [(x-r)+y]^2 [p(x) + p_1(x)y + y^2 \bar{p}(x,y)] \\
&= [(x-r)^2 + 2y(x-r) + y^2] [p(x) + p_1(x)y + y^2 \bar{p}(x,y)] \\
&= p(x)(x-r)^2 + y[2(x-r)p(x) + p_1(x)(x-r)^2] + y^2 h(x,y)
\end{aligned}$$

where $h(x,y)$ is some expression in x and y . Therefore, when we expand $f(x+y)$, the sum of the terms with y^1 as their power of y is $y[2(x-r)p(x) + p_1(x)(x-r)^2]$. By definition, this sum also equals $yf_1(x)$, so $f_1(x) = 2(x-r)p(x) + p_1(x)(x-r)^2$. Substituting $x=r$ into this equation gives $f_1(r) = 2(r-r)p(r) + p_1(r)(r-r)^2 = 0$. Therefore, r is a root of $f_1(x)$.

We now assume that r is a root of both $f(x)$ and $f_1(x)$ and will deduce that $(x-r)^2$ is a factor of $f(x)$.

Since we are assuming that $f(x)$ has a root of r , we can write $f(x) = p(x)(x-r)$ for some polynomial $p(x)$. Then

$$\begin{aligned}
f(x+y) &= (x+y-r)p(x+y) \\
&= [(x-r)+y][p(x) + p_1(x)y + y^2 \bar{p}(x,y)] \\
&= p(x)(x-r) + y[p(x) + (x-r)p_1(x)] + y^2 k(x,y)
\end{aligned}$$

where $k(x,y)$ is some expression in x and y . Similar to the argument for the other direction, this implies $yf_1(x) = y[p(x) + (x-r)p_1(x)]$, hence $f_1(x) = p(x) + (x-r)p_1(x)$. We are assuming that $f_1(r) = 0$, so we can substitute $x=r$ on both sides of this equation to get $f_1(r) = p(r) + (r-r)p_1(r)$ or $0 = p(r) + 0$. Therefore, $p(r) = 0$, so there is some polynomial $q(x)$ such that $p(x) = q(x)(x-r)$. Substituting into $f(x) = p(x)(x-r)$, we get $f(x) = q(x)(x-r)^2$, and so r is a repeated root of $f(x)$ by definition.

A proof using calculus. If you take a minute to verify that the polynomial $f_1(x)$ is the derivative of polynomial $f(x)$, denoted by $f'(x)$, then we can reframe this result as follows: Suppose that $f(x)$ is a polynomial and that r is a root of $f(x)$. Prove that r is a repeated root of $f(x)$ if and only if r is a root of $f'(x)$.

Assume that r is a root of $f(x)$.

Suppose that r is a repeated root of $f(x)$. Then $f(x) = p(x)(x-r)^2$ for some polynomial $p(x)$. By the product rule, $f'(x) = p'(x)(x-r)^2 + 2p(x)(x-r)$, so $f'(r) = p'(r)(r-r)^2 + 2p(r)(r-r) = 0$. Therefore, r is a root of $f'(x)$.

Now suppose that r is a root of $f'(x)$. Since r is a root of $f(x)$, there is a polynomial $p(x)$ such that $f(x) = p(x)(x-r)$. By the product rule, $f'(x) = p'(x)(x-r) + p(x)$. Substituting $x=r$ gives $f'(r) = p'(r)(r-r) + p(r)$ and since $f'(r) = 0$ by assumption, this implies $0 = 0 + p(r)$, so r is a root of $p(x)$. Therefore, there is a polynomial $q(x)$ such that $p(x) = q(x)(x-r)$. Substituting gives $f(x) = p(x)(x-r) = q(x)(x-r)^2$, so r is a repeated root of $f(x)$.

- (j) Suppose $p(x)$ and $q(x)$ are irreducible rational polynomials with a root, α , in common. Assume that $p(x)$ had degree n . We know that α is an algebraic number, so let $m(x)$ be its minimal polynomial.

By the division algorithm for polynomials, there are polynomials $f(x)$ and $r(x)$ such that $p(x) = f(x)m(x) + r(x)$ and the degree of $r(x)$ is *less* than the degree of $m(x)$. By the definition of the minimal polynomial, $m(\alpha) = 0$. By assumption, $p(\alpha) = 0$, so $p(\alpha) = f(\alpha)m(\alpha) + r(\alpha)$ implies that $0 = 0 + r(\alpha)$ or $r(\alpha) = 0$. Since $m(x)$ is the minimal polynomial of α and the degree of $r(x)$ is less than the degree of $m(x)$, we must have that r does not have positive degree. This means $r(x)$ is constant, and since $r(\alpha) = 0$, it must be the zero polynomial¹.

Therefore, $p(x) = f(x)m(x)$, which is a factorization of $p(x)$ into a product of two polynomials. Since $p(x)$ is irreducible and $m(x)$ has positive degree, we must have that $f(x)$ is constant. Therefore, there is a constant a such that $p(x) = am(x)$. Since $p(x)$ is irreducible, it has degree at least 1, so $a \neq 0$.

By an essentially identical argument, there is a constant $b \neq 0$ such that $q(x) = bm(x)$. Taking $c = \frac{a}{b}$, we have

$$cq(x) = cbm(x) = \frac{a}{b}bm(x) = am(x) = p(x)$$

Computing the degrees of the algebraic numbers from part (e). We can use parts (f) and (j) to prove the following: If an algebraic number is a root of an irreducible polynomial of degree d , then the degree of the algebraic number is d .

To see why this is true, suppose α is algebraic of degree d and is a root of an irreducible polynomial $p(x)$ of degree n . By part (f), the minimal polynomial, $m(x)$, of α has degree d . This means $p(x)$ and $m(x)$ are irreducible polynomials with a root, α , in common. By part (j), one is a scalar multiple of the other (and that scalar is nonzero), so they have the same degree. In other words, $n = d$.

It follows that if we are given an algebraic number α and produce an irreducible polynomial of degree d to which α is a root, we will have shown that the degree of α is d . In part (e), one can show that each of the four polynomials produced (for each algebraic number) is irreducible, so the degrees of the algebraic numbers are as stated at the end of the solution to (e).

- (k) Suppose $p(x)$ is an irreducible rational polynomial with a repeated root, α . By part (h)(ii), α is also a root of $p_1(x)$. Note that if $a_n x^n$ is the leading term of $p(x)$, then $na_n x^{n-1}$ is the leading term of $p_1(x)$, and so $p_1(x)$ has degree strictly lower than that of $p(x)$. As well, $p(x)$ has a repeated root, so $n \geq 2$, which means $n - 1 \geq 1$. Therefore, $p_1(x)$ has positive degree less than n .

Every polynomial can be factored into a product of irreducible factors. To see why, we can emulate the reasoning used to see that every positive integer is the product of prime numbers. For example, if $p_1(x)$ is irreducible, then we stop. Otherwise, it can be factored as $p_1(x) = f(x)g(x)$ where each of $f(x)$ and $g(x)$ has degree at least 1 and less than that of $p_1(x)$. Now either $f(x)$ and $g(x)$ are irreducible, or they can be factored into polynomials of lower degree. Critically, the degrees always go down but stay larger than 1 when we

¹For technical reasons, mathematicians usually distinguish the zero polynomial among the constant polynomials and do not consider it to have degree 0. Typically it is either not assigned a degree or assigned a “degree” of $-\infty$. For the purposes of this document, there is no harm in just taking the zero polynomial to have degree 0.

factor. Consequentially, this cannot go on forever, and we will eventually be left with $p_1(x)$ expressed as a product of irreducible polynomials.

Since $p_1(\alpha) = 0$, one of these irreducible factors, say $h(x)$, has α as a root. Since $h(x)$ is a factor of $p_1(x)$, its degree is no larger than the degree of $p_1(x)$, which means $h(x)$ has degree less than n .

We now have that α is a root of both $h(x)$ and $p(x)$. Both polynomials are irreducible, so part (j) implies that they must have the same degree. We have just argued that the degree of $h(x)$ is less than the degree of $p(x)$, so this is a problem.

This means our assumption that $p(x)$ has a repeated root must have been wrong, so we conclude that an irreducible rational polynomial cannot have a repeated root.



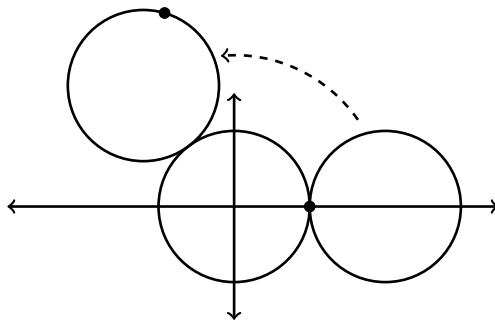
Problem of the Month

Problem 7: April 2024

Curves in the plane are often given as the set of points (x, y) that satisfy some equation in x and y . For example, the set of points (x, y) that satisfy $y = x^2$ is a parabola, the set of points (x, y) that satisfy $y = 3x + 4$ is a line, and the set of points (x, y) that satisfy $x^2 + y^2 = 1$ is the circle of radius 1 centred at the origin.

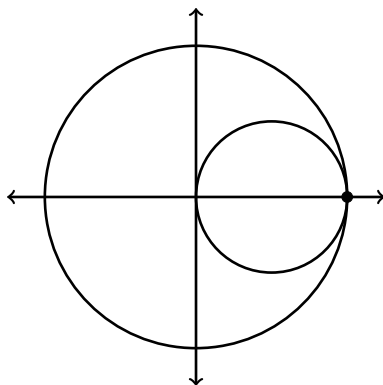
Another way to express a curve in the plane is using *parametric* equations. With this type of description, we introduce a third variable, t , called the *parameter*, and each of x and y is given as a function of t . This is useful for describing the position of a point that is moving around the plane. For example, imagine that an ant is crawling around the plane. If its x -coordinate at time t is $x = x(t)$ and its y -coordinate at time t is $y(t)$, then its position at time t is $(x(t), y(t))$.

- (a) A particle's position at time t is $(x, y) = (1 + t, -2 + 2t)$. That is, its x -coordinate at time t is $1 + t$ and its y -coordinate at time t is $-2 + 2t$.
- Plot the position of the particle at $t = 0, 1, 2, 3$, and 4.
 - Show that every position the particle occupies is on the line with equation $y = 2x - 4$.
 - Sketch the path of the particle from $t = 0$ through $t = 4$.
- (b) A particle's position at time t is $(\cos(t), \sin(t))$. Sketch the path of the particle from $t = 0$ through $t = 2\pi$.
- (c) A particle's position at time t is $(\cos(t), \sin(2t))$.
- Plot the position of the particle at $t = \frac{k\pi}{12}$ for the integers $k = 0$ through $k = 24$. Sketch the path of the particle from $t = 0$ through $t = 2\pi$.
 - Show that every position the particle occupies is on the curve with equation $y^2 = 4x^2 - 4x^4$.
- (d) Circle 1 is centred at the origin, Circle 2 is centred at $(2, 0)$, and both circles have radius 1. The circles are tangent at $(1, 0)$. Circle 2 is "rolled" in the counterclockwise direction along the outside of the circumference of Circle 1 without slipping. The point on Circle 2 that was originally at $(1, 0)$ (the point of tangency) follows a curve in the plane. Find functions $x(t)$ and $y(t)$ so that the points on this curve are $(x(t), y(t))$ for $0 \leq t \leq 2\pi$.



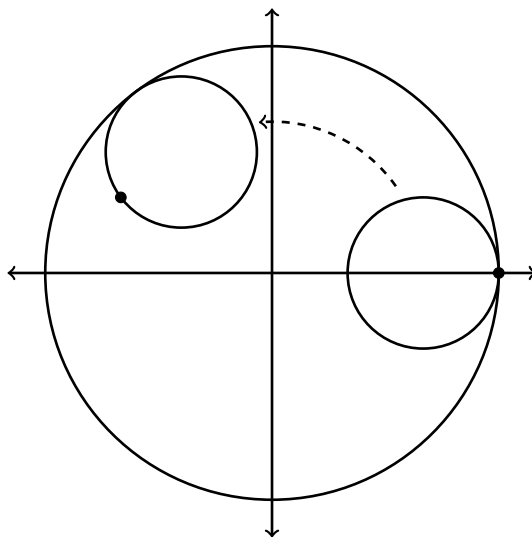


- (e) The setup in this problem is similar to (d). Circle 1 is centred at the origin and has radius 2 and Circle 2 is centred at $(1, 0)$ and has radius 1 so that the two circles are tangent at $(2, 0)$. Circle 2 is rolled around the inside of the circumference of Circle 1 in the counterclockwise direction. Describe the curve in the plane followed by the point on Circle 2 that is initially at $(2, 0)$.



- (f) Circle 1 is centred at the origin and has radius 1. Circle 2 has radius $r < 1$, is inside Circle 1, and the two circles are initially tangent at $(1, 0)$. When Circle 2 is rolled around the inside of Circle 1 in the counterclockwise direction, the point on Circle 2 that was initially at $(1, 0)$ will follow some curve in the plane.

- (i) Show that when $r = \frac{1}{4}$, the points on the curve satisfy the equation $(\sqrt[3]{x})^2 + (\sqrt[3]{y})^2 = 1$.
- (ii) Show that the curves for $r = \frac{1}{3}$ and $r = \frac{2}{3}$ are exactly the same and that the point initially at $(1, 0)$ travels this curve in opposite directions for the two radii.





Problem of the Month

Problem 7: April 2024

Hint

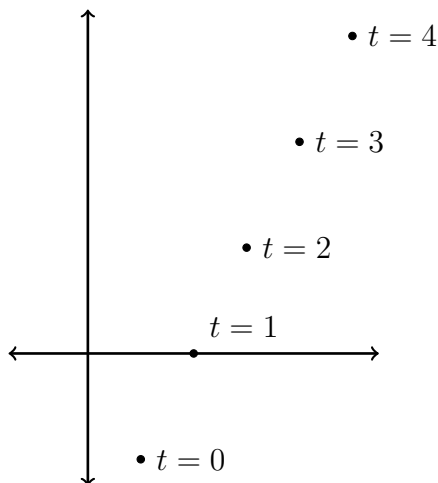
- (a) In part (ii), solve for t in terms of x and then substitute into the equation involving y .
 - (b) At time t , how far is the particle from the origin?
 - (c) Starting with $y^2 = (\sin 2t)^2$, use trigonometric identities to eliminate all appearances of the variable t . Remember that $x = \cos t$.
 - (d) As Circle 2 rolls around Circle 1, let t be the angle made by the positive x -axis and the line segment connecting the origin and the centre of Circle 2. It will help to draw a reasonably accurate diagram with Circle 2 rolled part of the way around Circle 1 (perhaps an angle of $\frac{\pi}{3}$ or so). Once you have done this, mark the point on the circumference of Circle 2 that was originally at $(1, 0)$ by P (or some other label). The objective is to find the coordinates of P in terms of t . Since the circles roll without slipping, the arclength from the point of tangency along Circle 1 to $(1, 0)$ should equal the arclength from the point of tangency along Circle 2 to P .
 - (e) As Circle 2 rolls along the inside of Circle 1, it (usually) intersects the x -axis at two points. Convince yourself that one of these two points must be the origin, then think about the other point.
 - (f) Using a strategy similar to part (d), find a general formula for the coordinates of P in terms of the angle t . Do this either in general for r or do it separately for the three relevant values of r in this question.
 - (i) Find identities for $\cos 3t$ in terms of $\cos t$ and $\sin 3t$ in terms of $\sin t$.
 - (ii) Find the parametric equations for the position of P when $r = \frac{1}{3}$ if Circle 2 is rolled clockwise around Circle 1 instead of counterclockwise.
-
-



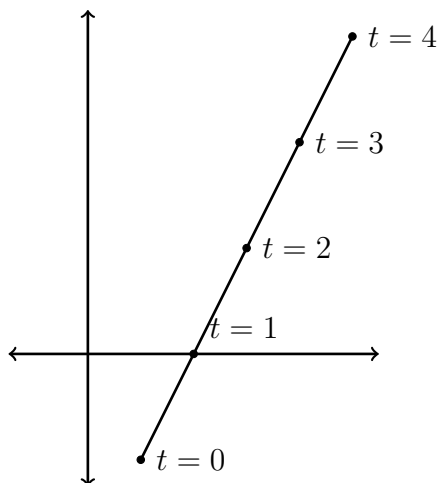
Problem of the Month

Solution to Problem 7: April 2024

- (a) (i) When $t = 0$, the position is $(1 + 0, -2 + 2(0)) = (1, -2)$. When $t = 1$, the position is $(1 + 1, -2 + 2(1)) = (2, 0)$. When $t = 2$, the position is $(3, 2)$, when $t = 3$, the position is $(4, 4)$, and when $t = 4$, the position is $(5, 6)$. The plot of these positions is below

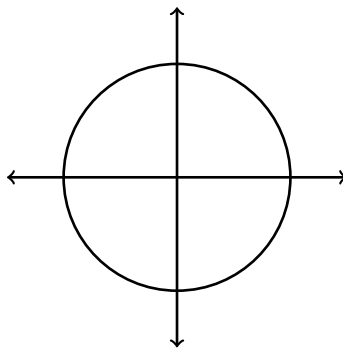


- (ii) For each t , we have $x = 1 + t$ and $y = -2 + 2t$. Solving for t in the equation $x = 1 + t$ gives $t = x - 1$, which can be substituted into the second equation to get $y = -2 + 2(x - 1) = -2 + 2x - 2 = 2x - 4$. Therefore, every point of the form $(x, y) = (t + 1, -2 + 2t)$ satisfies $y = 2x - 4$.
- (iii) From part (ii), the points that the particle occupies are all on the line with equations $y = 2x - 4$. We care about the points on this line from $t = 0$ to $t = 4$ inclusive, so the plot is just the line segment connecting $(1, -2)$ (the point for $t = 0$) to $(5, 6)$ (the point for $t = 4$). The plot is below.



- (b) We have $x = \cos t$ and $y = \sin t$, so for any position (x, y) that the particle occupies, $x^2 + y^2 = \cos^2 t + \sin^2 t = 1$. Therefore, the particle is always somewhere on the unit circle.

Indeed, every point on the unit circle has coordinates $(\cos t, \sin t)$ for exactly one real number t with $0 \leq t < 2\pi$. The graph of the path of the particle is



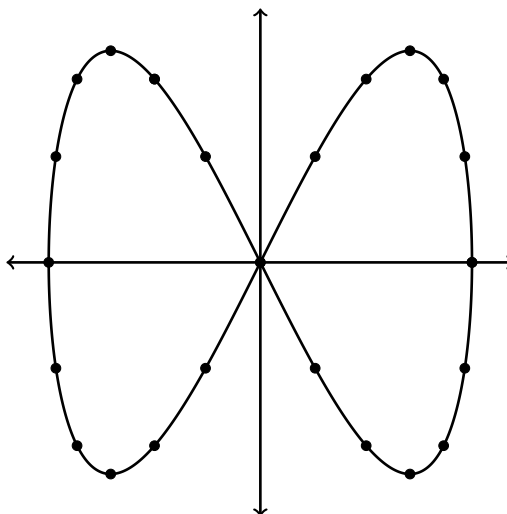
- (c) (i) In each of the three tables below, the left column contains values of t , the middle column contains the corresponding values of $\cos t$, and the right column contains the corresponding values of $\sin 2t$. Although it is not particularly difficult to write down exact values for each of the trigonometric ratios in the table, we want to plot the points so the values are all given rounded to three digits past the decimal point.

t	$\cos t$	$\sin 2t$
0	1	0
$\frac{\pi}{12}$	0.966	0.5
$\frac{2\pi}{12}$	0.866	0.866
$\frac{3\pi}{12}$	0.707	1
$\frac{4\pi}{12}$	0.5	0.866
$\frac{5\pi}{12}$	0.259	0.5
$\frac{6\pi}{12}$	0	0
$\frac{7\pi}{12}$	-0.259	-0.5

t	$\cos t$	$\sin 2t$
$\frac{8\pi}{12}$	-0.5	0.866
$\frac{9\pi}{12}$	-0.7071	-1
$\frac{10\pi}{12}$	-0.866	-0.866
$\frac{11\pi}{12}$	-0.966	-0.5
π	-1	0
$\frac{13\pi}{12}$	-0.966	0.5
$\frac{14\pi}{12}$	-0.866	0.866
$\frac{15\pi}{12}$	-0.707	1

t	$\cos t$	$\sin 2t$
$\frac{16\pi}{12}$	-0.5	0.866
$\frac{17\pi}{12}$	-0.259	0.5
$\frac{18\pi}{12}$	0	0
$\frac{19\pi}{12}$	0.259	-0.5
$\frac{20\pi}{12}$	0.5	-0.866
$\frac{21\pi}{12}$	0.707	-1
$\frac{22\pi}{12}$	0.866	-0.866
$\frac{23\pi}{12}$	0.966	-0.5
$\frac{24\pi}{12}$	1	0

Below is a plot of the points above including a sketch of the curve.

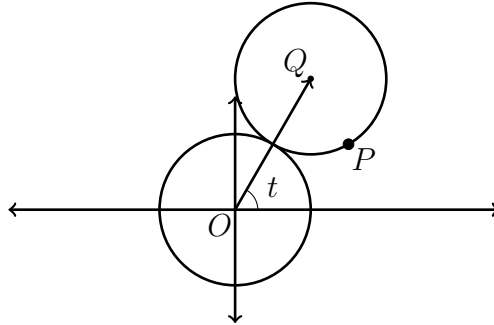


- (ii) Every point (x, y) on the parametric curve satisfies $x = \cos t$ and $y = \sin 2t$ for some real number t . Using the double-angle formula for $\sin 2t$, we get $y = 2 \sin t \cos t$,

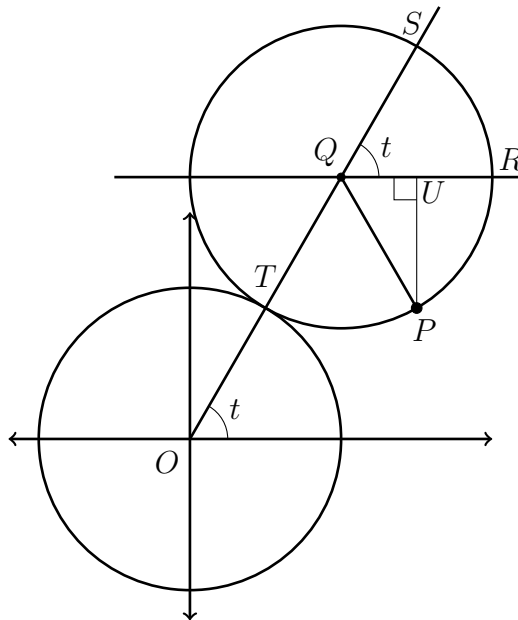
so $y^2 = 4 \sin^2 t \cos^2 t$. By the Pythagorean identity, $1 - x^2 = 1 - \cos^2 t = \sin^2 t$. Substituting gives $y^2 = 4(1 - x^2)x^2 = 4x^2 = 4x^4$.

- (d) We will label the point on Circle 2 that is originally at $(1, 0)$ by P and we will label the centre of Circle 2 by Q . Since the circumferences of the circles are the same, Circle 2 will return to its original position after rolling exactly once around Circle 1. This means Q will travel exactly once around the circle of radius $1 + 1 = 2$ centred at the origin.

We will let t represent the angle made by the positive x -axis and the ray from the origin to Q . For example, the diagram below depicts the position of the outer circle when $t = \frac{\pi}{3}$. The origin is labelled by O .



We wish to find the coordinates of P in terms of the angle t . In the diagram below, we have added to the previous diagram a line through the origin and Q , a line segment connecting Q to P , as well as a horizontal line through Q . The horizontal line intersects Circle 2 at R to the right of Q . The line through O and Q intersects the point of tangency of the two circles at T and it also intersects Circle 2 at S (the points S and T are different). A perpendicular has also been drawn from P to QR intersecting QR at U .



The line connecting the centres of tangent circles always passes through the point of tangency, which justifies the implicit claim in the previous paragraph that OQ passes through the point of tangency. It also implies that the length of OQ is $1 + 1 = 2$. Therefore, the coordinates of Q are $(2 \cos t, 2 \sin t)$. Since QR is horizontal by construction, QR is parallel

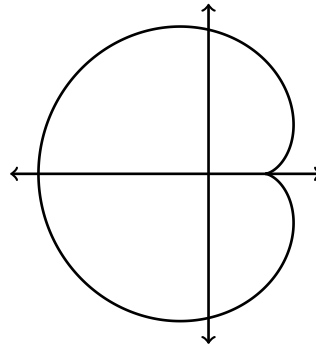
to the x axis, which implies $\angle SQR = t$.

Because Circle 2 is rolling along Circle 1 without slipping, the arc from T to the x axis along Circle 1 has the same length as the arc TP . This means $\angle TQP = t$ as well since the two circles have the same radius. Since TQS is a line, we get that $\angle RQP = \pi - 2t$.

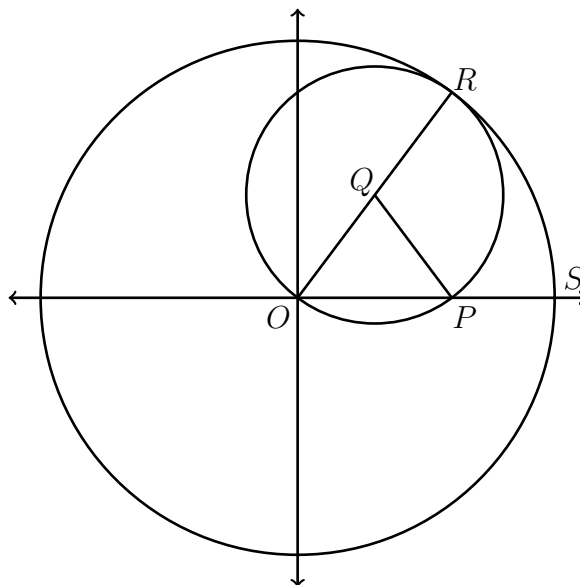
Suppose Q has coordinates (q_1, q_2) and P has coordinates (p_1, p_2) . The radius of Circle 2 is 1, so $QP = 1$. This means $p_1 - q_1 = \cos(\pi - 2t)$ and $q_2 - p_2 = \sin(\pi - 2t)$. Thus,

$$\begin{aligned} (p_1, p_2) &= (q_1 + (p_1 - q_1), q_2 - (q_2 - p_2)) \\ &= (2 \cos t + \cos(\pi - 2t), 2 \sin t - \sin(\pi - 2t)) \\ &= (2 \cos t - \cos 2t, 2 \sin t - \sin 2t) \end{aligned}$$

where the last equality is by trigonometric identities. Therefore, $x(t) = 2 \cos t - \cos 2t$ and $y(t) = 2 \sin t - \sin 2t$. A plot of this curve is given below.



- (e) As Circle 2 rolls around the inside of Circle 1, if a line is drawn through the point of tangency perpendicular to the mutual tangent, then the line will pass through the centre of both circles. Such a line contains a diameter of both circles, and since the radius of Circle 1 equals the diameter of Circle 2, the centre of Circle 1 is always on Circle 2. In the diagram below, Circle 2 has been rotated by some positive angle between 0 and $\frac{\pi}{2}$. The origin is labelled by O , the current point of tangency is labelled by R , the centre of Circle 2 is labelled by Q , the other point at which Circle 2 intersects the x -axis is labelled by P , and $(2, 0)$ is labelled by S .

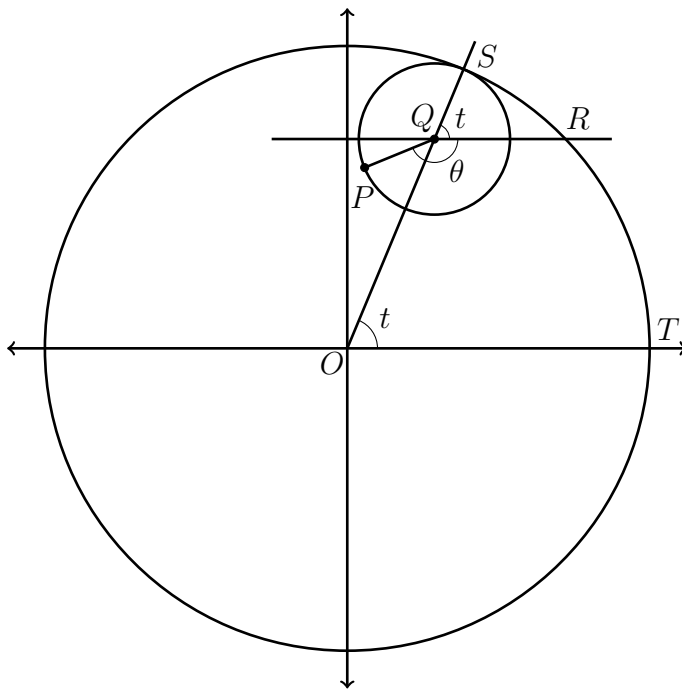


We wish to determine the current location of the point on Circle 2 that started at S . Circle 2 does not slip as it rolls, so this point is the same distance from R in the clockwise direction along both circles.

Since O is on the circumference of Circle 2 and Q is the centre of Circle 2, a well-known circle property implies that $\angle PQR = 2\angle POR$. The radius of Circle 1 is 2, so provided we measure angles in radians, the length of the arc PS is $2\angle POR$. The radius of Circle 2 is 1, so the length of arc PR is $\angle PQR$. These quantities are equal, so the arcs PR and PS have equal length. Therefore, the original point of tangency is at P .

By drawing similar diagrams for obtuse and reflex angles, it can be similarly shown that the original point of tangency is always on the diameter of Circle 1. Now note half the circumference of Circle 1 is equal in length to the circumference of Circle 2, so when Circle 2 has rolled exactly half way around Circle 1, the original point of tangency is exactly where Circle 1 intersects the negative x -axis. It follows by symmetry that the original point of tangency will go from $(2, 0)$ to $(-2, 0)$ and back to $(2, 0)$ as Circle 2 rolls around Circle 1.

- (f) We will first work out, in general, a pair of parametric equations for P , the original point of tangency. As in part (d), O is the origin, Q is the centre of Circle 2, and the measure of the angle made by the positive x -axis and the ray from O to Q is t . In the diagram below, a horizontal line is drawn through Q intersecting Circle 2 at R to the right of Q and the line defined by OQ , for the same reasoning that was used in part (d), passes through the mutual point of tangency labelled by S . We label $(1, 0)$ by T .



As mentioned, O , Q , and S are on a line, and so $OQ = OS - QS = 1 - r$. Therefore, the coordinates of Q are $((1 - r) \cos t, (1 - r) \sin t)$. If we let θ be $\angle PQR$, then by reasoning similar to that which was used in part (d), the coordinates of P are

$$((1 - r) \cos t + r \cos \theta, (1 - r) \sin t - r \sin \theta)$$

For now, assume that t is small enough that Circle 2 has not yet rolled far enough for P to have returned to the circumference of Circle 1. The length of arc SP on Circle 2 is

equal to the length of arc ST on Circle 1, which is equal to t since the radius of Circle 1 is 1 (provided we use radians as our unit of angle measure). Because QR is parallel to OT , $\angle SQR = \angle SOT = t$, so $\angle SQP = t + \theta$ (note that this angle is measured clockwise from S , so in the diagram, the angle measures more than π radians). The radius of Circle 2 is r , so the length of arc SP is $r(t + \theta)$. The arcs ST and SP are equal, so $t = rt + r\theta$, which can be solved for θ to get $\theta = \frac{t - tr}{r}$, and we note that $0 < r < 1$, so $t - tr$ is positive.

Now suppose that t is such that P has already returned to Circle 1 k times. Between consecutive times that P returns to Circle 1, the arc ST increases in length by $2\pi r$, the circumference of Circle 2. Therefore, instead of the equation $t = rt + r\theta$, we get $t = rt + r\theta + 2\pi rk$ since the arc length from S to T is still t , but to get equality, we need to account for the k complete revolutions of Circle 2.

Solving this equation for θ gives $\theta = \frac{t - rt - 2\pi rk}{r} = \frac{t - rt}{r} - 2\pi k$. However, since we will be taking \sin and \cos of θ , we can ignore the quantity $2\pi k$ since k is an integer. Therefore, the coordinates of P when Q makes an angle of t with the positive x -axis are

$$\left((1 - r) \cos t + r \cos \left(\frac{t - tr}{r} \right), (1 - r) \sin t - r \sin \left(\frac{t - tr}{r} \right) \right)$$

Before answering the given questions, we note that if $r = \frac{1}{2}$, (that is, the radius of Circle 2 is twice that of Circle 1), then the coordinates simplify to $(\cos t, 0)$, meaning the point P always has a y -coordinate of 0. This gives another proof of the result in (e).

(i) When $r = \frac{1}{4}$, the coordinates of P are

$$\left(\frac{3}{4} \cos t + \frac{1}{4} \cos(3t), \frac{3}{4} \sin t - \frac{1}{4} \sin(3t) \right)$$

Using standard trigonometric identities, one can show that $\cos(3t) = 4 \cos^3 t - 3 \cos t$ and $\sin(3t) = 3 \sin t - 4 \sin^3 t$. Substituting into the coordinates above, we get

$$\begin{aligned} x(t) &= \frac{1}{4}(3 \cos t + \cos(3t)) & y(t) &= \frac{1}{4}(3 \sin t - \sin(3t)) \\ &= \frac{1}{4}(3 \cos t + 4 \cos^3 t - 3 \cos t) & &= \frac{1}{4}(3 \sin t - 3 \sin t + 4 \sin^3 t) \\ &= \cos^3 t & &= \sin^3 t \end{aligned}$$

leading to the rather tidy expression for the coordinates of $(\cos^3 t, \sin^3 t)$. Therefore, $\sqrt[3]{x} = \cos t$ and $\sqrt[3]{y} = \sin t$, so $(\sqrt[3]{x})^2 + (\sqrt[3]{y})^2 = \cos^2 t + \sin^2 t = 1$.

(ii) For $r = \frac{1}{3}$, we get

$$\begin{aligned} x(t) &= \frac{2}{3} \cos t + \frac{1}{3} \cos(2t) & y(t) &= \frac{2}{3} \sin t - \frac{1}{3} \sin(2t) \\ &= \frac{1}{3}(2 \cos t + \cos 2t) & &= \frac{1}{3}(2 \sin t - \sin 2t) \end{aligned}$$

For $r = \frac{2}{3}$, we get

$$\begin{aligned} x(t) &= \frac{1}{3} \cos t + \frac{2}{3} \cos\left(\frac{t}{2}\right) & y(t) &= \frac{1}{3} \sin t - \frac{2}{3} \sin\left(\frac{t}{2}\right) \\ &= \frac{1}{3} \left(\cos t + 2 \cos \frac{t}{2} \right) & &= \frac{1}{3} \left(\sin t - 2 \sin \frac{t}{2} \right) \end{aligned}$$

Focusing for a moment on the equations for $r = \frac{2}{3}$, if we substitute $t = 2\pi$, we get $x(2\pi) = -\frac{1}{3}$ and $y(2\pi) = 0$, which means the point initially at $(1, 0)$ on Circle 2 has not returned to its original position yet. This is because the circumference of Circle 1 is not an integer multiple of the circumference of Circle 2. Indeed, the circumference of Circle 1 is 2π , but the circumference of Circle 2 is $\frac{4\pi}{3}$.

When Circle 2 first reaches the point where it is tangent to Circle 1 at $(1, 0)$, some points on Circle 2 have been tangent to Circle 1 more than once. To be precise, $2\pi - \frac{4\pi}{3} = \frac{2\pi}{3}$, so every point on Circle 2 has been tangent to Circle 1 at least once, but the points on an arc of length $\frac{2\pi}{3}$ on Circle 2 have been tangent to Circle 1 twice. Specifically, since $2 \times \frac{2\pi}{3} = \frac{4\pi}{3}$, exactly half of the points on Circle 2 have been tangent to Circle 1 twice and the rest have been tangent once.

If Circle 2 rolls around Circle 1 again, then half the points on Circle 2 will be tangent to Circle 1 exactly twice, and the other half will be tangent exactly once. This gives a total of three times for each point. Indeed, $3 \frac{4\pi}{3} = 4\pi = 2(2\pi)$, so three times the circumference of Circle 2 is an integer multiple of the circumference of Circle 1. Therefore, P returns to $(1, 0)$ for the first time when the circumference of Circle 2 wraps around the inside of the circumference of Circle 1 exactly 2 times.

Algebraically, if we substitute $t = 4\pi$, we get $x(4\pi) = 1$ and $y(4\pi) = 0$, so the curve for $r = \frac{2}{3}$ is travelled periodically every *two* times Circle 2 rolls along Circle 1.

To avoid confusion, we will refer to the circle with $r = \frac{1}{3}$ as Circle 3 and the circle with $r = \frac{2}{3}$ as Circle 2. The circumference of Circle 1 is an integer multiple of the circumference of Circle 3, so point P reaches $(1, 0)$ after Circle 3 rolls around Circle 1 exactly once.

Suppose the centre of Circle 3 revolves around the origin at a rate of 1 radian per second and the centre of Circle 2 revolves around the origin at a rate of 2 radians per second. By assuming this, we will have that both circles take exactly 2π seconds for P to return to its original position for the first time.

If Circle 3 is instead rolled clockwise around Circle 1, then the position of P will be $(x(2\pi - t), y(2\pi - t))$ where $x(t)$ and $y(t)$ are as given for $r = \frac{1}{3}$ above. This is because, for example, point P is in the same position after rolling $\frac{\pi}{3}$ radians counterclockwise

as it would be if it had rolled $2\pi - \frac{\pi}{3}$ radians clockwise. Thus, if Circle 3 rotated clockwise instead of counterclockwise, its position at time t is given by $(x(t), y(t))$ where

$$\begin{aligned} x(t) &= \frac{1}{3}(2 \cos(2\pi - t) + \cos(2(2\pi - t))) & y(t) &= \frac{1}{3}(2 \sin(2\pi - t) - \sin(2(2\pi - t))) \\ &= \frac{1}{3}(2 \cos t + \cos(2t)) & &= \frac{1}{3}(-2 \sin t + \sin 2t) \\ &= \frac{1}{3}(\cos 2t + 2 \cos t) & &= \frac{1}{3}(\sin 2t - 2 \sin t) \end{aligned}$$

If Circle 2 is rotated counterclockwise (as before) at 2 radians per second, then at time t the angle is $2t$, and so the coordinates of P are

$$\begin{aligned} x(t) &= \frac{1}{3} \left(\cos 2t + 2 \cos \frac{2t}{2} \right) & y(t) &= \frac{1}{3} \left(\sin 2t - 2 \sin \frac{2t}{2} \right) \\ &= \frac{1}{3}(\cos 2t + 2 \cos t) & &= \frac{1}{3}(\sin 2t - 2 \sin t) \end{aligned}$$

which are the exact same coordinates as Circle 3 at time t . Of course, if Circle 3 is rotated clockwise, then it travels the same path as if it were rotated counterclockwise but in the opposite direction. Thus, if both Circle 2 and Circle 3 are rotated counterclockwise, then P travels the same path in opposite directions.



Problem of the Month

Problem 8: May 2024

This month's problem is an extension of Question 4 from the 2024 Galois contest. It is not necessary to try the problem before attempting the questions below.

In an $m \times n$ rectangular grid, we say that two cells are *neighbours* if they share an edge. The *neighbourhood* of a cell is the cell itself along with its neighbours.

An $m \times n$ grid is called a *Griffin Grid* if each of its mn cells contains either a 1 or a -1 and the integer in every cell is equal to the product of the other integers in its neighbourhood.

For example, the 4×9 grid below is a Griffin Grid. The three shaded regions are the neighbourhoods of the cells in Row 1 and Column 1, Row 1 and Column 8, and Row 4 and Column 4.

-1	-1	1	1	1	1	1	-1	-1
1	1	-1	1	1	1	-1	1	1
-1	1	-1	-1	1	-1	-1	1	-1
-1	1	-1	1	1	1	-1	1	-1

The Galois problem restricted this definition to $m = 3$. Here we want to explore what happens more generally. If a question is marked with an asterisk (*), it means I was unable to solve it. Solutions will not be provided for these problems, but I would love to hear if you solve one!

- Show that an $m \times n$ grid with -1 or 1 in every cell is a Griffin Grid if and only if the cells in every neighbourhood have a product of 1 .
 - For every n , determine the number of $2 \times n$, $3 \times n$, and $4 \times n$ Griffin Grids. Determining the number of $3 \times n$ Griffin Grids in general is essentially what is required to answer part (c) of the Galois question.
 - Show that the number of $m \times n$ Griffin Grids is of the form 2^k for some integer k with $0 \leq k \leq m$.
 - * For general m , determine for which k there exists n with the property that the number of $m \times n$ Griffin Grids is exactly 2^k .
 - Show that for all m there exist infinitely many n for which there is exactly one $m \times n$ Griffin Grid.
 - Show that for all m there exist infinitely many n for which there are 2^m distinct $m \times n$ Griffin Grids.
 - * Find a simple general way to calculate the number of $m \times n$ Griffin Grids.
-
-



Problem of the Month

Problem 8: May 2024

Hint

- (b),(c) Once the m integers are placed in the leftmost column of a Griffin Grid, the rest of the integers in the grid are uniquely determined.

The definition of a Griffin Grid can be extended to infinite grids. It is useful to consider the infinite (in one direction) grid with m rows and infinitely many columns numbered 1, 2, 3, and so on indefinitely. In such an infinite Griffin Grid, the first k columns form an $m \times n$ Griffin Grid if and only if the $(n + 1)^{\text{st}}$ column consists entirely of 1s.

Instead of filling in the infinite grid with -1 s and 1 s, fill the first column with variables, then start to fill in the grid in general. Keep in mind that the only values that the variables will ever take is -1 and 1 , so you can assume that every variable squares to 1. For example, if $m = 2$ and the variables are a (in the top cell) and b (in the bottom cell), then both variables in the second column are ab , the third column is identical to the first column, and the fourth column has 1 in both cells. Using the observation at the end of the previous paragraph, the number of $m \times n$ Griffin Grids is always equal to the set of solutions to a very specific type of system of equations.

- (e),(f) Use the same approach as is outlined for part (c) above. You want to show that there are always lots of columns that give rise to systems of equations that have many solutions and lots that give rise to systems with very few solutions. For example, if the $(n + 1)^{\text{st}}$ column (in the general grid with variables) consists entirely of 1s, then there are 2^m different $m \times n$ Griffin Grids.
-
-



Problem of the Month

Solution to Problem 8: May 2024

- (a) Suppose an $m \times n$ grid with a 1 or -1 in every cell is a Griffin Grid. Consider an arbitrary cell C and suppose it contains the integer a (so $a = 1$ or $a = -1$). Let b be the product of the integers in the neighbours of C . Since the grid is a Griffin Grid, we have $a = b$. The product of the integers in the neighbourhood of C is $ab = a^2 = 1$, since $1^2 = (-1)^2 = 1$.

Now suppose an $m \times n$ grid has a 1 or a -1 in every cell and that the product of the integers in every neighbourhood is 1. Let C be an arbitrary cell, let a be the integer in C , and let b be the product of the integers in the neighbours of C . To verify that the grid is a Griffin Grid, we need to show that $a = b$.

We are assuming that $ab = 1$, and we know that $a = \pm 1$ and $b = \pm 1$ because it is the product of some integers, all of which are either 1 or -1 . If a and b were different, then their product would be -1 , not 1. Therefore, $a = b$.

- (b) Using similar reasoning to that which was used in part (a), it can be shown that an $m \times n$ grid with 1 or -1 in every cell is a Griffin Grid if and only if for every neighbourhood in the grid, the integer in each cell in the neighbourhood is equal to the product of the other cells in that neighbourhood.

In any $2 \times n$ grid with 1 or -1 in every cell, there are only four possibilities for the first column. They are shown below

1	1	-1	-1
1	-1	1	-1

As suggested in the hint, we will now try to understand Griffin Grids with 2 rows and infinitely many columns. We will refer to such a grid as a $2 \times \infty$ Griffin Grid.

The second column can be completely determined from the first column using the observation at the beginning of this part of the solution. Specifically, consider the two neighbourhoods highlighted below:



In both of these three-cell neighbourhoods, the integer in the second column is equal to the product of the integers in the first column. Thus, for each of the four possible first columns, we can compute the second column as follows:

1	1	...	1	-1	...	-1	-1	...	-1	1	...
1	1	...	-1	-1	...	1	-1	...	-1	1	...

From here, we can see that there is exactly one 2×2 Griffin Grid. This is because there are only four possible ways to fill in the first column, and if a 2×2 grid is to be a Griffin Grid, then its second column must be filled in according to the appropriate grid above. If we imagine “chopping off” the first two columns of the infinite grids above, only the first of the four 2×2 grids gives a Griffin Grid.

Continuing in this way, we can examine 2×3 Griffin Grids by filling in the next column. This can be done using the neighbourhoods shown below.



Similar to before, the integers in these neighbourhoods that are in the third column are equal to the product of the integers (in the respective neighbourhood) in the first two columns. Thus, if the cells in the first two columns contain the integers a , b , c , and d , as shown, then the cells in the third column are as shown.

a	c	acd		...
b	d	bcd		

However, we can simplify this a little bit. From earlier, we know that $c = d = ab$. As well, $a^2 = b^2 = c^2 = d^2 = 1$ since each of the variables a , b , c , and d can only take the values 1 and -1 . Therefore, $acd = a(ab)(ab) = aa^2b^2 = a$ and $bcd = b(ab)(ab) = ba^2b^2 = b$. We now have the following simpler and completely general version of the first three columns of a $2 \times \infty$ Griffin Grid.

a	ab	a		...
b	ab	b		

In the same way that the third column was computed from the first and second columns, we can compute the fourth column from the second and third columns. After doing this, we get the diagram below. There is a thick vertical line separating the first three columns from the rest of the $2 \times \infty$ grid.

a	ab	a	1	...
b	ab	b	1	

Consider the six cells to the left of the thick line as a 2×3 grid. The four integers in the first two columns are equal to the products of their neighbours by construction. Denote by C the cell in the first row and the third column. The integer in C is a . By construction, the 1 to its right has been chosen so that the product of the neighbourhood of C , including the 1 to its right, is 1. However, this means the product of the cells in its neighbourhood that are to the left of the vertical line is also equal to 1 (since 1 divided by 1 is 1). These three cells (left of C , below C , and C itself) contain integers that have a product of 1.

A similar argument can be used on the cell below C to confirm that the 2×3 grid to the left of the thick line is a Griffin Grid. This is completely independent of the choice of a and b . There are two possibilities for each of a and b , so there are $2 \times 2 = 4$ Griffin Grids of size 2×3 . They are

1	1	1
1	1	1

1	-1	1
-1	-1	-1

-1	-1	-1
1	-1	1

-1	1	-1
-1	1	-1

This approach can be generalized to determine the number of $2 \times n$ Griffin Grids. To explain this, we first continue to fill out a few more columns in the $2 \times \infty$ grid. We will call the Griffin Grid filled in by starting with variables in the first column and letting them propagate in the way described above the $2 \times \infty$ *variable Griffin Grid*.

a	ab	a	1	a	ab	a	1	a	ab	...
b	ab	b	1	b	ab	b	1	b	ab	

By similar reasoning to before, the first n columns of a $2 \times \infty$ Griffin Grid will form a $2 \times n$ Griffin Grid exactly when both integers in the $(n+1)^{\text{st}}$ column equal 1. If we examine the $2 \times \infty$ variable Griffin Grid, the number of ways to assign values to a and b so that both integers in the $(n+1)^{\text{st}}$ column are 1 will be equal to the number of $2 \times n$ Griffin Grids.

For example, to count the 2×5 Griffin Grids, we look at the 6th column of the $2 \times \infty$ variable Griffin Grid. Both cells contain ab . Thus, the first five columns form a 2×5 Griffin Grid if and only if $ab = 1$. This happens when $a = b = 1$ and when $a = b = -1$, so there are exactly two 2×5 Griffin Grids. Similarly, to count 2×4 Griffin Grids, we count the number of ways to choose a and b so that the 5th column of the $2 \times \infty$ variable Griffin Grid contains only 1. The 5th column contains a and b , so we get a 2×4 Griffin Grid only when $a = b = 1$. Thus, there is exactly one 2×4 Griffin Grid.

Now notice that the columns of the $2 \times \infty$ variable Griffin Grid appear to repeat. Indeed, since each column after the second is computed from the previous two columns, as soon as a pair of two consecutive columns appears for a second time, the columns must repeat. The first and second columns are identical to the fifth and sixth columns, so the columns in the $2 \times \infty$ variable Griffin Grid must repeat with period 4. This means the number of $2 \times n$ Griffin Grids is always the same as the number of $2 \times (n+4)$ Griffin Grids.

The number of 2×1 Griffin Grids is 2 (this can be seen using the method described above). The number of 2×2 Griffin Grids is 1 (this was noted earlier but can also be checked using the method described above). The number of 2×3 Griffin Grids is 4, and the number of 2×4 Griffin Grids is 1. This repeats, so we get the following:

- (i) If n is even, then there is only one $2 \times n$ Griffin Grid.
- (ii) If n is one more than a multiple of 4, then there are two $2 \times n$ Griffin Grids.
- (iii) If n is three more than a multiple of 4, then there are four $2 \times n$ Griffin Grids.

As with $2 \times n$ Griffin Grids, we can fill in the $3 \times \infty$ variable Griffin Grid starting with a , b , and c in the first column. As before, we can assume $a^2 = b^2 = c^2 = 1$. Filling out subsequent columns until we see a repetition, we get

a	ab	ac	bc	c	1	c	bc	ac	ab	a	1	a	ab	
b	abc	1	abc	b	1	b	abc	1	abc	b	1	b	abc	\dots
c	bc	ac	ab	a	1	a	ab	ac	bc	c	1	c	bc	

To determine the number of 3×1 Griffin Grids, we need to find all ways to choose a , b , and c so that the integers in the second column are all 1. In other words, we need all solutions to the system of

$$\begin{aligned} ab &= 1 \\ abc &= 1 \\ bc &= 1 \end{aligned}$$

where a , b , and c are all ± 1 .

The equations $ab = 1$ and $abc = 1$ can be multiplied to get $ababc = 1$ or $a^2b^2c = 1$, so $c = 1$. We can multiply this equation by $bc = 1$ to get $bc^2 = 1$ or $b = 1$. Finally, multiplying $bc = 1$ by $abc = 1$ gives $ab^2c^2 = 1$ so $a = 1$. This method of solving probably seems unnecessarily complicated, but it generalizes nicely. We have shown that the only solution to the system is $a = b = c = 1$. Therefore, there is only one 3×1 Griffin Grid.

To determine the number of 3×2 Griffin Grids, we need to count the number of ways to choose a , b , and c to each be ± 1 so that the entries in the third column are all 1. The middle integer is always 1, so this cell is 1 regardless of how a , b , and c are chosen. The other two equations are both $ac = 1$, so this time the number of Griffin Grids is equal to the number of solutions to the system with just the equation $ac = 1$, where a , b , and c must all be ± 1 .

This system does not restrict b beyond $b = \pm 1$. For $ac = 1$, we need either $a = c = 1$ or $a = c = -1$. This means there are $2 \times 2 = 4$ solutions (2 choices for b and two independent choices for the equal value of a and c). There are four 3×2 Griffin Grids.

Continuing in this way, we can count the number of Griffin Grids of size $3 \times n$ for $n = 1$ through $n = 12$. After that the number of Griffin Grids repeats with period 12. The number of $3 \times n$ Griffin Grids are summarized in the table below.

n	# of $3 \times n$ Griffin Grids
1, 3, 4, 6, 7, 9, 10, or 12 more than a multiple of 12	1
2 or 8 more than a multiple of 12	4
5 or 11 more than a multiple of 12	8

Finally, we leave as an exercise that the number of Griffin Grids of size $4 \times n$ is 16 if n is a multiple of 4 and 1 otherwise.

- (c) Suppose x_1, x_2, \dots, x_k are variables and P_1, P_2, \dots, P_r are each the product of some of the x_i . By convention, we consider a single variable to be a product. For example, it might be that $P_3 = x_9$. We will also allow the P_i to be 1, the so-called “empty product”. We will

show that the number of solutions to the system

$$\begin{aligned} P_1 &= 1 \\ P_2 &= 1 \\ &\vdots \\ P_r &= 1 \end{aligned}$$

is a power of 2, provided each x_i is restricted to the values 1 and -1 . Since each $x_i = \pm 1$, we have that $x_i^2 = 1$, so we can assume that no equation mentions a variable more than once. Following the work from part (b), the number of $m \times n$ Griffin Grids is always equal to the number of solutions to such a system of equations, so this will prove that the number of Griffin Grids is always a power of 2.

The key observation is the following: For any i with $1 \leq i \leq r$, the system of equations above has the exact same set of solutions as the system

$$\begin{aligned} P_1 &= 1 \\ P_2 &= 1 \\ &\vdots \\ P_{i-1} &= 1 \\ P_1 P_i &= 1 \\ P_{i+1} &= 1 \\ &\vdots \\ P_r &= 1 \end{aligned}$$

That is, if we modify the system by multiplying one equation by another equation, the solution set does not change. To see why this is true, we first suppose that $(x_1, x_2, \dots, x_m) = (c_1, c_2, \dots, c_m)$ is a solution to the original system. Every equation in the second system appears in the original system except $P_1 P_i = 1$, so all equations in the second system except $P_1 P_i = 1$ are automatically satisfied by the choice of the x_i . By assumption, $P_1 = 1$ and $P_i = 1$, so we also have $P_1 P_i = 1$ as well. We have confirmed that $(x_1, x_2, \dots, x_m) = (c_1, c_2, \dots, c_m)$ is a solution to the second system.

A nearly identical argument shows that a solution to the second system is also a solution to the first system, proving that the two systems have the exact same set of solutions.

We will now prove that the number of solutions is a power of 2 by induction on k , the number of variables.

If $k = 1$, then the only possible equations are $x_1 = 1$ and $1 = 1$. If $x_1 = 1$ appears, then there is one solution. If $x_1 = 1$ does not appear, then all equations are $1 = 1$, so the equations do not restrict the value of x_1 beyond $x_1 = 1$ and $x_1 = -1$. Therefore, the number of solutions is either 1 or 2, both of which is a power of 2.

Now suppose the number of solutions to such a system of equations in at most $k - 1$ variables is always a power of 2.

Consider a system of equations in k variables, x_1 through x_k . If it happens to be that none of the equations mention x_k , then the system can be viewed as having at most $k - 1$

variables. By induction, the number of solutions to this system (with x_k ignored) is a power of 2. That is, the number of ways to choose values of x_1 through x_{k-1} so that the equations are all satisfied is 2^d for some non-negative integer d .

Since x_k is not mentioned in the equations, every one of these 2^d solutions allows us to choose $x_k = 1$ or $x_k = -1$. Therefore, the system has $2 \times 2^d = 2^{d+1}$ solutions.

Now suppose x_k appears in at least one equation. By possibly reordering the equations, we can assume that $P_1 = 1$ mentions the variable x_k . Consider the system

$$\begin{aligned} P_1 &= 1 \\ Q_2 &= 1 \\ Q_3 &= 1 \\ &\vdots \\ Q_r &= 1 \end{aligned}$$

where $Q_i = P_i$ if P_i does not mention x_k , and $Q_i = P_1 P_i$ if P_i mentions x_k . In other words, equations from P_2 through P_r that mention x_k get multiplied by P_1 and the others are left alone. From earlier, the new system has the same solutions as the original system.

If P_i mentions x_k , then it must mention it exactly once. This means the equation $Q_i = 1$ mentions x_k exactly twice, but these occurrences can be deleted since $x_k^2 = 1$.

The point is that we have converted the given system to a new system that has the same solution set and $P_1 = 1$ is the only equation that mentions x_k . Now consider the system

$$\begin{aligned} Q_2 &= 1 \\ Q_3 &= 1 \\ &\vdots \\ Q_r &= 1 \end{aligned}$$

which does not mention x_k . By induction, the number of solutions to this system is 2^d for some non-negative integer d . Each such solution to this smaller system is an assignment of values to the variables x_1 through x_{k-1} . Once these values are chosen, the equation $P_1 = 1$, which mentions x_k , will determine the value of x_k .

Therefore, the number of solutions to the original system is 2^d . Using induction, we have now shown that the number of solutions to any system of this type is a power of 2. Therefore, the number of $m \times n$ Griffin Grids is always a power of 2.

Finally, note that the system arising from the $m \times \infty$ variable Griffin Grid always has m variables. There are at most 2^m ways to assign a value of 1 or -1 to each variable, so the number of solutions must be a power of 2 that is at most 2^m .

- (e) Once again, we imagine filling in the $m \times \infty$ variable Griffin Grid starting with variables a_1, a_2, \dots, a_m in the first column.

Every cell in the grid contains either 1 or the product of some of a_1 through a_m , where each variable appears at most once in such a product. There are exactly 2^m possible expressions that can go in the cells. In any two consecutive columns, there are $2n$ cells, which means

there are $(2^m)^{2n} = 2^{2mn}$ possible ways that the $2n$ cells in two consecutive columns can be filled. In practice, not all of these possibilities will actually occur.

Among the first $2^{2mn} + 2$ columns, there are $2^{2mn} + 1$ pairs of consecutive columns, which means there must be two pairs of consecutive columns that are identical.

Since there are two identical pairs of consecutive columns, we can choose the earliest ones. More precisely, let i be the smallest positive integer such that there is $j > i$ with the property that the i^{th} column equals the j^{th} column and the $(i + 1)^{\text{st}}$ column equals the $(j + 1)^{\text{st}}$ column.

Suppose that $i > 1$. Let P_1, P_2, \dots, P_m be the products in column $i - 1$ let Q_1, Q_2, \dots, Q_m be the products in column i , and let R_1, R_2, \dots, R_m be the products in column $i + 1$.

By construction, we have that $R_1 = P_1 Q_1 Q_2$ (after possibly removing the squares of some variables). This equation can be multiplied on both sides by $Q_1 Q_2$ to get the equation $R_1 Q_1 Q_2 = P_1 (Q_1)^2 (Q_2)^2$. Since the square of every variable in Q_1 and Q_2 is 1, we also have $(Q_1)^2 = (Q_2)^2 = 1$, so $P_1 = Q_1 Q_2 R_1$.

We also have that $R_2 = P_2 Q_1 Q_2 Q_3$, and it similarly follows that $P_2 = Q_1 Q_2 Q_3 R_2$.

Continuing in this way, we get that each P_i can be computed in terms of the Q_i and R_i . In other words, column $i - 1$ can be computed directly from columns i and $i + 1$.

Because of our assumptions about i and j , it follows that column $i - 1$ must be equal to column $j - 1$. However, this means columns $i - 1$ and i are identical to columns $j - 1$ and j , which contradicts the minimality of i .

We conclude that $i > 1$ is impossible, so it must be that $i = 1$ which means column j equals column 1 and column $j + 1$ equals column 2.

Therefore, the products in column j are all single variables, and in fact, they are the variables a_1, a_2, \dots, a_m in order from top to bottom. The only way to assign the variables so that the entire column contains 1 is to set $a_1 = 1, a_2 = 1$, and so on to $a_m = 1$. Therefore, there is exactly one $m \times (j - 1)$ Griffin Grid.

- (f) Most of the work was done in part (e). Let's assume that column $j > 1$ equals column 1 and column $j + 1$ equals column 2. Thus, we are assuming that the entries in column j are a_1, a_2, \dots, a_m from top to bottom and that the entries in column $j + 1$ are $a_1 a_2, a_1 a_2 a_3, a_2 a_3 a_4, \dots, a_{m-2} a_{m-1} a_m, a_{m-1} a_m$ from top to bottom.

Suppose that P_1, P_2, \dots, P_m are the products in column $j - 1$. The top cell in column $j + 1$ contains $a_1 a_2$ since column $j + 1$ equals column 2. The expression must also be $a_1 a_2 P_1$ by the way the $m \times \infty$ variable Griffin Grid is filled in. From $a_1 a_2 = a_1 a_2 P_1$, we get $P_1 = 1$.

Looking at the second cell in column $j + 1$, we have that it is equal to $a_1 a_2 a_3$, but it is also equal to $P_2 a_1 a_2 a_3$, and so $P_2 = 1$.

Continuing in this way, we get that every entry in column $j - 1$ equals 1. Since any assignment of values to a_1, a_2, \dots, a_m will lead to column $j - 1$ having all 1s, we conclude that there are 2^m Griffin Grids of size $m \times (j - 2)$. As a final note, observe that column 1 and column 2 always (for every m) contain products that are not equal to 1, so we know that $j - 1 > 2$, from which it follows that $j - 2 > 1$. This means we do not need to worry about the possibility that $j - 2 = 0$.