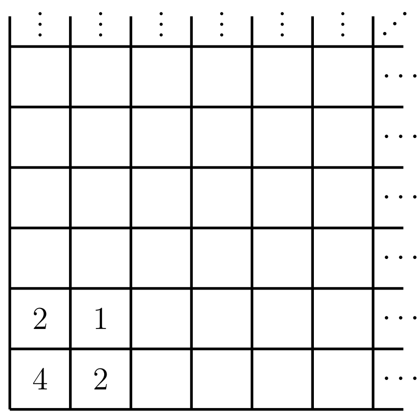
Problem of the Month

Problem 0: September 2020

Problem

A rectangular array extends up and to the right with infinitely many rows and infinitely many columns. Integers are placed in the four "bottom-left" cells as shown with 4 in the bottom-left corner, 2 in each of the cells sharing a side with the cell containing 4, and 1 in the cell immediately to the right and above the 4.



When referring to rows and columns, we start the enumeration from the bottom and the left. For example, the "third row" refers to the third row from the bottom.

Integers are placed in the remaining cells *recursively* as follows:

- In the first and second rows, each remaining cell contains the sum of the integer in the cell immediately to its left and twice the integer two cells to its left. For example, the third cell in the first row contains the integer $2 + 2(4) = 10$.
- Cells in or above the third row contain the sum of the integer in the cell immediately below and twice the integer in the cell two below. For example, the second cell in the third row contains the integer $1 + 2(2) = 5$.

We will denote by $f(m, n)$ the integer in the m^{th} row and the n^{th} column.

(a) Show that every cell other than those in the first two columns contains the sum of the integer in the cell immediately to its left and twice the integer in the cell two to its left. That is, show that $f(m, n) = f(m, n-1) + 2f(m, n-2)$ for all integers $m \geq 1$ and $n \geq 3$.

(b) Prove that $f(m, m)$ is a perfect square for every integer $m \geq 1$. In other words, prove that all of the cells on the diagonal contain perfect squares.

(c) Determine the value of $f(456, 789)$.

Hint

Before doing any of the parts of this problem, it is a good idea to fill in some of the table in order to gain some intuition. However, working out a few cases and convincing yourself something is true is not the same as writing a formal proof.

(a) By carefully using the rules defining the numbers in the cells, an informal argument can be given using a bit of algebra. However, to give a formal argument, it is highly recommended to use *mathematical induction*. You may want to read about this before trying to write a proof.

(b) Compute a few of the diagonal entries. Do you notice anything about what numbers were squared to get these diagonal entries? Once again, mathematical induction will be useful in formalizing any observations you make.

(c) Can you find a formula for the entries in the second row and second column?

Problem of the Month

Solution to Problem 0: September 2020

We will use the notation introduced in the problem statement and denote by $f(m,n)$ the integer in the cell in the m^{th} row and the n^{th} column. Using this notation, the example given in the first bullet point in the problem statement translates to $f(1,3) = f(1,2) + 2f(1,1) = 2 + 2(4) = 10$.

Before attempting any of these problems, it is a good idea to fill in a bit more of the array in order to gain some intuition:

$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\iddots$
						$\cdots$
34	17	85	119	289		$\cdots$
14	7	35	49	119		$\cdots$
10	5	25	35	85		$\cdots$
2	1	5	7	17		$\cdots$
4	2	10	14	34		$\cdots$

There are several observations you might make at this point. For example, 4, 1, 25, 49, and 289 are all perfect squares, so the claim in part (b) might seem plausible. You might also notice that the array appears to be symmetric in the diagonal. That is, it seems $f(m,n) = f(n,m)$ for all positive integers m and n.

(a) Using our notation, the task is to show that

$$f(m,n) = f(m,n-1) + 2f(m,n-2) \tag{1}$$

for every pair of positive integers (m,n) with $n \geq 3$. For cells in the bottom two rows ($m = 1$ or $m = 2$), the integers are defined in this way. That is, Equation (1) holds when $m = 1$ and $m = 2$ by definition. Looking back at the partially-filled array above, you might want to check that this identity holds for a few other cells. For example, $f(4,5) = 119$, $f(4,4) = 49$, and $f(4,3) = 35$ and indeed $119 = 49 + 2(35)$.

Now we note that we can use the fact that Equation (1) holds for cells in the first and second row (for $m = 1$ and $m = 2$) to show that Equation (1) also holds for cells in the third row (for $m = 3$). Consider the pair $(3,n)$ for some $n \geq 3$. We know that $f(3,n) = f(2,n) + 2f(1,n)$ by the rule in the second bullet point in the problem statement. Since Equation (1) holds for $m = 1$ and $m = 2$ we have $f(2,n) = f(2,n-1) + 2f(2,n-2)$ and $f(1,n) = f(1,n-1) + 2f(1,n-2)$. Putting these together we get

$$\begin{aligned}
f(3,n) &= f(2,n) + 2f(1,n) \\
&= [f(2,n-1) + 2f(2,n-2)] + 2[f(1,n-1) + 2f(1,n-2)] \\
&= [f(2,n-1) + 2f(1,n-1)] + 2[f(2,n-2) + 2f(1,n-2)]
\end{aligned}$$

Notice that $f(2,n-1) + 2f(1,n-1) = f(3,n-1)$ and $f(2,n-2) + 2f(1,n-2) = f(3,n-2)$, again by how the array is defined. This gives us that $f(3,n) = f(3,n-1) + 2f(3,n-2)$ which is Equation (1) for $m = 3$, and so the equation holds for the third row.

We could now proceed to argue that since Equation (1) holds for $m = 1$, $m = 2$ and $m = 3$ it must be the case that Equation (1) also holds for $m = 4$. You might find it useful to try to write down this argument yourself. (The argument will be very similar to the one presented above, and you should find that you only need to explicitly use the fact that Equation (1) holds for $m = 2$ and $m = 3$ in your argument.) Instead we proceed more generally to argue that if Equation (1) holds for $m = 1$, $m = 2$, and so on up to $m = r$ then Equation (1) must also hold for $m = r + 1$. The rough idea is to show that you can always get "the next row".

Suppose that $r \geq 2$ is an integer and that Equation (1) holds for the cells in the first r rows. That is, we assume Equation (1) holds for all $n \geq 3$ when $m = 1$, and when $m = 2$, and so on up to when $m = r$. This means we have the following:

Definition 1: Given a positive integer t, we have $f(m,t) = f(m-1,t) + 2f(m-2,t)$ for all $m \geq 3$. (This is by the rule in the second bullet point in the problem statement.)

Assumption 1: Given a positive integer $t \leq r$, we have $f(t,n) = f(t,n-1) + 2f(t,n-2)$ for all $n \geq 3$. (This is our assumption from above.) In particular, we are assuming that $f(r,n) = f(r,n-1) + 2f(r,n-2)$ and $f(r-1,n) = f(r-1,n-1) + 2f(r-1,n-2)$.

We will show that the above statements imply that the identity also holds for the cells in the $(r+1)^{\text{st}}$ row. Consider a cell in the $(r+1)^{\text{st}}$ row that is not in the first two columns. That is, consider the pair $(r+1,n)$ for some $n \geq 3$. We wish to show that Equation (1) holds for this pair, that is, $f(r+1,n) = f(r+1,n-1) + 2f(r+1,n-2)$. Here is the calculation, using the assumptions above:

$$\begin{aligned}
f(r+1,n) &= f(r,n) + 2f(r-1,n) && \text{(by Definition 1)} \\
&= \big[f(r,n-1) + 2f(r,n-2)\big] + 2\big[f(r-1,n-1) + 2f(r-1,n-2)\big] \\
& && \text{(by Assumption 1)} \\
&= f(r,n-1) + 2f(r-1,n-1) + 2f(r,n-2) + 4f(r-1,n-2) \\
&= \big[f(r,n-1) + 2f(r-1,n-1)\big] + 2\big[f(r,n-2) + 2f(r-1,n-2)\big] \\
&= f(r+1,n-1) + 2f(r+1,n-2) && \text{(by Definition 1)}
\end{aligned}$$

We have shown that if the identity holds in the first r rows, then it holds in the first $r+1$ rows. Since it holds in the first two rows, it holds in the first three rows. Since it holds in first three rows, it holds in the first four rows. This continues indefinitely to imply that the identity holds in every row. We have just used what is known as *strong induction.*

(b) You may have noticed that the entries in the second row are identical to those in the second column. This is true because the second row and the second column start with the same two integers (2 followed by 1), and all subsequent integers in each are determined in the same way (where each integer depends on the two integers before it). With our notation, this means that for every positive integer n we have $f(n,2) = f(2,n)$.

Something more subtle that you may have noticed is that for all positive integers m and n, we have $f(m,n) = f(m,2)f(2,n)$. You may want to go back to the array and check this for a few pairs (m,n). Using this, and the fact above that $f(n,2) = f(2,n)$ for every $n \geq 1$, we get that

$$f(n,n) = f(n,2)f(2,n) = f(2,n)^2$$

which establishes that $f(n,n)$ is a perfect square for every $n \geq 1$. In fact, it establishes that the integers on the diagonal are the squares of the integers in the second row (or second column).

We will finish the solution to part (b) with a somewhat informal explanation of why $f(m,n) = f(m,2)f(2,n)$ for all positive integers m and n. This fact will also be used in the solution to part (c). A formal proof of this fact can be found after the solution to part (c).

The main observation is that each row is a "scalar multiple" of the second row. To see this, we argue as follows: Suppose two sequences $a_1, a_2, a_3, \ldots$ and $b_1, b_2, b_3, \ldots$ satisfy $a_n = a_{n-1} + 2a_{n-2}$ and $b_n = b_{n-1} + 2b_{n-2}$ for $n \geq 3$. This means each sequence is entirely determined by its first two terms. Also, both sequences satisfy the same recursive rule.

Now suppose there is some constant c so that $b_1 = ca_1$ and $b_2 = ca_2$, that is, a_1 and a_2 can be scaled by the same factor to get b_1 and b_2, respectively. Then

$$b_3 = b_2 + 2b_1 = ca_2 + 2ca_1 = c(a_2 + 2a_1) = ca_3,$$

which says that b_3 can be obtained by scaling a_3 by the same factor. Continuing, we also have

$$b_4 = b_3 + 2b_2 = ca_3 + 2ca_2 = c(a_3 + 2a_2) = ca_4.$$

This reasoning can be continued to show that $b_n = ca_n$ for every positive integer n. The reason this happens is because the way in which subsequent terms rely on previous terms is *linear* (a slightly different use of the word than you may be used to). Roughly speaking, linear in this context means the terms in the sequence are obtained by scaling some of the previous terms and adding the results together.

Now note that the first two columns in the array contain sequences of integers with the above properties where b_n corresponds to the first column, a_n corresponds to the second column, and the constant is $c = 2$. Using the reasoning above, we can deduce that the integers in the first column are each exactly twice the integer to their right (in the second column). Now consider the second row and the m^{th} row. We know that the first two terms in the second row are 2 and 1, in that order, and the first two terms in the m^{th} row must be $2f(m,2)$ and $f(m,2)$, in that order. In part (a), we showed that the sequence in the m^{th} row satisfies the same recursive definition as the one in the second row. This means that the m^{th} row and second row contain sequences of integers with the above properties with $c = f(m,2)$. It follows that the integers in the m^{th} row must all be c times their corresponding integer in the second row. In other words, we have $f(m,n) = cf(2,n) = f(m,2)f(2,n)$.

(c) Using the identity $f(m,n) = f(m,2)f(2,n) = f(2,m)f(2,n)$, we get

$$f(456,789) = f(456,2)f(2,789) = f(2,456)f(2,789)$$

so we can find $f(456,789)$ by finding $f(2,456)$ and $f(2,789)$ and taking their product.

Since the rest of the solution will focus on the entries in the second row, we will simplify notation and define $g(n)$ to be $f(2,n)$ for each integer $n \geq 1$. From the table at the beginning of the solution, we get $g(1) = 2$, $g(2) = 1$, $g(3) = 5$, $g(4) = 7$, $g(5) = 17$. Continuing to compute terms, it can be checked that $g(6) = 31$, $g(7) = 65$, $g(8) = 127$, $g(9) = 257$, and $g(10) = 511$.

While the pattern may have been difficult to detect before, it may be easier to guess from the sequence

$$2, 1, 5, 7, 17, 31, 65, 127, 257, 511$$

as these numbers are all 1 away from a power of 2. In fact, 2 is 1 more than 2^0, 1 is 1 less than 2^1, 5 is 1 more than 2^2, 7 is 1 less than 2^3, and so on. Following this reasoning, we can see that when $1 \leq n \leq 10$, we have that

$$g(n) = 2^{n-1} + (-1)^{n-1}.$$

Proceeding once more by strong induction, we can prove that this identity holds for all $n \geq 1$. Assume for some $r \geq 2$ that $g(k) = 2^{k-1} + (-1)^{k-1}$ for every k from 1 to r inclusive. In particular, this implies $g(r) = 2^{r-1} + (-1)^{r-1}$ and $g(r-1) = 2^{r-2} + (-1)^{r-2}$. We will call these equations **Equation (2)** and **Equation (3)** respectively. By the definition of the integers in the second row, we have $g(r+1) = g(r) + 2g(r-1)$ which we will refer to as **Equation (4)**. Then we have

$$\begin{aligned}
g(r+1) &= g(r) + 2g(r-1) && \text{(Equation (4))} \\
&= \left(2^{r-1} + (-1)^{r-1}\right) + 2\left(2^{r-2} + (-1)^{r-2}\right) && \text{(Equations (2) and (3))} \\
&= 2^{r-1} + (-1)^{r-1} + 2^{r-1} + 2(-1)^{r-2} \\
&= 2^{r-1} + 2^{r-1} + (-1)^{r-1} + 2(-1)^{r-2} \\
&= 2\left(2^{r-1}\right) + (-1)^{r-2}(-1+2) \\
&= 2^r + (-1)^{r-2}(1) \\
&= 2^r + (-1)^{r-2}(-1)^2 \\
&= 2^r + (-1)^r
\end{aligned}$$

where the calculation after the second two equations is just arithmetic using exponent laws. This means $g(r+1) = 2^{(r+1)-1} + (-1)^{(r+1)-1}$, so by strong induction, we have that $g(n) = f(2,n) = 2^{n-1} + (-1)^{n-1}$ for all $n \geq 1$. Using the calculation from above, we get

$$\begin{aligned}
f(456, 789) &= f(2, 456)f(2, 789) \\
&= \left(2^{455} + (-1)^{455}\right)\left(2^{788} + (-1)^{788}\right) \\
&= \left(2^{455} - 1\right)\left(2^{788} + 1\right) \\
&= 2^{1243} - 2^{788} + 2^{455} - 1.
\end{aligned}$$

Proof that $f(m,n) = f(m,2)f(2,n)$ for all positive integers m and n.

This proof is by strong induction and features calculations very similar to those in the solution to part (a). For notational convenience, we will refer to the equation $f(m,n) = f(m,2)f(2,n)$ as Equation (2).

Since $f(2,2) = 1$, we have that $f(2,n) = f(2,2)f(2,n)$ and $f(m,2) = f(m,2)f(2,2)$, which verifies Equation (2) in the case that $m = 2$ or $n = 2$. To see that Equation (2) holds when $m = 1$ for all n, first observe that $f(1,1) = 2f(2,1)$ and $f(1,2) = 2f(2,2)$. Now assume for some integer $r \geq 2$ that $f(1,n) = 2f(2,n)$ for $n = 1$, $n = 2$, and so on up to $n = r$. In particular, we assume $f(1,r) = 2f(2,r)$ and $f(1,r-1) = 2f(2,r-1)$. Then

$$\begin{aligned} f(1,r+1) &= f(1,r) + 2f(1,r-1) \\ &= 2f(2,r) + 2\left(2f(2,r-1)\right) \\ &= 2\left(f(2,r) + 2f(2,r-1)\right) \\ &= 2f(2,r+1) \end{aligned}$$

so $f(1,r+1) = 2f(2,r+1)$ as well. By strong induction, it follows that $f(1,n) = 2f(2,n)$ for all $n \geq 1$. Noting that $f(1,2) = 2$, this means $f(1,n) = f(1,2)f(2,n)$, so Equation (2) holds for all n when $m = 1$.

We will now use induction again to continue to show row-by-row that Equation (2) holds for all positive integers m and n. We know that it is true for all n when $m = 1$ and when $m = 2$. Suppose that $r \geq 2$ is an integer and that Equation (2) holds in the first r rows. In particular, we are assuming $f(r,n) = f(r,2)f(2,n)$ and $f(r-1,n) = f(r-1,2)f(2,n)$ for all $n \geq 1$. For any $n \geq 1$, we have

$$\begin{aligned} f(r+1,n) &= f(r,n) + 2f(r-1,n) \\ &= \left(f(r,2)f(2,n)\right) + 2\left(f(r-1,2)f(2,n)\right) \\ &= \left(f(r,2) + 2f(r-1,2)\right)f(2,n) \\ &= f(r+1,2)f(2,n) \end{aligned}$$

where the first and last equalities are by the definition of the entries in the $(r+1)^{\text{st}}$ row. By strong induction, it follows that $f(m,n) = f(m,2)f(2,n)$ for all positive integers m and n.

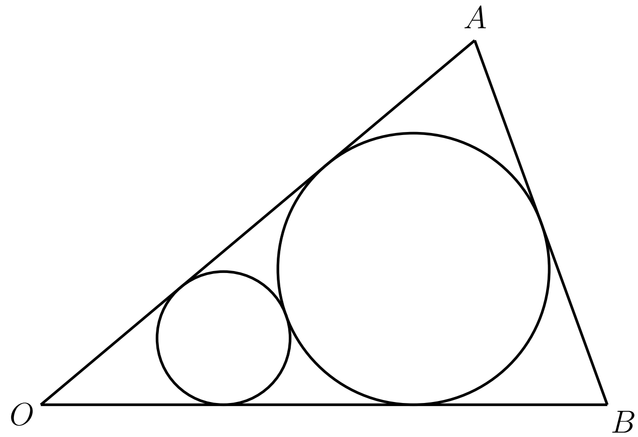
Can you see how to use the fact that $f(n,2) = f(2,n)$ for all positive integers n to prove that $f(m,n) = f(n,m)$ for all positive integers m and n?

Problem of the Month

Problem 1: October 2020

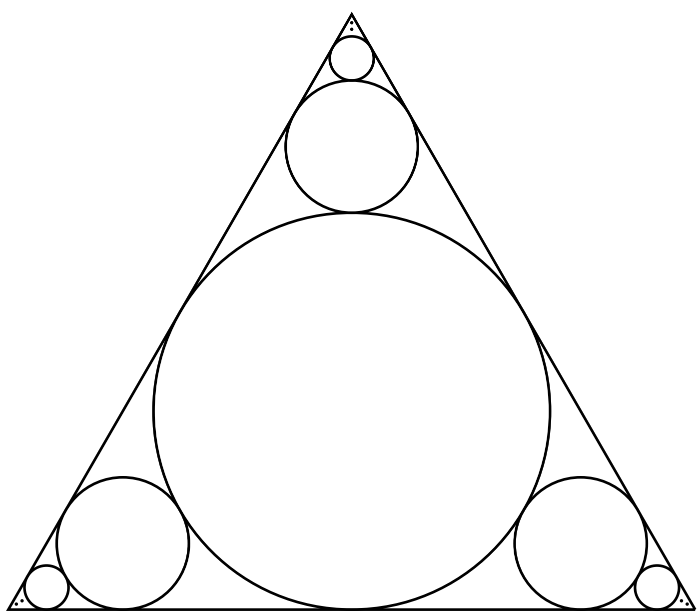
Problem

(a) Let θ be an angle with $0 < \theta < 45°$. In the diagram, points A and B are configured so that $\angle AOB = 2\theta$ and $\triangle AOB$ is isosceles with $AO = BO$.



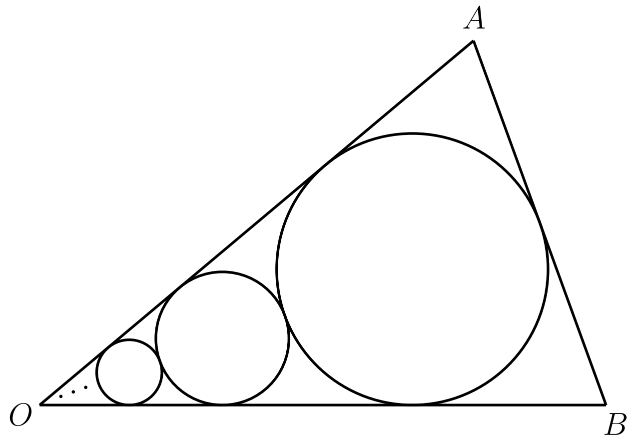
A circle is inscribed in $\triangle AOB$ and another circle is drawn so that it is tangent to the larger circle as well as OA and OB. In terms of θ, find the ratio of the radius of the larger circle to the radius of the smaller circle.

(b) Similar to part (a), an equilateral triangle has a circle inscribed in it. Three circles are then drawn, each tangent to two of the sides of the triangle as well as the larger circle. Another three circles are then drawn, each tangent to two of the three sides of the triangle as well as one of the circles drawn in the previous step.



If this process is continued indefinitely, what fraction of the area of the triangle is covered by circles?

(c) Suppose $\triangle AOB$ and θ are as they were defined in part (a). The process of drawing a circle tangent to OA, OB, and the smallest circle is repeated forever. What fraction of the area of $\triangle AOB$ is covered by circles? Your answer should be in terms of θ.



The result of part (c) can be applied to solve part (b). Can you see how?

Hint

(a) There are plenty of ways to approach this problem. One useful construction is to connect the centres of the circles to each other, then draw a perpendicular from each centre to the line OB.

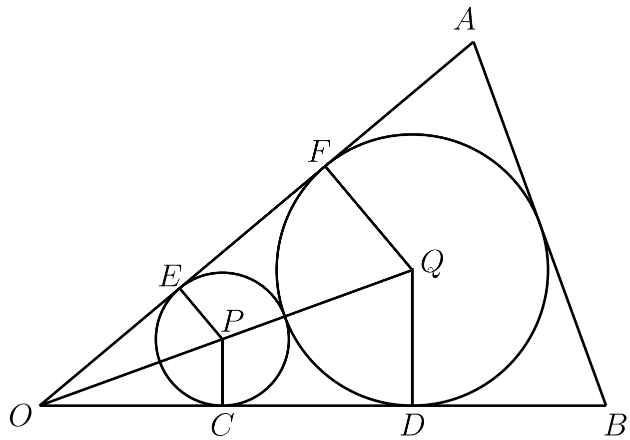
(b) It is possible to apply part (a) with $\theta = 30°$. While there are other ways to do this part, the easiest probably involves finding the sum of an infinite geometric series. You may want to look up how this is done.

(c) In some sense, this is a more general version of part (b). Finding the sum of a geometric series will be useful again here, but the terms of the series will be in terms of the variable θ. You might also find it useful to express the area of the triangle and the area of the largest circle in ways that are easy to compare.

Problem of the Month

Solution to Problem 1: October 2020

(a) **Solution 1**: Let P and Q be the centres of the smaller and larger circles, respectively, and let C and D be the points of tangency of the smaller and larger circles to OB. Similarly, let E and F be the points of tangency of the smaller and larger circles to OA.



Line segments PC and PE are radii of the smaller circle and thus are equal. Line segments OE and OC are equal because the distances from two points of tangency to the point where the tangents intersect are equal. We also have that $\angle OEP = \angle OCP = 90^\circ$ because a radius drawn to a point of tangency is perpendicular to that tangent. Therefore, $\triangle OCP$ is congruent to $\triangle OEP$ by side-angle-side congruence. This means $\angle EOP = \angle COP$, so OP is the angle bisector of $\angle EOC = \angle AOB$. By a similar argument, OQ is the angle bisector of $\angle FOD = \angle AOB$. This tells us that P and Q both lie on the angle bisector of $\angle AOB$. Therefore, OPQ is a line segment, and $\angle POC = \frac{2\theta}{2} = \theta$.

Let R be the radius of the larger circle and let r be the radius of the smaller circle. It follows from the fact that a radius is perpendicular to the corresponding tangent that line segment PQ passes through the point where the two circles are tangent. This means $PQ = r + R$. Also, since $\frac{PC}{OP} = \sin\theta$ and $PC = r$, we get $OP = \frac{r}{\sin\theta}$. Therefore, $OQ = \frac{r}{\sin\theta} + r + R$ ($\sin\theta \neq 0$ because $0^\circ < \theta < 45^\circ$). We also know $\sin\theta = \frac{QD}{OQ}$ and $QD = R$, so

$$\frac{R}{OQ} = \sin\theta = \frac{R}{\frac{r}{\sin\theta} + r + R}$$

Multiplying through by the denominator of the expression on the right gives

$$\begin{aligned} R &= \sin\theta\left(\frac{r}{\sin\theta} + r + R\right) \\ &= r + r\sin\theta + R\sin\theta. \end{aligned}$$

Bringing all terms with an R to one side and factoring, we get

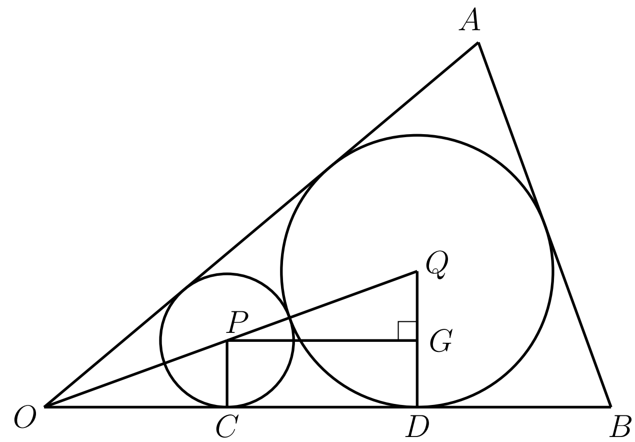
$$R(1 - \sin\theta) = r(1 + \sin\theta)$$

and so now we can solve for $\frac{R}{r}$ to get

$$\frac{R}{r} = \frac{1+\sin\theta}{1-\sin\theta}.$$

This expression is defined because $0 < \theta < 45^\circ$, so $\sin\theta \neq 1$.

Solution 2: Let P and Q be the centres of the smaller and larger circles, respectively, and let C and D be the points of tangency of the smaller and larger circles to OB. Let G be the point on QD so that PG is perpendicular to QD.



As mentioned in the first solution, $\angle PCD$ and $\angle QDC$ are both right angles and PQ passes through the point at which the two circles are tangent. As well, O, P, and Q lie on the angle bisector of $\angle AOB$.

This means PG is parallel to OD so $\angle QPG = \angle QOD$. We also have $\angle QGP = \angle QDO = 90^\circ$, so $\triangle PQG$ is similar to $\triangle OQD$. Therefore, since Q lies on the angle bisector of $\angle AOB$, we have $\frac{QG}{PQ} = \frac{QD}{OQ} = \sin\theta$.

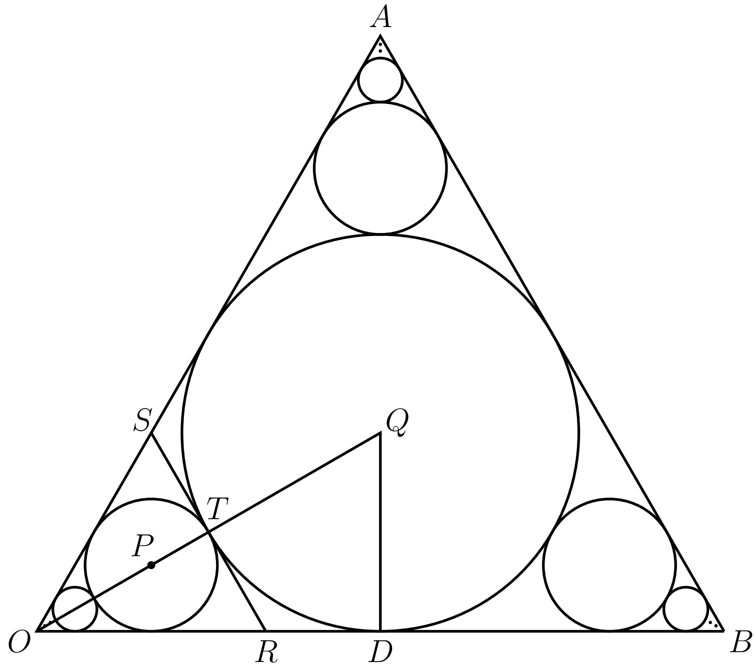
Let R be the radius of the larger circle and r be the radius of the smaller circle. Since quadrilateral $PGDC$ has three right angles, it is a rectangle, which means $GD = PC = r$. Thus, $QG = R - r$. We also have that $PQ = R + r$, so

$$\sin\theta = \frac{R-r}{R+r}.$$

This can be rearranged to get $R\sin\theta + r\sin\theta = R - r$ or $R(1-\sin\theta) = r(1+\sin\theta)$, and therefore

$$\frac{R}{r} = \frac{1+\sin\theta}{1-\sin\theta}.$$

(b) We will label the triangle $\triangle OAB$. Let P be the centre of the largest circle in the bottom-left corner and Q be the centre of the largest circle. Let the circles with centres P and Q be tangent at T, and suppose the common tangent intersects OA at S and OB at R. Finally, let D be the point at which the circle centred at Q is tangent to OB.



By the reasoning in part (a), points O, P, and Q all lie on the angle bisector of $\angle AOB$, so $\angle SOT = \angle ROT$. We also have, by circle properties, that $\angle STO = \angle RTO = 90°$, which means $\triangle STO$ is congruent to $\triangle RTO$ by angle-side-angle congruence (these two triangles share side OT). It follows that $\angle OST = \angle ORT$ and since $\angle SOR = 60°$, we get that $\triangle SOR$ is equilateral. [Note: If we only assume that $\triangle AOB$ is isosceles, this argument still shows that $\triangle SOR$ is similar to $\triangle AOB$.]

Suppose the side lengths of $\triangle AOB$ are equal to x.

Since OQ is the angle bisector of $\angle AOB$, we have that $\angle QOD = \dfrac{60°}{2} = 30°$. Since OB is tangent to the largest circle at point D, we have that $\angle ODQ = 90°$. Therefore, $\triangle ODQ$ is a $30°$-$60°$-$90°$ triangle which implies $\dfrac{QD}{OD} = \dfrac{1}{\sqrt{3}}$. Since $\triangle AOB$ is equilateral, D is the midpoint of OB. [The proof of this is left as an exercise. One way to show it is to connect Q to B and show that $\triangle QOD$ is congruent to $\triangle QBD$.] This means $OD = \dfrac{x}{2}$ so $QD = \dfrac{x}{2\sqrt{3}}$. We have found the radius of the largest circle in terms of the side length of the triangle.

By part (a) with $\theta = 30°$, the ratio of the radius of the circle centred at Q to the radius of the circle centred at P is

$$\frac{1+\sin 30°}{1-\sin 30°} = \frac{1+\frac{1}{2}}{1-\frac{1}{2}} = 3.$$

This means the radius of the circle centred at P is $\dfrac{1}{3} \times QD = \dfrac{x}{6\sqrt{3}}$. By symmetry, the other two circles tangent to the largest circle have this same radius.

We showed earlier that $\triangle SOR$ is equilateral. This means we can apply the argument above again to get that the radii of the three next largest circles are each $\dfrac{1}{3} \times \dfrac{x}{6\sqrt{3}}$. We could then draw the common tangent to the circle centred at P and the next largest circle to repeat the argument. The radius will be multiplied by $\frac{1}{3}$ each time.

Therefore, the total area of the circles is represented by the following infinite series:

$$\pi\left(\frac{x}{2\sqrt{3}}\right)^2 + 3\pi\left(\frac{x}{6\sqrt{3}}\right)^2 + 3\pi\left(\frac{x}{18\sqrt{3}}\right)^2 + \cdots.$$

The first term in the sum is equal to the area of the largest circle. The second term is equal to the total area of the three next largest circles (those tangent to the largest circle). The third term is equal to the total area of the three next largest circles, and so on.

After some simplification, the sum above is equivalent to

$$\frac{\pi x^2}{12} + \frac{3\pi x^2}{12}\left(\frac{1}{9} + \frac{1}{9^2} + \frac{1}{9^3} + \cdots\right).$$

If a and r are real numbers with $-1 < r < 1$, then we can find the sum of the geometric series $a + ar + ar^2 + \cdots$ using the formula $a + ar + ar^2 + \cdots = \dfrac{a}{1-r}$. Our expressions for the total area of the circles involves a geometric series with $a = r = \dfrac{1}{9}$, so

$$\frac{1}{9} + \frac{1}{9^2} + \frac{1}{9^3} + \cdots = \frac{\frac{1}{9}}{1-\frac{1}{9}} = \frac{1}{8}.$$

Therefore, the total area of the circles is

$$\begin{aligned}\frac{\pi x^2}{12} + \frac{3\pi x^2}{12} \times \frac{1}{8} &= \frac{8\pi x^2 + 3\pi x^2}{96}\\ &= \frac{11\pi x^2}{96}.\end{aligned}$$

There are several ways to determine the area of $\triangle AOB$ in terms of its side length, x. One way is to use that the area of a triangle with side lengths a and b meeting at angle α is $\dfrac{1}{2}ab\sin\alpha$. Thus, $\triangle AOB$ has area $\dfrac{1}{2}x^2\sin 60^\circ = \dfrac{\sqrt{3}x^2}{4}$ since each of its angles measures 60° and its sides all have length x. The fraction of the triangle that is covered by circles is

$$\frac{\frac{11\pi x^2}{96}}{\frac{\sqrt{3}x^2}{4}} = \frac{11\pi}{24\sqrt{3}}.$$

(c) This calculation will be rather similar to the one in part (b). Suppose the radius of the larger circle is R and set $\alpha = \dfrac{1-\sin\theta}{1+\sin\theta}$. From part (a), we have that $\dfrac{r}{R} = \alpha$, or $r = \alpha R$. Following the reasoning in the solution to part (b), we can show that the total area of the circles is given by

$$\pi R^2 + \pi(\alpha R)^2 + \pi(\alpha^2 R)^2 + \pi(\alpha^3 R)^2 + \cdots.$$

Factoring out πR^2, this is equal to

$$\pi R^2(1 + \alpha^2 + \alpha^4 + \alpha^6 + \cdots).$$

Since $0^\circ < \theta < 45^\circ$, we have that $0 < \sin\theta < 1$ (in fact, $\sin\theta < \frac{\sqrt{2}}{2}$, but having $\sin\theta < 1$ is good enough for what follows). This means $0 < 1 - \sin\theta < 1$. Furthermore, since $\sin\theta$ is positive, we have $1 - \sin\theta < 1 + \sin\theta$. It follows that

$$0 < \frac{1-\sin\theta}{1+\sin\theta} < 1$$

or $0 < \alpha < 1$ and so $0 < \alpha^2 < 1$. Therefore, using the formula for the sum of a geometric series (see part (b)), the total area of the circles is

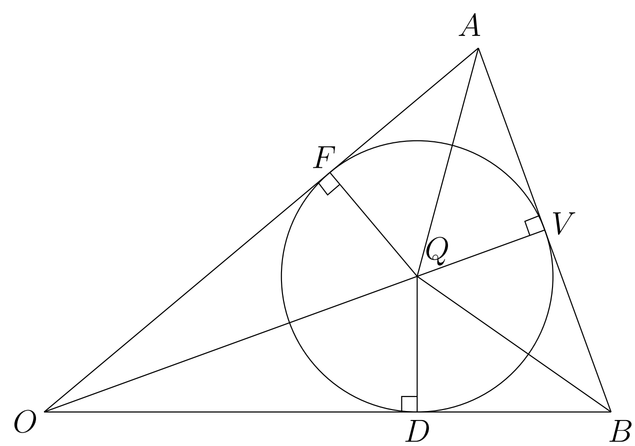
$$\pi R^2(1+\alpha^2+\alpha^4+\alpha^6+\cdots) = \frac{\pi R^2}{1-\alpha^2}.$$

Substituting the expression for α, we have that the total area of the circles is

$$\begin{aligned} \frac{\pi R^2}{1-\alpha^2} &= \frac{\pi R^2}{1-\left(\frac{1-\sin\theta}{1+\sin\theta}\right)^2} \\ &= \frac{\pi R^2(1+\sin\theta)^2}{(1+\sin\theta)^2-(1-\sin\theta)^2} \\ &= \frac{\pi R^2(1+\sin\theta)^2}{4\sin\theta}. \end{aligned}$$

We will return to the expression above later, but first we will find the area of $\triangle AOB$ in terms of R and θ in order to compute the ratio.

Let D, F, and V be the points of tangency of the largest circle to the three sides of $\triangle OAB$ as shown below. Connect the centre of the circle, Q, to O, A, B, D, F, and V. The rest of this page is devoted to proving that OQV is a line. You may wish to skip this part of the argument and come back to it later.



Similar to an observation in part (a), we have that $OF = OD$ because they are equal tangents. Also, $QF = QD = R$ and $\triangle OQF$ and $\triangle OQD$ have common side OQ. By side-side-side congruence, $\triangle OQF$ is congruent to $\triangle OQD$. This means $\angle OQF = \angle OQD$.

By similar arguments, $\angle BQV = \angle BQD$ and $\angle AQV = \angle AQF$.

It is given that $OA = OB$, and since $OF = OD$, we have

$$FA = OA - OF = OB - OD = DB.$$

Again, $QF = QD = R$ and $\angle AFQ = \angle BDQ = 90^\circ$ because they are each made by a tangent and a radius, so we have that $\triangle AFQ$ is congruent to $\triangle BDQ$ by side-angle-side congruence. This means $\angle BQD = \angle AQF$.

Using that $\angle OQF = \angle OQD$ and that $\angle BQV = \angle BQD = \angle AQF = \angle AQV$, we get

$$\begin{aligned}
360^\circ &= \angle OQD + \angle BQD + \angle BQV + \angle AQV + \angle AQF + \angle OQF \\
&= \angle OQD + \angle BQD + \angle BQV + \angle BQV + \angle BQD + \angle OQD \\
&= 2(\angle OQD + \angle BQD + \angle BQV) \\
&= 2\angle OQV
\end{aligned}$$

This means $\angle OQV = 180^\circ$, so OQV is a line segment.

By right-angle trigonometry and since the point Q lies on the angle bisector of $\angle AOB$, $\sin\theta = \dfrac{DQ}{OQ} = \dfrac{R}{OQ}$, so $OQ = \dfrac{R}{\sin\theta}$. Since $QV = R$ as well, we have that

$$OV = R + \frac{R}{\sin\theta}.$$

We also have that $\tan\theta = \dfrac{BV}{OV}$, which means

$$BV = OV\tan\theta = \left(R + \frac{R}{\sin\theta}\right)\tan\theta.$$

Since V is on the angle bisector of $\angle AOB$, we have $\angle AOV = \angle BOV$, so $\triangle AOV$ is congruent to $\triangle BOV$ by side-angle-side congruence, so $AV = BV$, which means $AB = 2BV$. Therefore, the area of $\triangle OAB$ is

$$\begin{aligned}
\frac{1}{2}\times AB \times OV &= \frac{1}{2}\times 2\left(R + \frac{R}{\sin\theta}\right)\tan\theta\left(R + \frac{R}{\sin\theta}\right) \\
&= R^2\tan\theta\left(1 + \frac{1}{\sin\theta}\right)^2
\end{aligned}$$

Recall that the total area of the circles is

$$\frac{\pi R^2(1+\sin\theta)^2}{4\sin\theta},$$

so the fraction of the triangle covered by circles is

$$\begin{aligned}
\frac{\frac{\pi R^2(1+\sin\theta)^2}{4\sin\theta}}{R^2\tan\theta\left(1+\frac{1}{\sin\theta}\right)^2} &= \frac{\pi(1+\sin\theta)^2}{4\sin\theta\tan\theta\left(1+\frac{1}{\sin\theta}\right)^2} \\
&= \frac{\pi(1+\sin\theta)^2}{4\frac{\sin^2\theta}{\cos\theta}\left(1+\frac{1}{\sin\theta}\right)^2} \\
&= \frac{\pi\cos\theta(1+\sin\theta)^2}{4\left[\sin\theta\left(1+\frac{1}{\sin\theta}\right)\right]^2} \\
&= \frac{\pi\cos\theta(1+\sin\theta)^2}{4(1+\sin\theta)^2} \\
&= \frac{\pi}{4}\cos\theta.
\end{aligned}$$

As mentioned in the statement of the problem, this result can be used to produce the answer from part (b). We show how to do this now, noting that this isn't necessarily a better way to solve part (b).

Suppose the side length of the equilateral triangle is x. We computed in part (b) that the area of the triangle is $\frac{\sqrt{3}x^2}{4}$. As well, the area of the largest circle is $\frac{\pi x^2}{12}$.

In part (b), there are three infinite "lines" of circles, each starting with the largest circle and extending toward a vertex of the triangle. By part (c), each of these three lines of circles covers the fraction $\frac{\pi}{4}\cos\theta$ of the area of the triangle, where $\theta = \frac{60^\circ}{2} = 30^\circ$. Therefore, the area of each of the three infinite lines of circles is

$$\begin{aligned}\frac{\pi}{4}\cos 30^\circ \times \frac{\sqrt{3}x^2}{4} &= \frac{\pi}{4} \times \frac{\sqrt{3}}{2} \times \frac{\sqrt{3}x^2}{4} \\ &= \frac{3\pi x^2}{32}.\end{aligned}$$

If we take three times this quantity, we will have computed the total area of all circles in the picture from part (b) but will have counted the area of the largest circle three times rather than once. Therefore, the total area of the circles is

$$\begin{aligned}3 \times \frac{3\pi x^2}{32} - 2 \times \frac{\pi x^2}{12} &= \pi x^2\left(\frac{9}{32} - \frac{1}{6}\right) \\ &= \pi x^2\left(\frac{27}{96} - \frac{16}{96}\right) \\ &= \frac{11\pi x^2}{96}.\end{aligned}$$

Therefore, the fraction of the triangle covered by circles is

$$\frac{\frac{11\pi x^2}{96}}{\frac{\sqrt{3}x^2}{4}} = \frac{11\pi}{24\sqrt{3}}$$

which is indeed the answer from part (b).

Problem of the Month
Problem 2: November 2020

Problem

Suppose n is a positive integer and that you have n pairs of socks. Within each pair of socks, the two socks are the same colour. Every pair has a unique colour. After doing laundry, all of the $2n$ socks are in a laundry basket. You begin to remove socks one at a time (always choosing randomly and never replacing the socks) until you find a pair. That is, you remove socks until you remove a sock that matches a sock that has already been removed.

For positive integers n and k with $k < 2n$, we define $P(n, k)$ to be the probability that the first k socks removed are all different colours and there is a pair among the first $k + 1$ socks that are removed. That is, $P(n, k)$ is the probability that the first pair is found upon removing the $(k+1)^{\text{st}}$ sock.

(a) Compute $P(4, k)$ for each k from 1 through 7. Some of these probabilities should equal 0.

(b) Find a general formula for $P(n, k)$ when $k \leq n$. It might be useful later if you can find a formula that only uses addition, subtraction, multiplication, division, exponentiation, as well as factorials and binomial coefficients. Notice that the question does not ask about $P(n, k)$ for $k > n$. You might want to think about what happens in this case.

(c) For a positive integer i, define T_i to be the sum of the first i positive integers. For example, $T_1 = 1$, $T_2 = 1 + 2 = 3$, and $T_3 = 1 + 2 + 3 = 6$. The numbers T_1, T_2, T_3, T_4, and so on, are often called the *triangular numbers*. You may already know that $T_i = \dfrac{i(i+1)}{2}$ for every positive integer i. If not, think about why!

(i) Suppose $n = T_i$ for some $i > 1$. Show that the largest value of $P(n, k)$ is achieved when $k = i$ and when $k = i + 1$.

(ii) Suppose n is a positive integer that is not a triangular number. This means n is strictly between T_i and T_{i+1} for some i. Show that the largest value of $P(n, k)$ is achieved when

$$k = \left\lfloor \frac{1 + \sqrt{8n+1}}{2} \right\rfloor .$$

(d) For a positive integer n, we call k a *peak* for n if $P(n, k) \geq P(n, \ell)$ for all integers $1 \leq \ell \leq n$. Part (c) suggests that there are two peaks for n when $n > 1$ is a triangular number and that there is a unique peak when n is not a triangular number. Find a positive integer k for which there are exactly 2020 positive integers n with the property that k is a peak for n.

Hint

(a) It might be helpful to write out in words what $P(4,1)$, $P(4,2)$, $P(4,3)$, and so on, actually mean.

(b) There are many ways to approach this. It might be useful to calculate $P(n,1)$, $P(n,2)$, $P(n,3)$, and so on, for some specific values of n (in (a), this was done for $n = 4$). To tidy up your expression, it might also be useful to read about *factorials* and *binomial coefficients*.

(c) Using your formula from (b), consider the quantity $\dfrac{P(n,k+1)}{P(n,k)}$.

(d) Use part (c) and look for a pattern.

Problem of the Month

Solution to Problem 2: November 2020

(a) The quantity $P(4,1)$ is the probability that there is no pair among the first sock removed *and* there is a pair among the first two socks removed. There can never be a pair among one sock, so $P(4,1)$ is simply the probability that the second sock matches the first. After drawing one sock, there will be seven socks remaining and exactly one of them matches the first. Therefore, $P(4,1) = \frac{1}{7}$.

The quantity $P(4,2)$ is the probability that the first two socks are different and the third sock matches one of the first two. The probability that the first two socks are different is the probability that the second sock does not match the first sock, which is $\frac{6}{7}$. If the first two socks drawn are different, then there will be six socks remaining, two of which match one of the first two. Therefore, $P(4,2) = \frac{6}{7} \times \frac{2}{6} = \frac{2}{7}$.

In a similar way, we compute $P(4,3)$ by first computing the probability that the first three socks are different. This is the probability that the first two socks are different times the probability that the third sock is different from both of the first two. From before, the probability that the first two socks are different is $\frac{6}{7}$. Since two socks of the remaining six would match at least one of the first two (different) socks, the probability that the first three socks are different is $\frac{6}{7} \times \frac{4}{6} = \frac{4}{7}$. After three distinct socks are drawn, there are five remaining, and three of them will make a pair. Therefore, $P(4,3) = \frac{4}{7} \times \frac{3}{5} = \frac{12}{35}$.

Suppose the first four socks that were removed are all different. Then the fifth sock must match one of these first four since there are only four pairs in total. Therefore, $P(4,4)$ is equal to the probability that the first four socks are different. Following similar reasoning to that in the previous few cases,

$$P(4,4) = \frac{6}{7} \times \frac{4}{6} \times \frac{2}{5} = \frac{8}{35}.$$

Among any five of the eight socks, there must be a pair since there are only four pairs overall. Therefore, the probability that the first five, six, or seven socks are all different is equal to zero, which means $P(4,5) = P(4,6) = P(4,7) = 0$. Note that this reasoning generalizes. That is, $P(n,k) = 0$ when $n < k < 2n$.

Since a pair must eventually be found, the sum of the probabilities $P(4,1)$ through $P(4,7)$ should be 1. Indeed,

$$\begin{aligned}
&P(4,1) + P(4,2) + P(4,3) + P(4,4) + P(4,5) + P(4,6) + P(4,7) \\
=\ &\frac{1}{7} + \frac{2}{7} + \frac{12}{35} + \frac{8}{35} + 0 + 0 + 0 \\
=\ &\frac{5 + 10 + 12 + 8}{35} \\
=\ &\frac{35}{35} = 1
\end{aligned}$$

(b) We will give two derivations of the same formula using two different kinds of reasoning.

Using similar reasoning to that in the solution to part (a), we first compute the probability that the first $k \leq n$ socks removed from a basket of $2n$ socks are all different. The probability that the first two socks are different is $\frac{2n-2}{2n-1}$ because after the first sock is removed, there are $2n-1$ socks remaining in the basket, one of which matches the first sock, so the other $2n-2$ do not match the first sock.

After two different socks are removed, there are $2n-2$ socks remaining in the basket, two of which will match one of the first two socks. Therefore, $(2n-2)-2 = 2n-4$ of the remaining socks will *not* match either of the first two socks. This means the probability that the first three socks are different is

$$\frac{2n-2}{2n-1} \times \frac{2n-4}{2n-2}.$$

After three distinct socks are removed, there are $2n-3$ socks remaining, three of which will make a pair. Hence, $(2n-3)-3 = 2n-6$ socks can be drawn so that the first four socks are different. Therefore, the probability that the first four socks are different is

$$\frac{2n-2}{2n-1} \times \frac{2n-4}{2n-2} \times \frac{2n-6}{2n-3}.$$

Continuing in this way, the probability that the first k socks are different is

$$\begin{aligned}&\frac{2n-2}{2n-1} \times \frac{2n-4}{2n-2} \times \frac{2n-6}{2n-3} \times \cdots \times \frac{2n-2(k-2)}{2n-(k-2)} \times \frac{2n-2(k-1)}{2n-(k-1)}\\ =&\frac{(2n-2)(2n-4)(2n-6)\cdots(2n-2(k-2))(2n-2(k-1))}{(2n-1)(2n-2)(2n-3)\cdots(2n-(k-2))(2n-(k-1))}\end{aligned}$$

and after factoring a 2 out of each term in the numerator, this is equal to

$$\frac{2^{k-1}(n-1)(n-2)(n-3)\cdots(n-k+2)(n-k+1)}{(2n-1)(2n-2)(2n-3)\cdots(2n-k+2)(2n-k+1)}.$$

It may not be obvious how to interpret the formula above when $k=1$, or even when $k=2$. You may want to think about this now, but some explanation is given after the simplification that follows. In order to tidy up this expression, we will multiply by 1 in the forms $\frac{(n-k)!}{(n-k)!}$, $\frac{(2n-k)!}{(2n-k)!}$, and $\frac{2n}{2n}$. This will allow us to "complete" some factorials:

$$\begin{aligned}&\frac{2^{k-1}(n-1)(n-2)(n-3)\cdots(n-k+2)(n-k+1)}{(2n-1)(2n-2)(2n-3)\cdots(2n-k+2)(2n-k+1)}\\ =&\frac{2^{k-1}(n-1)(n-2)(n-3)\cdots(n-k+2)(n-k+1)}{(2n-1)(2n-2)(2n-3)\cdots(2n-k+2)(2n-k+1)} \times \frac{(n-k)!}{(n-k)!} \times \frac{(2n-k)!}{(2n-k)!} \times \frac{2n}{2n}\\ =&\frac{2^k n!(2n-k)!}{(2n)!(n-k)!}.\end{aligned}$$

Before moving on, note that when $k=1$, the expression above simplifies to

$$\frac{2^1 n!(2n-1)!}{(2n)!(n-1)!} = \frac{2n(2n-1)!}{(2n)!} = \frac{(2n)!}{(2n)!} = 1$$

and when $k = 2$ it simplifies to

$$\frac{2^2 n!(2n-2)!}{(2n)!(n-2)!} = \frac{4(n)(n-1)}{2n(2n-1)} = \frac{2n-2}{2n-1}.$$

For $k = 2$, this agrees with the probability computed earlier. For $k = 1$, there can never be a pair among one sock, so the probability that there is no pair among the first sock should indeed be 1.

After the first k different socks are drawn, there are $2n - k$ socks remaining and k of them can be drawn to make a pair with one of the first k. Therefore, we have

$$P(n,k) = \frac{2^k n!(2n-k)!}{(2n)!(n-k)!} \times \frac{k}{2n-k} = \frac{k2^k n!(2n-k-1)!}{(2n)!(n-k)!}.$$

In the derivation above, we directly computed the probability that the first pair was found when the $(k+1)^{\text{th}}$ sock was drawn. In the second derivation, the approach will be to count the total number of ways in which all $2n$ socks can be drawn (without stopping at a pair), then count the number of ways in which all $2n$ socks can be drawn so that the first pair occurs when the $(k+1)^{\text{th}}$ sock is drawn. The probability $P(n, k)$ is the ratio of the result of these two counts.

The second derivation of the formula assumes an understanding of binomial coefficients.

For the total number of ways to draw the socks, first observe that there are $(2n)!$ ways to order the $2n$ socks. Since the two socks in each pair are indistinguishable, this over counts by a factor of 2 for each of the n pairs. Therefore, the number of ways to draw the $2n$ socks is $\frac{(2n)!}{2^n}$.

We now count the number of ways that the socks can be drawn so that the first pair occurs when the $(k+1)^{\text{st}}$ sock is drawn. There are n choices for the colour of the first pair, and k choices for where the first of these socks was drawn (the second sock in this pair is the $(k+1)^{\text{st}}$ sock drawn). The other $k - 1$ socks drawn among the first k must be different from eachother and different from the first pair. Thus, there are $\binom{n-1}{k-1}$ possible choices for the colours of the other $k-1$ socks, and $(k-1)!$ orders in which they can be drawn. Thus, if the first pair is made when the $(k+1)^{\text{st}}$ sock is drawn, then there are

$$n \times k \times \binom{n-1}{k-1} \times (k-1)!$$

ways in which the first $k + 1$ socks can be drawn.

Of the remaining $2n-k-1$ socks to be drawn, there are $n-k$ pairs. Similar to the earlier count, this means there are $\frac{(2n-k-1)!}{2^{n-k}}$ ways to draw the remaining socks. Therefore, there are

$$\begin{aligned} n \times k \times \binom{n-1}{k-1} \times (k-1)! \times \frac{(2n-k-1)!}{2^{n-k}} &= \frac{nk(n-1)!(k-1)!(2n-k-1)!}{(k-1)!(n-k)!2^{n-k}} \\ &= \frac{k2^{k-n} n!(2n-k-1)!}{(n-k)!} \end{aligned}$$

ways to draw the $2n$ socks so that the first pair occurs when the $(k+1)^{\text{th}}$ sock is drawn. Dividing this by $\frac{(2n)!}{2^n}$ and simplifying, we have

$$P(n,k) = \frac{k2^k n!(2n-k-1)!}{(2n)!(n-k)!}.$$

(c) Before solving either (i) or (ii), we will analyze the quantity $\dfrac{P(n,k+1)}{P(n,k)}$, being careful to assume that $k \le n$ so that the denominator is not 0. Also note that in both parts we have that $n > T_1 = 1$. Notice that for any integer $m > 1$, we have that $\dfrac{m!}{(m-1)!} = m$ and $\dfrac{(m-1)!}{m!} = \dfrac{1}{m}$. Using the formula from (b) for $P(n,k+1)$ and $P(n,k)$, we have that

$$\begin{aligned}\frac{P(n,k+1)}{P(n,k)} &= \frac{(k+1)2^{k+1}n!(2n-k-2)!}{(2n)!(n-k-1)!} \times \frac{(2n)!(n-k)!}{k2^k n!(2n-k-1)!} \\ &= \frac{2(k+1)}{k} \times \frac{n-k}{2n-k-1}\end{aligned}$$

Since we are interested in comparing the sizes of $P(n,k+1)$ and $P(n,k)$, it will be useful to determine when their ratio is greater than, less than, or equal to 1. To do this, we will expand the numerator and denominator and rearrange the ratio to take the form $1+x$ and examine when the quantity x is greater than, less than, or equal to 0.

$$\begin{aligned}\frac{P(n,k+1)}{P(n,k)} &= \frac{2kn-2k^2+2n-2k}{2kn-k^2-k} \\ &= \frac{2kn-k^2-k+(2n-k^2-k)}{2kn-k^2-k} \\ &= \frac{2kn-k^2-k}{2kn-k^2-k} + \frac{2n-k^2-k}{2kn-k^2-k} \\ &= 1 + \frac{2n-k^2-k}{2kn-k^2-k}.\end{aligned}$$

Recall that $1 \le k \le n$, and since $n > 1$, we have that $k+1 < 2n$, so $2n-k-1 > 0$ and hence $2nk-k^2-k > 0$. This means the denominator in the above expression is positive, so the sign of $\dfrac{2n-k^2-k}{2kn-k^2-k}$ (and whether or not it is 0) is completely determined by its numerator. Three pieces of information can now be extracted from the expression above, still subject to the restrictions $n > 1$ and $1 \le k \le n$:

- $P(n,k+1) < P(n,k)$ if and only if $2n-k^2-k < 0$.
- $P(n,k+1) > P(n,k)$ if and only if $2n-k^2-k > 0$.
- $P(n,k+1) = P(n,k)$ if and only if $2n-k^2-k = 0$.

Notice that the equation $2n-k^2-k=0$ is equivalent to $n = \dfrac{k(k+1)}{2}$, which means $P(n,k+1) = P(n,k)$ if and only if n is the k^{th} triangular number. Finally, we examine what happens when $k=n$. In this situation, $2n-k^2-k=0$ is the same as $n-n^2=0$ which implies $n=0$ or $n=1$. Similarly, the inequality $2n-k^2-k>0$ is the same as $n-n^2>0$, which means n is strictly between 0 and 1. We are assuming that $n>1$, so neither of these situations can occur. Therefore, if $k=n$, we must have $2n-k^2-k<0$ which makes sense since $P(n,n+1)=0$ but $P(n,n)>0$.

(i) Suppose $n = T_i = \dfrac{i(i+1)}{2}$, the i^{th} triangular number. From above, we have that $P(n,i) = P(n,i+1)$. Furthermore, since the list $T_1, T_2, T_3, \ldots$ of triangular numbers is increasing (each is obtained from the previous by adding a positive number), there is no positive integer $j \ne i$ for which $n = T_j$ as well. Therefore, $P(n,k+1) = P(n,k)$ if and only if $k = i$.

Notice that the quantity $2n - k^2 - k$ is decreasing as k increases. This means it must be positive until $k = i$, at which point it equals 0, and after which it must be negative. Using the earlier discussion, this means the list

$$P(n,1), P(n,2), \ldots, P(n,i), P(n,i+1), \ldots, P(n,n)$$

is increasing until $P(n,i)$ and decreasing from $P(n,i+1)$. We have observed that $P(n,i) = P(n,i+1)$, so $P(n,k)$ is largest when k takes the two values $k = i$ and $k = i+1$.

(ii) We now assume for some integer i that $T_i < n < T_{i+1}$. Since n is not a triangular number, we know that there is no $k \leq n$ for which $P(n,k) = P(n,k+1)$. Rearranging the first two conditions in the bulleted list above, we have that $P(n,k) < P(n,k+1)$ for any k satisfying $\dfrac{k(k+1)}{2} < n$ and $P(n,k) > P(n,k+1)$ for any k satisfying $\dfrac{k(k+1)}{2} > n$. This means the probabilities $P(n,k)$ strictly increase while k satisfies $\dfrac{k(k+1)}{2} < n$ and decrease thereafter. By our assumption, i is the largest integer with the property that $n > \dfrac{i(i+1)}{2} = T_i$. This means that among all positive integers $k \leq n$, $P(n,k)$ is largest when $k = i+1$.

To answer the question, we need to show that

$$i + 1 = \left\lfloor \frac{1 + \sqrt{1+8n}}{2} \right\rfloor.$$

We are assuming that $T_i < n < T_{i+1}$, which means

$$\frac{i(i+1)}{2} < n < \frac{(i+1)(i+2)}{2}.$$

The left inequality rearranges to $i^2 + i - 2n < 0$. The roots of $i^2 + i - 2n$ are

$$i = \frac{-1 \pm \sqrt{1+8n}}{2}.$$

and so for the inequality $i^2 + i - 2n < 0$ to be satisfied, we need to have i less than the larger of these roots. We conclude that

$$i < \frac{-1 + \sqrt{1+8n}}{2}.$$

The other inequality rearranges to $0 < i^2 + 3i + 2 - 2n$. The polynomial on the right has roots

$$i = \frac{-3 \pm \sqrt{3^2 - 4(2-2n)}}{2} = \frac{-3 \pm \sqrt{1+8n}}{2}.$$

For the inequality $0 < i^2 + 3i + 2 - 2n$ to be satisfied, i must be either smaller than the smaller root, or larger than the larger root. The smaller of these two roots is negative, so since we also require that $i > 0$, we have that

$$\frac{-3 + \sqrt{1+8n}}{2} < i.$$

Using that this quantity is exactly one less than $\dfrac{-1+\sqrt{1+8n}}{2}$, we now have that the integer i satisfies

$$\frac{-1+\sqrt{1+8n}}{2} - 1 < i < \frac{-1+\sqrt{1+8n}}{2}.$$

Thus, we have that the integer i is between two quantities differing by exactly 1. As long as the bounding quantities are not integers, this means i must be the largest integer that is less than or equal to $\dfrac{-1+\sqrt{1+8n}}{2}$. In other words, as long as $\dfrac{-1+\sqrt{1+8n}}{2}$ is not an integer, we will have

$$i = \left\lfloor \frac{-1+\sqrt{1+8n}}{2} \right\rfloor.$$

To finish the proof, suppose there is an integer m such that $m = \dfrac{-1+\sqrt{1+8n}}{2}$. Rearranging leads to $(2m+1)^2 = 1+8n$ or $4m^2+4m+1 = 1+8n$. Solving for n gives

$$n = \frac{4m^2+4m}{8} = \frac{m^2+m}{2} = \frac{m(m+1)}{2}$$

which would mean that n is a triangular number. We are assuming this is not the case, so $\dfrac{-1+\sqrt{1+8n}}{2}$ is not an integer. Therefore,

$$i+1 = \left\lfloor \frac{-1+\sqrt{1+8n}}{2} \right\rfloor + 1 = \left\lfloor \frac{1+\sqrt{1+8n}}{2} \right\rfloor$$

(d) We will begin by computing the peaks for a few small positive integers. To add to the work from part (c), we note that $P(1,1) = 1$, and so $n = 1$ has a unique peak of 1. We will use that when $n = T_i$ for some integer $i > 1$, there are two peaks for n and they occur at $k = i$ and $k = i+1$, as well as the fact that when n is not a triangular number, there is a unique peak for n and it occurs at

$$k = \left\lfloor \frac{1+\sqrt{1+8n}}{2} \right\rfloor.$$

Keeping in mind that the first few triangular numbers are 1, 3, 6, 10, 15, and 21, we have

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

You may notice that a rather tidy pattern has started to emerge. It appears that with the exceptions of $k = 1$ and $k = 2$, the integer k occurs as a peak $k+1$ times, and it occurs as a peak for the numbers between T_{k-1} and T_k inclusive. Therefore, we expect $k = 2019$ to occur 2020 times as a peak.

First, suppose $k = 2019$ occurs as a peak for some triangular number n. From earlier work, this means $n = 2\,039\,190 = T_{2019}$ or $n = 2\,037\,171 = T_{2018}$. This gives two integers n for which 2019 is a peak. Otherwise, for 2019 to be a peak for n, we must have

$$2019 = \left\lfloor \frac{1+\sqrt{1+8n}}{2} \right\rfloor$$

which means

$$2019 \leq \frac{1+\sqrt{1+8n}}{2} < 2020.$$

This can be rearranged to get $4037 \leq \sqrt{1+8n} < 4039$ which can be further rearranged to get $2\,037\,171 \leq n < 2\,039\,190$. Combining this with the other two numbers for which 2019 is a peak, the integers n for which 2019 is a peak are exactly those that satisfy

$$2\,037\,171 \leq n \leq 2\,039\,190.$$

This is a total of $2\,039\,190 - 2\,037\,170 = 2020$ integers.

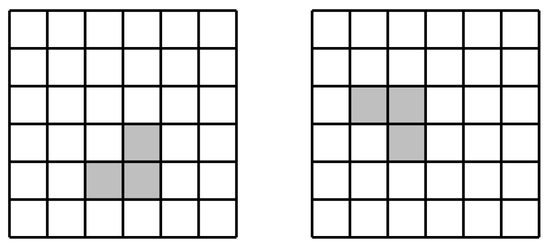
Problem of the Month

Problem 3: December 2020

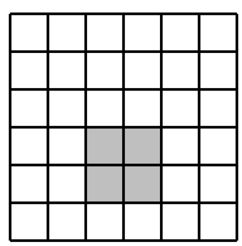
Problem

You and a friend are playing games involving a 6×6 grid with a coin in each cell. In each game, your friend arranges the coins so that each coin shows either a head or a tail. An arrangement of coins is called *winnable* if it is possible to perform a sequence of legal moves that results in all 36 coins showing a head. Each game has a different set of legal moves.

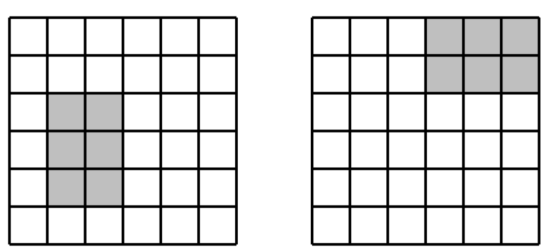
(a) In the first game, a legal move consists of flipping over exactly three of the four coins in a 2×2 subgrid. Each grid below has three cells highlighted. In each of these two grids, flipping over the coins in the highlighted cells is an example of a legal move.



(b) In the second game, a legal move consists of flipping over all four of the coins in a 2×2 subgrid. In the grid below, flipping over the four coins in the highlighted cells is an example of a legal move.



(c) In the third game, a legal move consists of flipping over all 6 of the coins in a 3×2 or 2×3 subgrid. In each of the two grids below, flipping over the coins in the highlighted cells is an example of a legal move.



For each of the three games, determine how many of the 2^{36} arrangements are winnable. In all three games, subgrids must be "connected". For example, the four corners of the 6×6 grid is *not* a 2×2 subgrid.

Hint

In all three parts, it is a good idea to try some examples. It is also useful to note that the order in which moves are performed does not affect the outcome of the moves. As well, performing a move twice has no overall effect on the coins.

(a) Consider an arrangement of coins that has 1 coin showing a tail and the rest showing heads. Is this arrangement winnable?

(b) What happens to the *parity* of the number of tails in each row and column when a legal move is performed?

(c) Are there any legal moves that you do not need? That is, are there any moves with the property that if an arrangement is winnable, then it is winnable without using that move?

Problem of the Month

Solution to Problem 3: December 2020

(a) In the first game, every arrangement is winnable. We will show this by describing how to win any arrangement.

Observe that every cell is a member of either one, two, or four 2×2 subgrids depending on if the cell is on an edge, in a corner, or in the "interior" of the grid. What is important for this argument is that every cell is in at least one 2×2 subgrid.

Suppose a coin is showing a tail and that it is the top-left coin in some 2×2 subgrid:

T	?
?	?

The question marks indicate that the coin in that cell could be showing either a head or a tail.

It is possible to change the indicated tail to a head without changing the other three coins. This can be done by performing the three moves below where the cells in which coins will be flipped are marked by an X.

X	
X	X

X	X
X	

X	X
	X

After performing these three legal moves, the coin in the top left corner has been flipped three times, so it will now be showing a head. The other three coins in the 2×2 subgrid were flipped twice each, so they will each be showing what they were showing before the moves were performed. No other coins in the grid were flipped, so this sequence of three legal moves has the effect of changing one coin from showing a tail to showing a head and does not change what any other coins are showing.

A similar sequence of three moves can be used to change a tail to a head if it is in one of the other corners of a 2×2 subgrid. Thus, it is possible to change any one tail to a head without changing any other coins, so every arrangement can be won by changing the tails to heads one at a time. We note that the technique for winning described above may not win in the smallest possible number of moves.

(b) A simple yet useful observation for this game is that the *parity* of the number of tails in any given row or column will not change as a result of a legal move.

Each legal move affects two rows and two columns and flips two coins in each of these two rows and two columns. For the cells of a row affected by a legal move, one of the following four situations occurs:

H	H	$\longrightarrow$	T	T		T	T	$\longrightarrow$	H	H
H	T	$\longrightarrow$	T	H		T	H	$\longrightarrow$	H	T

In the two situations illustrated on the top, the number of tails in that row will either decrease by two or increase by two. In the two situations illustrated on the bottom, the number of tails in the affected row does not change. A similar argument shows that the number of tails in a column either increases by two, decreases by two, or stays the same after any legal move.

This means that if some row or column has an odd number of tails in the initial arrangement, then it will still have an odd number of tails after any number of legal moves. For an arrangement to be winnable, it must be possible to perform a sequence of legal moves resulting in the number of coins showing tails in each row and column to be 0. Since 0 is even, this means a winnable arrangement must have an even number of tails in each row and each column. We will call an arrangement of the coins "good" if there is an even number of tails in each row and an even number of tails in each column.

We just argued that if an arrangement is winnable, then it is good. We will next show that if an arrangement is good, then it is winnable. This will show that the winnable arrangements are exactly the good arrangements, so we will be able to count the winnable arrangements by counting the good arrangements.

Suppose an arrangement of the coins is good. A fundamental fact, which we argued above, is that a legal move transforms a good arrangement into an other good arrangement.

One strategy of winning is to systematically change all coins to show heads, row by row, starting with the top row. If there is a tail in the top row, locate the leftmost tail in the top row. Since the arrangement is good, the number of tails in the top row is even, so there must be at least one tail in the top row to the right of this tail. That is, the leftmost tail in the top row is in one of the first five columns. In the example illustrated below, it occurs in the third column. This means it is a legal move to flip the four coins with this tail in the top-left corner. Doing so will change the leftmost tail to a head.

H	H	T	?	?	?		H	H	H	?	?	?
?	?	?	?	?	?		?	?	?	?	?	?
?	?	?	?	?	?	$\longrightarrow$	?	?	?	?	?	?
?	?	?	?	?	?		?	?	?	?	?	?
?	?	?	?	?	?		?	?	?	?	?	?
?	?	?	?	?	?		?	?	?	?	?	?

This move may change some heads to tails in the second row and may even change a head to a tail in the top row (if there was a head immediately to the right of the leftmost tail). However, as mentioned above, the new arrangement is still good, and the leftmost tail in the top row must now be further to the right.

After no more than four legal moves, the top row will either contain no tails or the leftmost tail will be in the fifth column. In the latter situation, the rightmost coin in the top row must also be showing a tail because the number of tails in the top row is even. Flipping the four coins in the top right corner of the grid will now make every coin in the top row show a head.

Once the top row has no tails in it, this procedure can be applied to the second row. Once the second row contains to tails, it can be applied to the third row, then the fourth, and finally the fifth. This means that a sequence of legal moves can be performed to leave the coins arranged so that every coin in the first five rows is showing a head.

However, this procedure cannot be applied to the bottom row since there is no "room". Flipping a coin in the bottom row necessarily flips coins in the fifth row, which has already been changed to show only heads. However, it turns out that since the arrangement was good, changing the coins in the first five rows to all show heads will force the coins in the sixth row to all show heads as well, meaning the game has already been won.

To see that this is true, suppose a good arrangement has no coins showing tails in the first five rows. In any column, if the coin in the bottom row showed a tail, then there would be an odd number of tails in that column. The arrangement is good, so this is not possible. Therefore, the coins in the bottom row must also be showing heads.

We have shown that every good arrangement is winnable, which shows that the good arrangements are exactly the winnable arrangements. To answer the question, we will count the good arranagements.

To do this, denote by x the number of subsets of a set of 6 objects that have an even number of elements. With this notation, in any given row (or column), there are x ways to arrange the coins so that an even number of them are showing tails. We will show that the number of good arrangements is x^5 and compute the exact value of x later.

Notice that there are x^5 ways to arrange the coins in the first 5 rows so that there is an even number of tails in each of these rows. We will now prove the following claim: Every arrangement of an even number of tails in the first 5 rows can be "extended" in a unique way to a good arrangement.

To prove the claim, suppose the coins in the first 5 rows are arranged so that they each contain an even number of tails. To extend this to a good arrangement, we must decide how to arrange the coins in the bottom row. In order to have an even number of tails in each column, there is no choice to make: if there are an even number of tails in the first 5 cells of a column, then the final cell in that column must show a head. If there are an odd number of tails in the first 5 cells of a column, then the final coin must show a tail. Thus, if the arrangement of the coins in the first 5 rows can be extended to a good arrangement, then there is only one way to do it. All that remains is to show that the way of arranging the coins in the bottom row described above will produce an even number of tails in the bottom row.

This can be seen by observing that since there are an even number of tails in each column, there must be an even number of tails *in total*. Since the number of tails in the first 5 rows is even, it must be the case that the number of tails in the last row is even as well. Otherwise, the total number of tails in the grid would be odd.

You may wish to pause to dwell on the logic, but we have now shown that the number of good arrangements is equal to the number of ways to arrange the coins so that there are an even number of tails in each of the first five rows. As mentioned above, this leads to the number of good arrangements being equal to x^5. It remains to compute x.

Suppose X is a set of 6 objects. We will count the number of subsets that have an even number of elements. That is, we will count the number of subsets of X that contain exactly 0, 2, 4, or 6 elements.

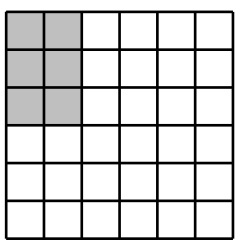
The empty set is the only subset of X that contains 0 elements. Likewise, the set X itself is the only subset of X that contains 6 elements. There are $\dfrac{6 \times 5}{2} = 15$ subsets with 2 elements. This is because there are 6 ways to choose one element, 5 ways to choose a second, and this counts each subset exactly twice. "Choosing" 2 elements is the same as "ignoring" 4 elements, so the number of subsets with 4 elements is also equal to 15. Therefore, we have that $x = 1 + 15 + 15 + 1 = 32$.

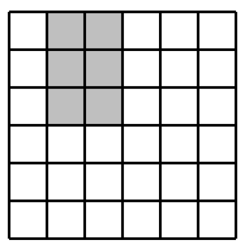
In fact, it is not a coincidence that $x = 2^5$ and the total number of subsets of X is 2^6. Indeed, if X is a set of n elements for any any positive integer n, then there are 2^n subsets of X in total, exactly 2^{n-1} of which contain an even number of elements. Put another way, exactly half of the subsets of a set have an even number of elements. You might want to try to prove this in general.

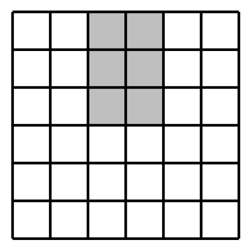
We can now answer the question: There are $(2^5)^5 = 2^{25}$ winnable arrangements.

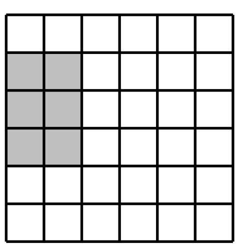
(c) Throughout this solution, a sequence of legal moves will be called a "winning sequence" if it causes all coins in a given arrangement to show heads.

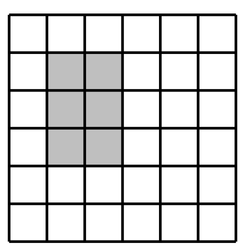
There are a total of 40 legal moves, but we will give names to 9 of them, pictured below.

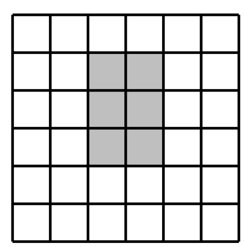

Move 1

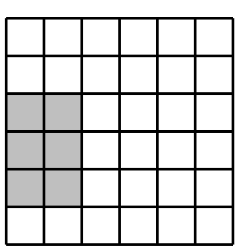

Move 2

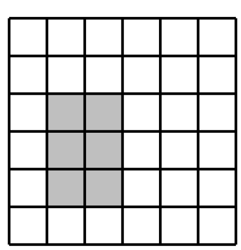

Move 3

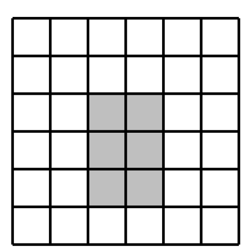

Move 4


Move 5


Move 6

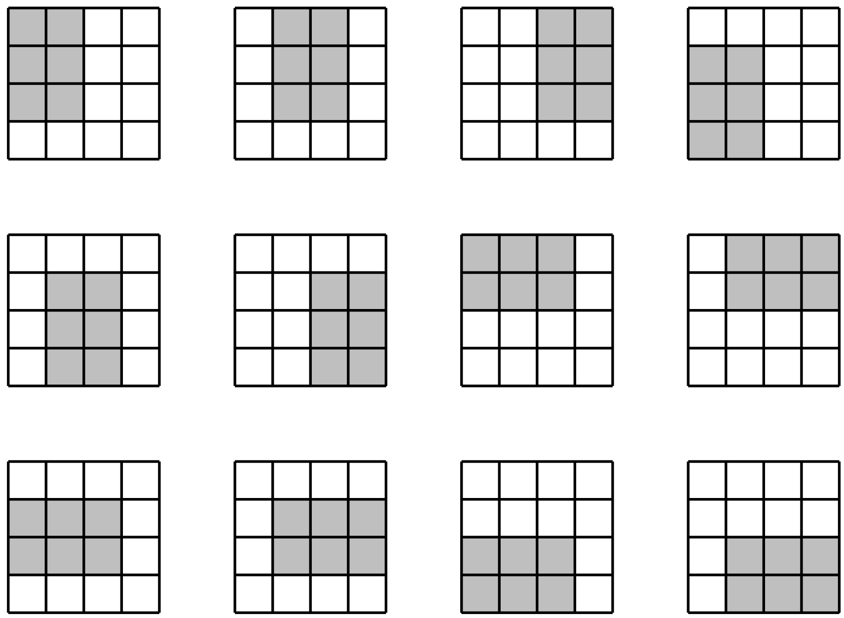

Move 7


Move 8


Move 9

In the first part of this solution, we will argue that any winnable arrangement has a winning sequence that does not use any of Moves 1 through 9. This will allow us to show that there are at most $2^{40-9} = 2^{31}$ winnable arrangements. After that, we will demonstrate that there are at least 2^{31} winnable arrangements. These two facts together will imply that there are exactly 2^{31} winnable arrangements.

Consider any 4×4 subgrid. There are 12 moves that flip only coins in this subgrid:



Suppose we perform each of the moves above once, except the first one. The grid below indicates how many times each coin will be flipped. The highlighted cells are those in which the coin is flipped an odd number of times.

1	3	4	2
3	7	8	4
3	7	8	4
2	4	4	2

If a coin is flipped an odd number of times, it has changed from showing a head to showing a tail, or vice versa. If a coin is flipped an even number of times, there will be no change in what it shows. This means the effect of the first move can be “simulated” by the other eleven. That is, if a move flips coins in the top-left 3×2 subgrid of a 4×4 subgrid, its effect can be achieved using the other eleven moves that flip coins only in that 4×4 subgrid. This is what will allow us to “eliminate” Moves 1 through 9.

Suppose we have a winnable arrangement and a winning sequence. By replacing each (if any) occurrence of Move 1 by the eleven other moves described above, we get a new winning sequence that does not use Move 1. This new sequence will be longer than the original and will have new occurrences of Moves 2, 3, 4, 5, and 6. The key is that if an arrangement is winnable, then there is a winning sequence that does not use Move 1.

Next, consider Move 2. The eleven moves that can be used to replace an occurrence of Move 2 include Moves 3, 5, and 6 (among others), but do not include Move 1. We can now take our winning sequence (that does not include Move 1) and replace every occurrence of Move 2 by eleven other moves. Again, the sequence will get longer and there will be new occurrences of Moves 3, 5, and 6, but Move 1 will not be reintroduced. Thus, if an arrangement is winnable, then there is a winning sequence that does not use Move 1 and does not use Move 2.

Next, each occurrence of Move 3 can be replaced by eleven other moves. Considering the diagrams above carefully, this will not reintroduce Move 1 or Move 2 to the sequence. We can then continue this way to see that if there is a winning sequence, then there is a winning sequence that does not use any of Moves 1 through 9. The new sequence will have more moves, but the key is that it uses at most $40 - 9 = 31$ distinct moves. [It is not directly important for this argument, but removing nine moves is in fact the "best" that we can do. By the end of the solution, you might have a better idea of what this means and why it is true.]

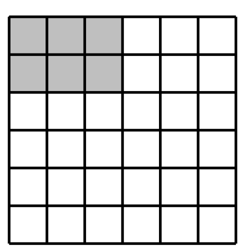
We can further refine the winning sequence by removing repetitions. As noted earlier and mentioned in the hint, the order in which moves are applied does not matter. This is because the overall effect of a sequence of moves on an individual coin is only influenced by how many times that coin is flipped, not by when it is flipped. Thus, to know what effect a sequence has on a given coin, one needs only to count how many times it is flipped. Performing any individual move twice contributes an even number of flips of each coin (either 0 or 2), which means that we can get a shorter winning sequence by eliminating "pairs" of the same move.

Therefore, if an arrangement is winnable, then there is a winning sequence that uses at most 31 moves and does not use any move more than once. Since the order does not matter, there are essentially only 2^{31} winning sequences: For each of the 31 moves, we either use it or do not use it. Since every winning arrangement can be won in one of 2^{31} ways, there cannot be more than 2^{31} winnable arrangements.

We will show that there are at least 2^{31} winnable arrangements by building as many. This will be done using reasoning similar to that in part (b). It should be pointed out that in this part of the argument, we are only concerned with whether or not an arrangement is winnable, so we will not worry about whether or not we are using Moves 1 through 9.

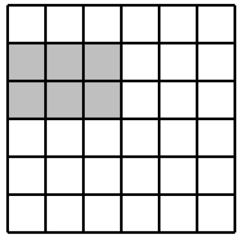
Suppose three consecutive coins in a row are flipped. If the number of tails in that row was even, it will now be odd, and if the number of tails in that row was odd, it will now be even. This can be seen using a similar argument to the one at the beginning of the solution to part (b).

Following similar reasoning to part (b), we can change the coins in the grid to heads row by row. Starting with the first row, if there are an odd number of tails, apply the move indicated below:



This will flip three coins in the first row making the number of tails in the first row even. By using the five moves that flip exactly two coins in the top row (the "3×2" moves), the top row can be changed so that it shows all heads using the same reasoning as in part (b). This will potentially change some heads to tails *below* the top row, but the important thing is that the top row will contain all heads.

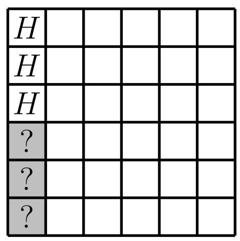
Next, the same idea can be used to change all coins in the second row to show heads. If there are an odd number of tails in this row, apply the move



This will change the coins so that there are an even number of tails in the second row. Using the five moves that flip exactly two coins in the second row and no coins in the top row, the second row can be changed so that all coins show a head. In a similar way, the coins can be changed so that there are no tails in the top three rows and all remaining tails are in the bottom three rows.

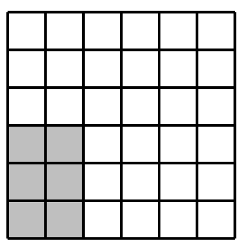
We will now argue that it is possible to perform a sequence of moves to get all remaining tails in the 3×2 subgrid in the bottom-right corner of the array.

First, focus on the three cells highlighted below:



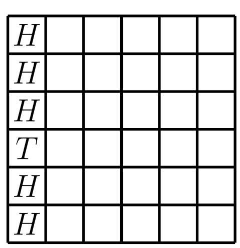
If the coins in these cells all show heads, then there is nothing to do. Otherwise, there are seven possibilities.

If all three of them show a tail, then the move

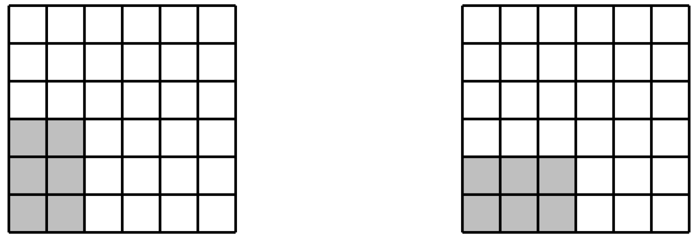


will not change any coins in the top three rows, and will make all coins in the first (leftmost) column show heads.

For each of the other six possibilities, it is possible to change the coins in the first column to all show heads without changing any coins in the first three rows. For example, if the coins in the first column are arranged



then the two moves below will achieve this:



The other five situations can be dealt with similarly, in each case using at most three moves.

This argument can be repeated to ensure that the first four columns contain all heads. Therefore, it is possible, no matter how the coins were arranged, to use a sequence of legal moves to convert the arrangement to

H	H	H	H	H	H
H	H	H	H	H	H
H	H	H	H	H	H
H	H	H	H	?	?
H	H	H	H	?	?
H	H	H	H	?	?

so that all remaining tails (if any) are in the cells marked with question marks.

We will now explain how to construct 2^{31} winnable arrangements. For brevity, we will refer to the coins in the 30 cells outside those marked with question marks above as the "first 30" and the other 6 coins as the "last 6". We will show that for every arrangement of the first 30 coins, there are at least two arrangements of the last 6 coins that make the overall arrangement winnable.

Consider an arbitrary arrangement of the first 30 coins and arrange the last 6 coins to all show heads. We do not claim that this arrangement is winnable, but it will help us to find winnable arrangements.

In the way described earlier, perform a sequence of legal moves so that all remaining tails (if any) are among the last 6 coins. For example, perhaps the last 6 coins are arranged as indicated below:

H	H
T	T
H	H

The overall effect of this sequence of legal moves is to do two things: change all of the first 30 coins to show heads, and flip the coins in the last 6 that are now showing tails. In the example above, it flips the two coins in the middle. To construct a winnable arrangement, arrange the first 30 coins in the same way as before but arrange the last 6 exactly as they are shown above. The sequence of legal moves will still make the first 30 coins show heads, but since we know the sequence flips the "middle two" of the last 6 coins, the sequence will actually change the whole grid to show heads.

In general, we can generate a winnable arrangement by following these steps.

(i) Arrange the first 30 coins arbitrarily and the last 6 to show heads.

(ii) Perform a sequence of legal moves so that all tails (if any) are among the last 6 coins.

(iii) A winnable arrangement can now be found by arranging the first 30 coins in the same way as in (i), but arranging the last 6 in the way they appeared *after* (ii).

The sequence of moves in (ii) will change the arrangement in (iii) to show all heads. This is because in (ii) we learn the overall effect of the moves on the last 6 coins, so in (iii) we can set them up so that they will be flipped to show heads. Thus, we get at least one winnable arrangement corresponding to the given arrangement of the first 30 coins.

To get another, in step (iii) instead arrange the last 6 coins in a way opposite to how they appeared after step (ii). This is a different arrangement because all of the last 6 coins will be different from the other arrangement. Performing the sequence of legal moves will now result in the last 6 coins all showing tails. The arrangement can then be changed to show all heads by flipping all of the last 6 coins, which is a legal move.

Therefore, there are at least $2 \times 2^{30} = 2^{31}$ winnable arrangements. In fact, this also explains how to win. That is, perform a sequence of legal moves to force all tails to the last 6 coins. If the arrangement was winnable, then these coins will either all show tails or all show heads. At most one more move will convert all coins to heads. If any other arrangement appears in the last 6 coins, then the arrangement was not winnable in the first place.

It is interesting to note that things actually do not get much more complicated if the games are played on an $n \times n$ grid with $n > 6$. In part (a), the same argument shows that every arrangement is winnable. In part (b), if the game were played on an $n \times n$ grid, then $2^{(n-1)\times(n-1)}$ of the $2^{n\times n}$ arrangements are winnable. Notice that this means that if the coins are arranged randomly, then the probability that the arrangement is winnable is

$$\frac{2^{n^2-2n+1}}{2^{n^2}} = \frac{1}{2^{2n-1}}$$

which gets very small as n gets large. Thus, if your friend arranges the coins randomly in the game in (b), there is a very small chance that the arrangement is winnable.

On the other hand, extending the reasoning in part (c) shows that there are 2^{n^2-5} winnable arrangements, so the probability that a random arrangement is winnable in the game in (c) is

$$\frac{2^{n^2-5}}{2^{n^2}} = \frac{1}{2^5} = \frac{1}{32}$$

which does not depend on n.

Problem of the Month
Problem 4: January 2021

Problem

In this problem, we will explore when a quadratic polynomial of the form $x^2 + ux + v$ can be decomposed as the sum of the squares of two other polynomials. Keep in mind that a constant function is a polynomial. All polynomials in the problem statements below are assumed to have real coefficients, though they may not have real roots.

(a) Find at least three pairs $\big(p(x), q(x)\big)$ of polynomials such that $(p(x))^2 + (q(x))^2 = x^2 + 2x + 2$.

(b) Suppose $f(x) = x^2 + ux + v$ has the property that $f(x) \geq 0$ for all real numbers x. Prove that there are polynomials $p(x)$ and $q(x)$ such that $x^2 + ux + v = (p(x))^2 + (q(x))^2$.

In the remaining parts of this problem, we will say that the pair of polynomials $\big(p(x), q(x)\big)$ is *special* for the polynomial $x^2 + ux + v$ if

- the coefficients of $p(x)$ and $q(x)$ are all rational, and
- $x^2 + ux + v = (p(x))^2 + (q(x))^2$.

(c) Prove that there are no special pairs for $x^2 + x + 1$.

(d) Prove that if there is a special pair for $x^2 + ux + v$, then u and v are both rational and there is a rational number r such that $4v - u^2 = r^2$.

(e) Prove that if there is a special pair for $x^2 + ux + v$, then there are infinitely many special pairs for $x^2 + ux + v$.

Hint

(a) Expand the expression on the left side of $(ax+b)^2+(cx+d)^2 = x^2+2x+2$ and look at the resulting coefficients.

(b) What if $q(x)$ is constant?

(c) If real numbers s and t satisfy $s^2+t^2=1$, then there is a real number θ with the property that $s=\cos\theta$ and $t=\sin\theta$.

(d) Compute $4v-u^2$ in terms of the coefficients of $p(x)$ and $q(x)$.

(e) Try to link Pythagorean triples to angles θ with the property that $\cos\theta$ and $\sin\theta$ are both rational.

Problem of the Month

Solution to Problem 4: January 2021

(a) One such pair of polynomials is rather easy to spot:

$$(x+1)^2+1^2.$$

Here are two other pairs that are less easy to spot:

$$\left(\frac{3}{5}x+\frac{7}{5}\right)^2+\left(\frac{4}{5}x+\frac{1}{5}\right)^2$$

and

$$\left(\frac{5}{13}x+\frac{17}{13}\right)^2+\left(\frac{12}{13}x+\frac{7}{13}\right)^2.$$

In fact, if $m^2+n^2=p^2$ for integers m, n, and p, then

$$\left(\frac{m}{p}x+\frac{m+n}{p}\right)^2+\left(\frac{n}{p}x+\frac{n-m}{p}\right)^2=x^2+2x+2.$$

We will explore how one might discover such a parameterization of solutions in part (d).

(b) By completing the square, we can write $f(x)$ in vertex form. That is, there are h and k such that $f(x)=(x-h)^2+k$. Since $f(x)\geq 0$ for all x, $f(h)=k\geq 0$, so $\sqrt{k}$ is a real number. This means $f(x)=(x-h)^2+\left(\sqrt{k}\right)^2$. The function $q(x)=\sqrt{k}$ is a constant polynomial. In terms of u and v, one can check that $h=-\dfrac{u}{2}$ and $k=v-\dfrac{u^2}{4}$.

(c) Using the idea from part (b), we have that

$$x^2+x+1=\left(x+\frac{1}{2}\right)^2+\left(\frac{\sqrt{3}}{2}\right)^2,$$

so there are polynomials $p(x)$ and $q(x)$ that satisfy $x^2+x+1=(p(x))^2+(q(x))^2$. We will show that if $p(x)$ and $q(x)$ satisfy $x^2+x+1=(p(x))^2+(q(x))^2$, then at least one of $p(x)$ and $q(x)$ has an irrational coefficient. Observe that since $\dfrac{\sqrt{3}}{2}$ is irrational, this is the case for the particular example above.

Since x^2+x+1 is a quadratic, the degrees of both $p(x)$ and $q(x)$ must be no larger than 1. Otherwise, the sum of their squares would have degree at least 4. You may wish to think about this point. Keep in mind that the leading coefficient of the square of a polynomial must be positive. Therefore, there are real numbers a, b, c, and d (some of which could be zero) so that $p(x)=ax+b$ and $q(x)=cx+d$. So

$$\begin{aligned}x^2+x+1&=(ax+b)^2+(cx+d)^2\\&=(a^2+c^2)x^2+2(ab+cd)x+(b^2+d^2).\end{aligned}$$

If two polynomials written in standard form are equal, then they must have the same coefficients. Therefore, we get the equations

$$\begin{aligned} a^2 + c^2 &= 1 \\ b^2 + d^2 &= 1 \\ ab + cd &= \frac{1}{2}. \end{aligned}$$

From the first two equations, the points (a, c) and (b, d) lie on the unit circle. This means there are α, β in the interval $[0, 2\pi)$ satisfying the equations

$$\begin{aligned} \cos\alpha &= a \\ \sin\alpha &= c \\ \cos\beta &= b \\ \sin\beta &= d. \end{aligned}$$

Substituting this into $ab + cd = \frac{1}{2}$, we get

$$\cos\alpha\cos\beta + \sin\alpha\sin\beta = \frac{1}{2}$$

and after noticing that the left side of this equation is equal to $\cos(\alpha - \beta)$, we conclude that

$$\cos(\alpha - \beta) = \frac{1}{2}.$$

We have assumed that both α and β are in the interval $[0, 2\pi)$, so this means $\alpha - \beta$ is in the interval $(-2\pi, 2\pi)$. Together with $\cos(\alpha-\beta) = \frac{1}{2}$, this means $\alpha-\beta$ is one of $\pm\frac{\pi}{3}, \pm\frac{5\pi}{3}$.

We now assume that all of a, b, c, and d are rational and reach a contradiction. There are four cases corresponding to the possible values of $\alpha-\beta$. We will go through the case where $\alpha - \beta = -\frac{\pi}{3}$. The other cases can be handled similarly.

Assume $\alpha - \beta = -\frac{\pi}{3}$ which can be rearranged to $\beta = \alpha + \frac{\pi}{3}$. Then

$$\begin{aligned} d &= \sin\beta \\ &= \sin\left(\alpha + \frac{\pi}{3}\right) \\ &= \sin\alpha\cos\left(\frac{\pi}{3}\right) + \cos\alpha\sin\left(\frac{\pi}{3}\right) \\ &= \frac{c}{2} + a\frac{\sqrt{3}}{2}. \end{aligned}$$

Rearranging, we have

$$a\frac{\sqrt{3}}{2} = d - \frac{c}{2}.$$

If $a \neq 0$, then this can be further rearranged to $\sqrt{3} = \frac{2}{a}\left(d - \frac{c}{2}\right)$. The expression on the right is rational since a, c, and d are rational. However, $\sqrt{3}$ is irrational so this equality cannot be true. This means we must have $a = 0$.

From the equation $a^2 + c^2 = 1$, we then get $c^2 = 1$ so $c = \pm 1$. As well, $cd = \frac{1}{2}$, so $d = \pm\frac{1}{2}$, which means $d^2 = \frac{1}{4}$. Therefore, $b^2 = 1 - \frac{1}{4} = \frac{3}{4}$, so $b = \pm\frac{\sqrt{3}}{2}$, neither of which is rational. This contradicts the assumption that b is rational. We conclude that $\alpha - \beta \neq -\frac{\pi}{3}$ if all four of a, b, c, and d are to be rational. It can be shown that each of the other three possibilities for the value of $\alpha - \beta$ leads to a similar contradiction. Therefore, if $x^2 + x + 1 = (p(x))^2 + (q(x))^2$, then at least one of $p(x)$ and $q(x)$ must have an irrational coefficient.

(d) Suppose there is a special pair for $x^2 + ux + v$. This means there are polynomials $p(x)$ and $q(x)$, with all coefficients rational, such that $x^2 + ux + v = (p(x))^2 + (q(x))^2$. By the same reasoning as in part (c), $p(x)$ and $q(x)$ both have degree at most 1. Therefore, there are rational numbers a, b, c, and d so that

$$x^2 + ux + v = (ax + b)^2 + (cx + d)^2$$

Expanding and equating coefficients gives

$$\begin{aligned} a^2 + c^2 &= 1 \\ 2(ab + cd) &= u \\ b^2 + d^2 &= v. \end{aligned}$$

Since a, b, c, and d are rational, so are u and v (since products and sums of rational numbers are rational), which verifies the first part of the problem. To see that $4v - u^2$ is a rational square, observe that

$$\begin{aligned} 4v - u^2 &= 4(b^2 + d^2) - (2ab + 2cd)^2 \\ &= 4b^2 + 4d^2 - 4a^2b^2 - 8abcd - 4c^2d^2 \\ &= 4b^2(1 - a^2) + 4d^2(1 - c^2) - 8abcd \\ &= 4b^2c^2 + 4a^2d^2 - 8abcd && \text{(since } a^2 + c^2 = 1\text{)} \\ &= 4\left((ad)^2 - 2(ad)(bc) + (bc)^2\right) \\ &= 4(ad - bc)^2 \\ &= (2ad - 2bc)^2 \end{aligned}$$

Since a, b, c, and d are rational, the quantity $2ad - 2bc$ is also rational. Therefore, if we take $r = 2ad - 2bc$, then we have $4v - u^2 = r^2$.

(e) We assume that there is at least one special pair for x^2+ux+v. Using this assumption, our goal is to produce infinitely many quadruples of rational numbers (a, b, c, d) that satisfy $x^2 + ux + v = (ax + b)^2 + (cx + d)^2$.

As in part (d), we can expand the right side of $x^2 + ux + v = (ax + b)^2 + (cx + d)^2$ and equate coefficients to get

$$\begin{aligned} a^2 + c^2 &= 1 \\ 2(ab + cd) &= u \\ b^2 + d^2 &= v, \end{aligned}$$

so we seek rational solutions to the above system of equations. For the first equation to be satisfied, there must be a real number θ such that $\cos\theta = a$ and $\sin\theta = c$. Substituting these into the second equation above, we get

$$b\cos\theta + d\sin\theta = \frac{u}{2}. \tag{1}$$

Since there is at least one special pair for $x^2 + ux + v$, part (d) implies that there is a rational number r satisfying $r^2 = 4v - u^2$. By the calculation at the end of part (d), we have $4v - u^2 = (2ad - 2bc)^2$. Combining these two equations and taking square roots, we have $ad - bc = \pm\frac{r}{2}$. This means any quadruple satisfying the system must also satisfy either $ad - bc = \frac{r}{2}$ or $ad - bc = -\frac{r}{2}$. We will proceed by solving for the quadruples satisfying $ad - bc = \frac{r}{2}$ as this will be sufficient to find infinitely many special pairs. (You may wish to consider what happens if we solve for quadruples satisfying $ad - bc = -\frac{r}{2}$ as well.)

Substituting $a = \cos\theta$ and $c = \sin\theta$ into, $ad - bc = \frac{r}{2}$, we have

$$d\cos\theta - b\sin\theta = \frac{r}{2}. \tag{2}$$

We will now argue that if θ, b, and d satisfy equations (1) and (2), then taking $a = \cos\theta$ and $c = \sin\theta$ results in a quadruple (a, b, c, d) satisfying $x^2 + ux + v = (ax + b)^2 + (cx + d)^2$. (Keep in mind that in this context u, v, and r are fixed quantities and a, b, c, d, and θ are variables. Also, at this point we are not worrying about whether or not a, b, c, and d are rational. This part of the argument will wait until the end.) To that end, assume θ, b, and d satisfy equations (1) and (2) and set $a = \cos\theta$ and $c = \sin\theta$. We need to show that $a^2 + c^2 = 1$, $2(ab + cd) = u$, and $b^2 + d^2 = v$. That $a^2 + c^2 = 1$ follows from the Pythagorean identity. That $2(ab + cd) = u$ follows simply by multiplying equation (1) by 2. To see that $b^2 + d^2 = v$, we first square both sides of equations (1) and (2) to get

$$b^2\cos^2\theta + d^2\sin^2\theta + 2bd\cos\theta\sin\theta = \frac{u^2}{4}$$

and

$$d^2\cos^2\theta + b^2\sin^2\theta - 2bd\cos\theta\sin\theta = \frac{r^2}{4}.$$

Adding these two equations gives

$$(b^2 + d^2)\cos^2\theta + (b^2 + d^2)\sin^2\theta = \frac{1}{4}(u^2 + r^2).$$

Factoring and using that $r^2 = 4v - u^2$ we get

$$(b^2 + d^2)(\cos^2\theta + \sin^2\theta) = \frac{1}{4}(u^2 + 4v - u^2) = v$$

and since $\cos^2\theta + \sin^2\theta = 1$, we have the desired result.

We will next argue that for every real number θ, there are real numbers b and d so that equations (1) and (2) are both true. Since u and r are fixed quantities, if we fix θ and view b and d as variables, equations (1) and (2) give a system of two linear equations in two unknowns. To solve for b and d, we can first multiply equation (1) by $\cos\theta$ and equation (2) by $-\sin\theta$ to get

$$\begin{aligned} b\cos^2\theta + d\sin\theta\cos\theta &= \frac{u}{2}\cos\theta \\ b\sin^2\theta - d\sin\theta\cos\theta &= -\frac{r}{2}\sin\theta \end{aligned}$$

and adding these equations gives

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

which implies $b = \frac{1}{2}(u\cos\theta - r\sin\theta)$. A similar calculation can be used to find that $d = \frac{1}{2}(u\sin\theta + r\cos\theta)$. Thus, for any real θ, if we set $a = \cos\theta$, $b = \frac{1}{2}(u\cos\theta - r\sin\theta)$, $c = \sin\theta$, and $d = \frac{1}{2}(u\sin\theta + r\cos\theta)$, then we have $x^2 + ux + v = (ax+b)^2 + (cx+d)^2$.

We will now use this to show that there are infinitely many rational quadruples (a, b, c, d) satisfying $x^2+ux+v = (ax+b)^2+(cx+d)^2$. First, from part (d) we have that u and r are rational. This means that if $\cos\theta$ and $\sin\theta$ are both rational, then a, b, c, and d as defined above will also be rational. If we take θ to be one of the non-right angles in a right-angled triangle with integer side lengths, then we will have that $\cos\theta$ and $\sin\theta$ are both rational. (This is the idea that was used at the end of the solution to part (a).) However, this does not immediately imply that we get infinitely many distinct rational values of $\cos\theta$ because some of these triangles will be similar.

To deal with this issue, we will consider a particular family of right-angled triangles with integer side lengths. For each integer $n > 1$, the triangle with side lengths $n^2 - 1$, $2n$, and $n^2 + 1$ is a right-angled triangle. For each $n > 1$, define θ_n to be the angle at which the sides of length $2n$ and n^2+1 meet. Note that the side of length n^2+1 is the longest, so it must be the hypotenuse. Then $\cos\theta_n = \dfrac{2n}{n^2+1}$ and $\sin\theta_n = \dfrac{n^2-1}{n^2+1}$. Therefore, for every integer $n > 1$, we get a special pair $\big(ax+b, cx+d\big)$ by setting

$$\begin{aligned} a &= \frac{2n}{n^2+1} \\ c &= \frac{n^2-1}{n^2+1} \\ b &= \frac{1}{2}\left(\frac{2un}{n^2+1} - \frac{r(n^2-1)}{n^2+1}\right) \\ d &= \frac{1}{2}\left(\frac{u(n^2-1)}{n^2+1} + \frac{2rn}{n^2+1}\right) \end{aligned}$$

Furthermore, it is not a difficult exercise to show that $\cos\theta_{n+1} < \cos\theta_n$, which means that the values of a that arise this way are all different, so we indeed get infinitely many special pairs for $x^2 + ux + v$.

Problem of the Month
Problem 5: February 2021

Problem

For an integer $n \geq 3$, we define T_n to be the triangle with side lengths $n-1$, n, and $n+1$, and define A_n to be the area of T_n. We will say that an integer $n \geq 3$ is *remarkable* if A_n is an integer.

(a) Determine all integers n for which T_n is right-angled.

(b) Suppose n is a remarkable integer. Prove that

  (i) $\dfrac{n^2-4}{3}$ is a perfect square,

  (ii) n is not a multiple of 3, and

  (iii) n is even.

(c) There are three remarkable integers less than or equal to 100. Determine these three integers.

(d) The only remarkable integers between 100 and 10 000 are $n = 194$, $n = 724$, and $n = 2702$. Find a polynomial function $f(n)$ of degree greater than 1 with the property that if n is a remarkable integer, then $f(n)$ is also a remarkable integer. Use this polynomial to deduce that there are infinitely many remarkable integers.

(e) Explain how to find all remarkable integers. This should involve somehow describing an infinite set of remarkable integers as well as justification that your set is complete. Keep in mind that the infinite set from part (d) may not include *all* remarkable integers.

Hint

(a) No hint given.

(b) Using Heron's formula, it is possible to find an expression for A_n in terms of n.

(c) Sometimes it is faster to check all possibilities than to find a more clever approach. The conditions in part (b) can be used to eliminate about two thirds of the integers between 3 and 100.

(d) There is at least one such polynomial of degree 2.

(e) Try to find a few pairs (a, b) of positive integers that satisfy the equation $a^2 - 3b^2 = 1$. Compare the values of a to the known remarkable integers. Factoring the equation above as $(a + b\sqrt{3})(a - b\sqrt{3}) = 1$ and taking small positive integer powers of both sides may provide some insight into how one might generate more integer solutions to $a^2 - 3b^2 = 1$.

Problem of the Month

Solution to Problem 5: February 2021

This solution has an appendix containing various additional material relating to this problem. It is our hope that the solution can be read and understood without needing to look at the appendix.

(a) In a right-angled triangle, the hypotenuse is always the longest side. Therefore, if T_n is right-angled, then the Pythagorean Theorem implies that $(n-1)^2 + n^2 = (n+1)^2$ which can be simplified to $n^2 - 4n = 0$. Factoring the left side leads to $n(n-4) = 0$ so $n = 0$ or $n = 4$. Since $n \geq 3$, we must have $n = 4$. Indeed, the triangle with side lengths 3, 4, and 5 is right-angled by the *converse* of the Pythagorean Theorem (see Appendix (1)), so $n = 4$ is the only n for which T_n is right-angled.

(b) We will find a formula for A_n in terms of n. Here we present a derivation using Heron's formula. Appendix (2) contains a derivation using trigonometry.

Heron's formula states that if a triangle has side lengths a, b, and c, its area is

$$\sqrt{s(s-a)(s-b)(s-c)}$$

where the quantity $s = \dfrac{a+b+c}{2}$ is called the *semiperimeter* of the triangle.

The semiperimeter of T_n is $s = \dfrac{(n-1)+n+(n+1)}{2} = \dfrac{3n}{2}$. The other three quantities needed for Heron's formula are $s-(n-1)$, $s-n$, and $s-(n+1)$, which can be simplified to

$$\begin{aligned} s-(n-1) &= \frac{3n}{2} - \frac{2(n-1)}{2} \\ &= \frac{n+2}{2} \\ s-n &= \frac{3n}{2} - \frac{2n}{2} \\ &= \frac{n}{2} \\ s-(n+1) &= \frac{3n}{2} - \frac{2(n+1)}{2} \\ &= \frac{n-2}{2}. \end{aligned}$$

Using Heron's formula and simplifying gives

$$\begin{aligned} A_n &= \sqrt{s\big(s-(n-1)\big)\big(s-n\big)\big(s-(n+1)\big)} \\ &= \sqrt{\left(\frac{3n}{2}\right)\left(\frac{n+2}{2}\right)\left(\frac{n}{2}\right)\left(\frac{n-2}{2}\right)} \\ &= \frac{1}{4}\sqrt{3n^2(n-2)(n+2)} \\ &= \frac{n}{4}\sqrt{3(n^2-4)} && \text{(Ssince } n > 0\text{)} \end{aligned}$$

Now suppose that n is a remarkable integer, which means that A_n is an integer. We will verify that (i), (ii), and (iii) are true. Rearranging the formula for A_n from above, we get $\dfrac{4A_n}{n} = \sqrt{3(n^2-4)}$. The quantity $3(n^2-4)$ is a positive integer. Positive integers have the property that their square roots are either integers or irrational numbers. (For a proof of this, see Appendix (3)). Since $\sqrt{3(n^2-4)}$ is equal to $\dfrac{4A_n}{n}$ which is rational, this means $\sqrt{3(n^2-4)}$ not irrational and hence must be an an integer. Suppose $\sqrt{3(n^2-4)} = k$ for some positive integer k. Then $3(n^2-4) = k^2$, so k^2 is a multiple of 3. Since 3 is prime, k must be a multiple of 3, so $\dfrac{k}{3}$ is an integer and

$$\frac{n^2-4}{3} = \frac{k^2}{9} = \left(\frac{k}{3}\right)^2.$$

Thus, $\dfrac{n^2-4}{3}$ is a perfect square which proves (i).

Since $\dfrac{n^2-4}{3}$ is an integer, n^2-4 is a multiple of 3. If n^2 were a multiple of 3, then n^2-3 would be a multiple of 3, but n^2-3 and n^2-4 are consecutive integers, so they cannot both be multiples of 3. Thus, n^2 is not a multiple of 3, so n is not a multiple of 3. This proves (ii).

To see that n must be even, we can again set k to be an integer such that $3(n^2-4) = k^2$ which means $A_n = \dfrac{kn}{4}$. If n is odd, then so are n^2 and n^2-4. This means $k^2 = 3(n^2-4)$ is odd, which means k is odd. If n and k are both odd, then $A_n = \dfrac{kn}{4}$ cannot be an integer. Therefore, n is not odd, and hence must be even. This proves (iii).

(c) By part (b), if n is remarkable, then it is even and is not a multiple of 3. The table below contains each integer $n \le 100$ that is neither odd nor a multiple of 3, along with the corresponding value of $\dfrac{n^2-4}{3}$:

n	$\frac{n^2-4}{3}$	n	$\frac{n^2-4}{3}$	n	$\frac{n^2-4}{3}$
4	4	38	480	70	1632
8	20	40	532	74	1824
10	32	44	644	76	1924
14	64	46	704	80	2132
16	84	50	832	82	2240
20	132	52	900	86	2464
22	160	56	1044	88	2580
26	224	58	1120	92	2820
28	260	62	1280	94	2944
32	340	64	1364	98	3200
34	384	68	1540	100	3332

Again by part (b), if n is remarkable, then $\dfrac{n^2-4}{3}$ must be a perfect square. The only values of n in the table above for which $\dfrac{n^2-4}{3}$ is a perfect square are $n=4$, $n=14$, and $n=52$, so these are the only possibilities for remarkable integers less than or equal to

100. Since the problem states that there are three remarkable integers less than or equal to 100, the integers $n = 4$, $n = 14$, and $n = 52$ must be remarkable. For completeness, we compute the values of A_4, A_{14}, and A_{52} to verify that they are integers:

$$\begin{aligned} A_4 &= \frac{4}{4}\sqrt{3(4^2 - 3)} \\ &= 6 \\ A_{14} &= \frac{14}{4}\sqrt{3(14^2 - 4)} \\ &= 84 \\ A_{52} &= \frac{52}{4}\sqrt{3(52^2 - 4)} \\ &= 1170. \end{aligned}$$

(d) As suggested in the hint, we will try to find a polynomial function $f(n)$ of degree 2 with the desired property.

The first six remarkable integers are 4, 14, 52, 194, 724, and 2702. Squaring the first three, we get $4^2 = 16$, $14^2 = 196$, and $52^2 = 2704$. Subtracting 2 from these values gives 14, 194, and 2702, which are all remarkable integers. From this, we guess that the polynomial function $f(n) = n^2 - 2$ has the property that if n is a remarkable integer, then $f(n)$ is also a remarkable integer. Before verifying that this is true, we test the guess with $n = 194$, in which case $f(n) = 194^2 - 2 = 37\,634$. Indeed, with $n = 37\,634$, we have

$$\begin{aligned} A_n &= \frac{37\,634}{4}\sqrt{3(37\,634^2 - 4)} \\ &= 613\,283\,664 \end{aligned}$$

which is large, but it is an integer. We will now show that if n is remarkable, then $f(n)$ is remarkable. By the definition of remarkable, we must show that if A_n is an integer, then A_{n^2-2} is an integer. Suppose n is remarkable. By part (b), n is even and there is some k for which $n^2 - 4 = 3k^2$. By the formula for A_n, we have

$$\begin{aligned} A_{n^2-2} &= \frac{n^2-2}{4}\sqrt{3\big((n^2-2)^2 - 4\big)} \\ &= \frac{n^2-2}{4}\sqrt{3(n^4 - 4n^2)} \\ &= \frac{n^2-2}{4}\sqrt{3n^2(3k^2)} \\ &= \frac{3nk(n^2-2)}{4} && \text{(since } n, k > 0) \end{aligned}$$

which is an integer since both n and $n^2 - 2$ are even. This proves our claim that if n is remarkable, then $f(n) = n^2 - 2$ is remarkable.

To exhibit infinitely many remarkable integers, we need only observe that when $n > 2$, $n^2 - 2 > n$. To see this, note that $n^2 - n - 2 = (n - 2)(n + 1)$ has roots -1 and 2. Since $n^2 - n - 2$ is a quadratic with a positive leading coefficient, $n^2 - n - 2 > 0$ for all $n > 2$, which can be rearranged to $n^2 - 2 > n$ for all $n > 2$. Thus, starting at any remarkable integer greater than 2 (which they all are), the sequence $n, f(n), f(f(n)), f(f(f(n))), \ldots$ is an infinite sequence of remarkable integers. Starting with $n = 4$ gives the sequence

$$4, 14, 194, 37\,634, 1\,416\,317\,954, \ldots.$$

As a final observation, we notice that the infinite sequence above starting with 4 does *not* contain all remarkable integers. We can see immediately that $n = 52$, $n = 724$, and $n = 2702$ are missing. In part (e), we will explain how to find all remarkable integers.

(e) Appendix (5) contains a potentially interesting approach to finding all of the remarkable integers, but we give no proof that it works. What follows is an algebraic argument that relates the set of remarkable integers to the set of integer solutions to the equation $x^2 - 3y^2 = 1$, which is often called *Pell's equation*.

We first prove that $n \geq 3$ is remarkable if and only if there is a positive integer k for which $n^2 - 3k^2 = 4$. In part (b), we showed that if n is remarkable, then $n^2 - 4 = 3k^2$ for some k. Furthermore, $n \geq 3$, so k is nonzero and hence can be taken to be positive. We need to prove that if $n^2 - 4 = 3k^2$ for some positive integer k, then n is remarkable.

Note that if r is any odd integer, then r^2 is one more than a multiple of 4. (We will use this fact below.) To see this, observe that we must have $r = 2t + 1$ for some integer t and hence $r^2 = 4t^2 + 4t + 1 = 4(t^2 + t) + 1$.

Suppose that $n^2 - 4 = 3k^2$ for some positive integer k. We will prove that k and n are both even. We rewrite the equation as $n^2 = 3k^2 + 4$ and investigate what happens if k is odd. If k is odd, then $n^2 = 3k^2 + 4$ is odd, so n is odd. From the fact above, this means there are integers u and v so that $n^2 = 4u + 1$ and $k^2 = 4v + 1$. Substituting this into $n^2 = 3k^2 + 4$ gives $4u + 1 = 3(4v + 1) + 4$ which can be rearranged to $4(u - 3v - 1) = 2$. However, since $u - 3v - 1$ is an integer, this says that 4 is a factor of 2, which is not true. Therefore k cannot be odd and hence must be even. This implies $n^2 = 3k^2 + 4$ is also even, so n is even as well.

Computing A_n using $n^2 - 4 = 3k^2$, we have

$$A_n = \frac{n}{4}\sqrt{3(n^2 - 4)} = \frac{n}{4}\sqrt{3^2k^2} = \frac{3nk}{4}$$

which is an integer since both n and k are even.

Thus, we can find all remarkable integers n by finding all n for which there exists a positive integer k satisfying $n^2 - 4 = 3k^2$ or $n^2 - 3k^2 = 4$. We just argued that for this equation to hold, both n and k must be even. If we set $a = \frac{n}{2}$ and $b = \frac{k}{2}$, the equation becomes $a^2 - 3b^2 = 1$. We now have that the remarkable integers n are exactly the integers $n = 2a$ where a and b are positive integers satisfying $a^2 - 3b^2 = 1$.

Now what remains is solving Pell's equation: $x^2 - 3y^2 = 1$. More specifically, we want to find all pairs of nonnegative integers $(x, y) = (a, b)$ satisfying $x^2 - 3y^2 = 1$. It is easily checked that $(x, y) = (1, 0)$ is the only solution to Pell's equation where one of the variables is 0. All other solutions have both a and b positive.

For any nonnegative integer m, define a_m and b_m to be the unique integers satisfying $a_m + b_m\sqrt{3} = (2 + \sqrt{3})^m$. In practice, a_m and b_m can be found by expanding and collecting like terms. We will show that the solutions to Pell's equation are exactly the pairs (a_m, b_m).

Keep in mind that when $m = 0$, we have $(2 + \sqrt{3})^0 = 1$, so $a_0 = 1$ and $b_0 = 0$, which indeed gives a solution as observed earlier. For the exponents $m = 1$, $m = 2$, and $m = 3$, we get $(2+\sqrt{3})^1 = 2+\sqrt{3}$, $(2+\sqrt{3})^2 = 7+4\sqrt{3}$, and $(2+\sqrt{3})^3 = 26+15\sqrt{3}$. Thus, $(a_1,b_1) = (2,1)$, $(a_2,b_2) = (7,4)$, and $(a_3,b_3) = (26,15)$. Notice that $2^2 - 3(1^2) = 1$, $7^2 - 3(4^2) = 1$, and $26^2 - 3(15^2) = 1$. (As well, it is worth noting that the quantities $2 \times a_1 = 2 \times 2 = 4$, $2 \times a_2 = 2 \times 7 = 14$, and $2 \times a_3 = 2 \times 26 = 52$ are the first three remarkable integers.)

The rest of the solution is devoted to proving the following claims:

- (a_m, b_m) is a solution to Pell's equation for all $m \geq 0$.
- If (a,b) is a solution to Pell's equation, then $a = a_m$ and $b = b_m$ for some $m \geq 0$.

For the first point, the critical observation is that $(2-\sqrt{3})^m = a_m - b_m\sqrt{3}$ for all m. This can be easily checked for a few small m. For a complete proof, see Appendix (4).

Next, factor the right side of $2^2 - 3 = 1$ to get $(2+\sqrt{3})(2-\sqrt{3}) = 1$. Taking m^{th} powers of both sides and using exponent laws, we get

$$\begin{aligned} 1 = 1^m &= [(2+\sqrt{3})(2-\sqrt{3})]^m \\ &= (2+\sqrt{3})^m(2-\sqrt{3})^m \\ &= (a_m + b_m\sqrt{3})(a_m - b_m\sqrt{3}) \\ &= a_m^2 - 3b_m^2 \end{aligned}$$

which proves that (a_m, b_m) is a solution for each m.

For the second point, we first establish some terminology. Suppose (u_1, v_1) and (u_2, v_2) are both solutions. We say that (u_1, v_1) is smaller than (u_2, v_2) if $v_1 < v_2$. As it turns out, if $v_1 < v_2$ then $u_1 < u_2$ as well, but this fact will not be used. It was mentioned earlier that $(1,0)$ is the only solution with a 0 in the second coordinate, so this means $(1,0)$ is the unique *smallest* solution with respect to this notion of size.

Suppose (a,b) is a solution with $b > 0$. We will show that $2a - 3b > 0$ and $b > 2b - a \geq 0$. To do this, first divide $a^2 = 1 + 3b^2$ by b^2 to get $\left(\frac{a}{b}\right)^2 = \frac{1}{b^2} + 3$. Since b is a positive integer, it is at least 1, so $0 < \frac{1}{b^2} \leq 1$ which implies $3 < \left(\frac{a}{b}\right)^2 \leq 4$. Taking square roots, we get $\sqrt{3} < \frac{a}{b} \leq 2$ and since $\sqrt{3} \approx 1.732 > 1.5$, we have $\frac{3}{2} < \frac{a}{b} \leq 2$. The two inequalities rearrange to $2a - 3b > 0$ and $2b - a \geq 0$. Again from the observation that $\frac{a}{b} > \sqrt{3}$, we have that $a > b$, which implies $-b > -a$, and adding $2b$ to both sides gives $b > 2b - a$. Thus, $2a - 3b > 0$ and $b > 2b - a \geq 2$ as claimed.

Suppose again that (a,b) is a solution with $b > 0$. Consider

$$(a + b\sqrt{3})(2-\sqrt{3}) = (2a - 3b) + (2b - a)\sqrt{3}.$$

If we set $a' = 2a - 3b$ and $b' = 2b - a$, then $a' > 0$ and $b' \geq 0$ from the observations above, and

$$\begin{aligned} (a')^2 - 3(b')^2 &= (2a-3b)^2 - 3(2b-a)^2 \\ &= 4a^2 - 12ab + 9b^2 - 3(4b^2 - 4ab + a^2) \\ &= 4a^2 - 12ab + 9b^2 - 12b^2 + 12ab - 3a^2 \\ &= a^2 - 3b^2 \\ &= 1 \end{aligned}$$

so (a', b') is a solution as well. Furthermore, $b > b'$ as observed earlier, so (a', b') is a smaller solution than (a, b).

We have shown that if (a, b) is a solution with b positive, then multiplying $(a + b\sqrt{3})$ by $(2 - \sqrt{3})$ and collecting like terms gives a smaller solution. In this smaller solution, the value of y will be nonnegative. If it is positive, we can again multiply by $2 - \sqrt{3}$ to get an even smaller solution. Continuing, this can be repeated as long as the resulting y-value is positive to get smaller and smaller solutions. Since the y-values are decreasing and the process can be continued as long as the y-value is positive, we must eventually have that $(a + b\sqrt{3})(2 - \sqrt{3})^m$ gives the smallest solution. In other words, we eventually get $(a + b\sqrt{3})(2 - \sqrt{3})^m = 1 + 0\sqrt{3} = 1$. Multiplying both sides by $(2 + \sqrt{3})^m$, we get

$$(a + b\sqrt{3})(2 - \sqrt{3})^m(2 + \sqrt{3})^m = (2 + \sqrt{3})^m$$

and since $(2-\sqrt{3})^m(2+\sqrt{3})^m = [(2-\sqrt{3})(2+\sqrt{3})]^m = 1$ and $(2+\sqrt{3})^m = a_m + b_m\sqrt{3}$ by definition, we have $a + b\sqrt{3} = a_m + b_m\sqrt{3}$, so $a = a_m$ and $b = b_m$. This proves the second claim.

Putting this all together, we have proved the desired result:

If $c_1, c_2, c_3, \ldots, c_m, \ldots$ is the complete list of remarkable integers, in increasing order, then $c_m = 2a_m$ where a_m is the "integer part" of $(2 + \sqrt{3})^m$ as defined above.

Note: You may have observed that the c_m satisfy the *recurrence* $c_{m+2} = 4c_{m+1} - c_m$ for all $m \geq 1$. It is not too difficult to prove this using the description in the previous paragraph.

With $c_1 = 4$ and $c_2 = 14$, this recurrence gives

$$\begin{aligned}
c_3 &= 4c_2 - c_1 \\
&= 4(14) - 4 \\
&= 52 \\
c_4 &= 4c_3 - c_2 \\
&= 194 \\
c_5 &= 724 \\
c_6 &= 2702 \\
c_7 &= 10\,084
\end{aligned}$$

and so on.

One final interesting observation is that for all $m \geq 1$ we have

$$c_m = (2 + \sqrt{3})^m + (2 - \sqrt{3})^m,$$

This is also not too hard to justify using earlier observations. Since $0 < 2 - \sqrt{3} < 1$, this means for "large" m, $c_m \approx (2 + \sqrt{3})^m$. You might want to explore this with a calculator.

Appendix to Solution

1. (Converse of the Pythagorean Theorem) The Pythagorean Theorem is typically phrased in a way similar to "if a and b are the lengths of the legs in a right triangle and c is the length of its hypotenuse, then $a^2 + b^2 = c^2$.". The converse of the Pythagorean theorem says that if a, b, and c are real numbers satisfying $a^2 + b^2 = c^2$, then the triangle with side lengths a, b, and c is right-angled with a and b the lengths of the legs and c the length of the hypotenuse. It is not difficult to find proofs of this fact online.

2. (Derivation of formula for A_n using trigonometry) For the other derivation of this formula, we first let θ be the angle opposite the side of length $n+1$. By the Law of Cosines, we have
$$(n+1)^2 = n^2 + (n-1)^2 - 2n(n-1)\cos\theta$$
which leads to
$$\cos\theta = \frac{n^2-4n}{2n(n-1)} = \frac{n-4}{2(n-1)}.$$
By the Pythagorean identity, $\cos^2\theta + \sin^2\theta = 1$. As well, θ is an angle in a triangle, so $\sin\theta$ is positive. Therefore,
$$\begin{aligned}\sin\theta &= \sqrt{1-\left(\frac{n-4}{2n-2}\right)^2}\\ &= \sqrt{\left(\frac{2n-2}{2(n-1)}\right)^2 - \left(\frac{n-4}{2(n-1)}\right)^2}\\ &= \frac{1}{2(n-1)}\sqrt{4n^2-8n+4-n^2+8n-16}\\ &= \frac{1}{2(n-1)}\sqrt{3n^2-12}\\ &= \frac{1}{2(n-1)}\sqrt{3(n^2-4)}\end{aligned}$$
The area of a triangle with sides a and b meeting at angle θ is equal to $\frac{1}{2}ab\sin\theta$. Using this and the result of the calculation above, we get
$$\begin{aligned}A_n &= \frac{1}{2}n(n-1)\sin\theta\\ &= \frac{1}{2}n(n-1)\frac{1}{2(n-1)}\sqrt{3(n^2-4)}\\ &= \frac{n}{4}\sqrt{3(n^2-4)}.\end{aligned}$$

3. (Square root of an integer is an integer or irrational) To prove that the square root of a positive integer must be an integer or irrational, we assume that N is a positive integer with $\sqrt{N}$ rational and prove that $\sqrt{N}$ must be an integer. Hence, assume m and n are integers with $\frac{m^2}{n^2} = N$. We can assume that $\frac{m}{n}$ is in lowest terms. Since $\frac{m^2}{n^2}$ is an integer, m^2 must be a multiple of n^2, which means any prime factor of n^2 is also a prime factor of m^2. If n has a prime factor p, then p is a factor of m^2, so p is also a factor of m. We have assumed $\frac{m}{n}$ is in lowest terms, so we are forced to conclude that n has no prime factor. This means $n \pm 1$ so $\frac{m}{n} = \pm m$. Therefore, $\sqrt{N}$ is an integer.

4. ($(2-\sqrt{3})^m = a_m - b_m\sqrt{3}$) This argument assumes a knowledge of the binomial theorem and of binomial coefficients. To establish that $(2-\sqrt{3})^m = a_m - b_m\sqrt{3}$, we will use the

binomial theorem. First, observe that

$$
\begin{aligned}
(2+\sqrt{3})^m &= \sum_{k=0}^{m} \binom{m}{k} 2^{m-k}\sqrt{3}^k \\
&= \sum_{\substack{0\le k\le m\\ k \text{ even}}} \binom{m}{k} 2^{m-k}\sqrt{3}^k + \sum_{\substack{0\le k\le m\\ k \text{ odd}}} \binom{m}{k} 2^{m-k}\sqrt{3}\sqrt{3}^{k-1} \\
&= \sum_{\substack{0\le k\le m\\ k \text{ even}}} \binom{m}{k} 2^{m-k}3^{\frac{k}{2}} + \sum_{\substack{0\le k\le m\\ k \text{ odd}}} \binom{m}{k} 2^{m-k}(\sqrt{3})3^{\frac{k-1}{2}} \\
&= \sum_{\substack{0\le k\le m\\ k \text{ even}}} \binom{m}{k} 2^{m-k}3^{\frac{k}{2}} + \sqrt{3}\sum_{\substack{0\le k\le m\\ k \text{ odd}}} \binom{m}{k} 2^{m-k}3^{\frac{k-1}{2}}
\end{aligned}
$$

and so we see that

$$a_m = \sum_{\substack{0\le k\le m\\ k \text{ even}}} \binom{m}{k} 2^{m-k}3^{\frac{k}{2}}$$

and

$$b_m = \sum_{\substack{0\le k\le m\\ k \text{ odd}}} \binom{m}{k} 2^{m-k}3^{\frac{k-1}{2}}$$

In a similar way, we can compute $(2-\sqrt{3})^m$ to get

$$
\begin{aligned}
(2-\sqrt{3})^m &= \sum_{k=0}^{m} \binom{m}{k} 2^{m-k}(-\sqrt{3})^k \\
&= \sum_{\substack{0\le k\le m\\ k \text{ even}}} \binom{m}{k} 2^{m-k}(-\sqrt{3})^k + \sum_{\substack{0\le k\le m\\ k \text{ odd}}} \binom{m}{k} 2^{m-k}(-\sqrt{3})(-\sqrt{3})^{k-1} \\
&= \sum_{\substack{0\le k\le m\\ k \text{ even}}} \binom{m}{k} 2^{m-k}3^{\frac{k}{2}} + \sum_{\substack{0\le k\le m\\ k \text{ odd}}} \binom{m}{k} 2^{m-k}(-\sqrt{3})3^{\frac{k-1}{2}} \\
&= \sum_{\substack{0\le k\le m\\ k \text{ even}}} \binom{m}{k} 2^{m-k}3^{\frac{k}{2}} - \sqrt{3}\sum_{\substack{0\le k\le m\\ k \text{ odd}}} \binom{m}{k} 2^{m-k}3^{\frac{k-1}{2}} \\
&= a_m - \sqrt{3}b_m.
\end{aligned}
$$

5. (Different approach to finding all remarkable integers) Below are the values of A_n for $n = 4$, $n = 14$, $n = 52$, $n = 194$, $n = 724$, and $n = 2702$.

$$
\begin{aligned}
A_4 &= 6 \\
A_{14} &= 84 \\
A_{52} &= 1170 \\
A_{194} &= 16\,296 \\
A_{724} &= 226\,974 \\
A_{2702} &= 3\,161\,340
\end{aligned}
$$

There may not appear to be a pattern here, but consider the ratios below:

$$\begin{aligned}
\frac{A_{14}}{A_4} &= 14 \\
\frac{A_{52}}{A_{14}} &\approx 13.92857142857142 \\
\frac{A_{194}}{A_{52}} &\approx 13.92820512820512 \\
\frac{A_{724}}{A_{194}} &\approx 13.92820324005891 \\
\frac{A_{2702}}{A_{724}} &\approx 13.92820323032594
\end{aligned}$$

and so it appears that the sequence of integer areas is very close to being geometric with a ratio around 13.92820323032594. In fact, the quantity these ratios are approaching is $(2+\sqrt{3})^2 = 13.9282032302755\ldots$. Try to verify this using the result at the very end of the solution to part (e)!

Using this apparent pattern, if we suppose N is the next remarkable integer after 2702, then we expect $\frac{A_N}{A_{2702}} \approx 13.9282032302755$. Indeed, if we multiply A_{2702} by 13.928203230275509, we get $44031785.99999918\ldots$, which is extremely close to the integer 44 031 786. Thus, we guess that $A_N = 44\,031\,786$, which would mean that

$$\frac{N}{4}\sqrt{3(N^2-4)} = 44\,031\,786.$$

Multiplying by 4, squaring both sides, then dividing by 3 gives

$$N^2(N^2-4) = 10\,340\,256\,951\,198\,912$$

which can be rearranged to

$$N^4 - 4N^2 - 10\,340\,256\,951\,198\,912 = 0$$

which is a quadratic in N^2. Using a calculator (or a lot of paper on a rainy afternoon), we can use the quadratic formula to get

$$N^2 = \frac{4 \pm \sqrt{16 + 4 \times 10\,340\,256\,951\,198\,912}}{2} = 2 \pm 101\,687\,054$$

which means $N^2 = 101\,687\,056$ and so $N = 10\,084$. Indeed, 10 084 is the next remarkable integer after 2702. Repeating this (rather impractical) process will indeed find all n for which A_n is an integer, provided one possesses arbitrarily good rational approximations of $(2+\sqrt{3})^2$.

Problem of the Month

Problem 6: March 2021

Problem

Here is a simple activity that leads to an interesting math problem.

- For a positive integer $n > 1$, draw n dots on a piece of paper. Draw a line to connect each pair of dots. The lines do not need to be straight, but should be drawn so that they do not pass through any dots other than the two they connect. If two lines intersect, the intersection does **not** define a new dot.
- Colour each dot either red or blue in any way that you like.
- Colour each line as follows: If the line connects two dots of the same colour, colour the line red. Otherwise, colour the line blue.

Call a colouring of the dots *balanced* if it leads to the lines being coloured so that there is the same number of blue lines as red lines.

(a) Show that there is no balanced colouring when $n = 5$.

(b) Show that there is a balanced colouring when $n = 9$. Find all possibilities for the number of red dots in a balanced colouring when $n = 9$.

(c) Determine all n for which there is a balanced colouring. For each such n, determine all possibilities for the number of red dots in a balanced colouring.

For part (d), the dots can now be coloured red, blue, or green. The table below describes how the lines should be coloured once the dots are coloured. For example, the letter R is in the cell corresponding to the row for B and the column for G. This means that if a line connects one blue dot and one green dot, then it is to be coloured red.

	R	G	B
R	R	G	B
G	G	B	R
B	B	R	G

For part (d), we redefine a *balanced* colouring of the dots to mean a colouring leads to equal numbers of red, blue, and green lines.

(d) Describe all n for which there is a balanced colouring.

Hint

In all three parts, it is useful to introduce a variable for the number of dots of each colour. For example, you might set r to be the number of red dots and b to be the number of blue dots.

(a) Express the number of blue lines in terms of r and b.

(b)/(c) Determine expressions for the number of red lines and blue lines and set them equal to each other.

(d) An approach similar the one suggested for parts (b) and (c) should work here. There should be a new variable for the number of green dots and there will now be three quantities that must be equal to each other.

Problem of the Month

Solution to Problem 6: March 2021

(a) Since there are 5 dots, each dot is connected to $5 - 1 = 4$ dots. This gives a total of $\frac{5 \times 4}{2} = 10$ lines where the division by 2 is because the product 5×4 counts each line twice.

Suppose r is the number of red dots and b is the number of blue dots. Since $n = 5$ and every dot must be coloured, $r + b = 5$. Only the lines connecting a blue dot to a red dot are coloured blue, and since each blue dot is connected to each red dot, there are exactly rb blue lines. If a colouring is balanced, then the number of blue lines would be $\frac{10}{2} = 5$. Thus, if the colouring is balanced, then $rb = 5$. Since r and b are nonnegative integers, this means $r = 1$ and $b = 5$ or $r = 5$ and $b = 1$. In each case, $r + b \neq 5$, so there can be no balanced colouring when $n = 5$.

(b) Similar to the argument in part (a), the number of lines when $n = 9$ is $\frac{9 \times 8}{2} = 36$. As well, if we let r be the number of red dots and b be the number of blue dots, then the number of blue lines is rb. For a colouring to be balanced, we need $rb = 18$. Therefore, we are looking for nonnegative integers r and b such that $rb = 18$ and $r + b = 9$.

Since $rb = 18$ and both r and b are nonzero, we have $b = \frac{18}{r}$. Substituting this expression into $r+b = 9$ gives $r+\frac{18}{r} = 9$. Multiplying through by r and rearranging gives $r^2-9r+18 = 0$, which can be factored to get $(r-6)(r-3) = 0$. Therefore, the number of red dots must be either 3 or 6.

If $r = 3$, then $b = 9 - 3 = 6$, so the number of blue lines is $3 \times 6 = 18$. This means the number of red lines is $36 - 18 = 18$. If $r = 6$, then $b = 9 - 6 = 3$, so the number of blue lines is $6 \times 3 = 18$ and the number of red lines is $36 - 18 = 18$ as well.

Thus, colouring the dots so that 3 are red or 6 are red gives a balanced colouring, and there are no other possibilities.

(c) When there are n dots, there are $\frac{n(n-1)}{2}$ lines. Once again, we set r to be the number of red dots and b to be the number of blue dots, so that $rb = \frac{1}{2} \times \frac{n(n-1)}{2} = \frac{n(n-1)}{4}$ is the number of blue lines in a balanced colouring.

Therefore, we wish to find all integers $n > 1$ for which there are nonnegative integers r and b satisfying $r + b = n$ and $rb = \frac{n(n-1)}{4}$.

By squaring both sides of the equation $r+b = n$, we obtain $r^2+2rb+b^2 = n^2$. Multiplying both sides of the equation $rb = \frac{n(n-1)}{4}$ by 4 gives $4rb = n^2-n$. Subtracting this equation from $r^2 + 2rb + b^2 = n^2$ gives $r^2 - 2rb + b^2 = n$ which factors as $(r - b)^2 = n$. Therefore, if there is a balanced colouring, then n must be a perfect square.

To finish the argument, we will show that if n is a perfect square, then there is a balanced colouring. To get an idea how to do this, let us suppose $n = m^2$ for some positive integer m. We know that if a balanced colouring exists, then $(r - b)^2 = m^2$. If $r > b$, then $r - b = m$. Adding this to $r + b = n$ and dividing by 2, we have $r = \frac{n+m}{2}$. It then follows that $b = \frac{n-m}{2}$. If $r < b$, we get that $r = \frac{n-m}{2}$ and $b = \frac{n+m}{2}$. This shows that if a balanced colouring exists, then we must have that $r = \frac{n \pm \sqrt{n}}{2}$.

If n is a perfect square, then we can let $r = \frac{n+\sqrt{n}}{2}$ and $b = \frac{n-\sqrt{n}}{2}$ which gives

$$rb = \frac{n^2 - \sqrt{n}^2}{4} = \frac{1}{2} \times \frac{n^2 - n}{2}$$

and so the colouring is balanced. A nearly identical calculation shows that we can let $r = \frac{n-\sqrt{n}}{2}$ and $b = \frac{n+\sqrt{n}}{2}$ and we would also get a balanced colouring.

Therefore, there is a balanced colouring exactly when n is a perfect square. Moreover, if n is a perfect square, then the number of red dots in a balanced colouring must be either $\frac{n+\sqrt{n}}{2}$ or $\frac{n-\sqrt{n}}{2}$. We point out that when n is a perfect square, n and $\sqrt{n}$ are either both even or both odd. This means the numerators $n + \sqrt{n}$ and $n - \sqrt{n}$ are even, so the numbers of red dots given above are both integers.

(d) Suppose r is the number of red dots, b is the number of blue dots, and g is the number of green dots. The number of lines is $\frac{n(n-1)}{2}$, so in a balanced colouring, we need to have $\frac{n(n-1)}{6}$ lines of each colour.

The lines that are coloured red are the lines connecting two red dots or the lines connecting a blue dot to a green dot. The number of lines connecting red dots to red dots is $\frac{r(r-1)}{2}$. This is because each of the r red dots is connected to the $r - 1$ other red dots, so $r(r - 1)$ counts each such line twice. The number of lines connecting blue dots to green dots is bg. Therefore, the number of red lines is

$$\frac{r(r-1)}{2} + bg$$

Similar reasoning shows that the number of blue lines is

$$\frac{g(g-1)}{2} + rb$$

and that the number of green lines is

$$\frac{b(b-1)}{2} + rg.$$

Suppose the numbers of red lines, blue lines, and green lines are all equal. In particular, the number of red lines equals the number of blue lines, so

$$\frac{r(r-1)}{2} + bg = \frac{g(g-1)}{2} + rb.$$

Multiplying this equation by 2 and expanding gives $r^2-r+2bg = g^2-g+2rb$. Rearranging this, we get

$$(r^2 - g^2) + (g - r) + (2bg - 2rb) = 0$$

which has a common factor of $r - g$ and can be rewritten as

$$(r - g)(r + g - 1 - 2b) = 0. \tag{1}$$

Similarly, the number of blue lines equals the number of green lines, so

$$\frac{g(g-1)}{2} + rb = \frac{b(b-1)}{2} + rg$$

which is equivalent to

$$g^2 - b^2 + b - g + 2rb - 2rg = 0$$

and then factored as

$$(g - b)(g + b - 1 - 2r) = 0. \tag{2}$$

Equating the number of red lines and the number of green lines gives

$$\frac{r(r-1)}{2} + bg = \frac{b(b-1)}{2} + rg,$$

which is equivalent to

$$(r - b)(r + b - 1 - 2g) = 0. \tag{3}$$

Suppose r, b, and g are three distinct integers. This means $r - g \neq 0$, so Equation (1) implies $r + g - 1 - 2b = 0$ or $r + g = 1 + 2b$. Similarly, $g - b \neq 0$ and $r - b \neq 0$, so Equations (2) and (3) imply $b + g = 1 + 2r$ and $r + b = 1 + 2g$, respectively. Adding these three equations, we get

$$(r + g) + (b + g) + (r + b) = (1 + 2b) + (1 + 2r) + (1 + 2g)$$

which implies $2(r + b + g) = 3 + 2(r + b + g)$ so $3 = 0$. Of course, this is not true, which means r, b, and g cannot be three distinct integers. In other words, at least two of r, b, and g are equal to each other.

One way for Equations (1), (2), and (3) to be satisfied simultaneously is when $r = b = g$. This implies that there is some positive integer k with $n = 3k$ and $r = b = g = k$.

Otherwise, there are three possibilities: $r = b$ with g different from r and b, $r = g$ with b different from r and g, and $b = g$ with r different from b and g.

If $r = b$ and g is different from r and b, then $r \neq g$, so Equation (1) implies $r+g-1-2b = 0$. Substituting $r = b$, we get $g - 1 - b = 0$ or $g = 1 + b$. This shows that g is one more than the common value of r and b. Thus, there is some positive integer k so that $n = 3k + 1$ and $r = b = k$ and $g = k + 1$.

Similar analysis shows that if $r = g$ with b different from r and b, then $n = 3k+1$ for some positive integer k and $r = g = k$ with $b = k + 1$. As well, if $b = g$ and r is different form b and g, then $n = 3k + 1$ with $b = g = k$ and $r = k + 1$.

We have now argued that if there is a balanced colouring of n dots, then one of these two statements must be true:

- $n = 3k$ for some positive integer k and there are k dots of each colour.

- $n = 3k + 1$ for some positive integer k and (r, b, g) is one of $(k, k, k + 1)$, $(k, k + 1, k)$, and $(k + 1, k, k)$.

To finish the solution, we will check that the colouring is indeed balanced in each of the four situations described.

If $n = 3k$ for some positive integer k and $r = b = g = k$, then the number of red lines is

$$\begin{aligned}\frac{k(k-1)}{2} + k^2 &= \frac{k^2 - k}{2} + k^2 \\ &= \frac{3k^2 - k}{2} \\ &= \frac{9k^2 - 3k}{6} \\ &= \frac{3k(3k-1)}{6} \\ &= \frac{n(n-1)}{6}.\end{aligned}$$

A similar calculation shows that there are the same number of blue and green lines. Therefore, there is a balanced colouring when $n = 3k$ for some positive integer k.

If $n = 3k + 1$ for some positive integer k and $r = b = k$ with $g = k + 1$, then the number of red lines is

$$\begin{aligned}\frac{k(k-1)}{2} + k(k+1) &= \frac{k^2 - k + 2k^2 + 2k}{2} \\ &= \frac{3k^2 + k}{2} \\ &= \frac{9k^2 + 3k}{6} \\ &= \frac{3k(3k+1)}{6} \\ &= \frac{(n-1)n}{6}.\end{aligned}$$

The number of green lines is also $\dfrac{n(n-1)}{6}$ by essentially an identical calculation, and the number of blue lines is

$$\begin{aligned}\frac{(k+1)k}{2} + k^2 &= \frac{3k^2 + k}{2} \\ &= \frac{n(n-1)}{6}\end{aligned}$$

by another similar calculation. The situations when $r = g = k$ and $b = k+1$ and $b = g = k$ and $r = k + 1$ can be handled similarly.

We have now shown that when $n = 3k$ or $n = 3k + 1$, there is a balanced colouring. Therefore, a balanced colouring exists exactly when $n = 3k$ or $n = 3k+1$ for some positive integer k.

Further Remark

If you are familiar with *modular arithmetic*, then you might have noticed a pattern in the rules for colouring the lines. In the first three parts, we could have said that each point is "coloured" by either 0 or 1, and the colour of a line is the *sum modulo* 2 of the "colours" of the dots it connects. "Modulo 2" means the remainder after division by 2. So in this system of arithmetic, $0 + 0 = 0$, $0 + 1 = 1 + 0 = 1$, and $1 + 1 = 0$ since the remainder after dividing 2 by 2 is 0. In our problem, red corresponds to 0 and blue corresponds to 1. Similarly, in part (d) we might instead label each dot by 0, 1, or 2 (red is 0, blue is 1, and green is 2). The label is the *sum modulo 3*, the remainder after division by 3. This means $0 + 0 = 0$, $0 + 1 = 1 + 0 = 1$, $0 + 2 = 2 + 0 = 2$, $1 + 2 = 2 + 1 = 0$ (since the remainder after dividing 3 by 3 is 0), and $2 + 2 = 1$ (since the remainder after dividing 4 by 3 is 1).

With this in mind, we can imagine labelling the dots by the integers 0, 1, 2, or 3 and labelling the lines by the sum *modulo* 4, or even more generally, we could label dots label by 0, 1, 2, 3, and so on up to $n-1$ and label the lines by the sum modulo n. What should the definition of "balanced" be in general, and what can you say about balanced colourings in general?

Problem of the Month

Problem 7: April 2021

Problem

For an integer $n \geq 3$, n^2 points form an $n \times n$ square grid.

Define $P(n)$ to be the probability that three distinct points randomly selected from the grid are the vertices of a triangle with positive area. Also define $f(n)$ to be the number of sets of three distinct points from the grid that lie on a common line. We can think of $f(n)$ as the number of sets of three distinct points from the grid that are the vertices of a triangle with area 0.

For instance, with $n = 3$, it can be shown that there are 84 possible ways to select three distinct points, that 8 of the sets of three points lie on a line, and that 76 of the sets of three points form the vertices of a triangle with positive area. Thus, $f(3) = 8$ and $P(3) = \dfrac{76}{84} = \dfrac{19}{21}$.

The goal of this problem is to estimate $P(n)$ for large n. The approach outlined will be to estimate $f(n)$ and use it to estimate $P(n)$.

(a) When $n = 3$, $f(n) = 8$ and $P(3) = \dfrac{19}{21}$. Compute $f(n)$ and $P(n)$ for $n = 4$ and $n = 5$.

(b) For $n \geq 3$, prove that $f(n+1) < f(n)+5n^4+5n^3+5n^2+5n$. This will allow us to understand how quickly $f(n)$ grows which will help to estimate $P(n)$.

(c) Using part (b), prove that $f(n) < n^5$ for all $n \geq 3$.

(d) Prove that there is a constant c with the property that $P(n) > 1 - \dfrac{c}{n}$ for all $n \geq 3$. Use this to explain why the following statement makes sense: "For very large n, it is nearly certain that three points selected randomly from an $n \times n$ grid will be the vertices of a triangle with positive area."

As indicated in part (d), this problem is meant to examine what happens to $P(n)$ as n gets large. Since it seems very difficult to calculate $f(n)$ (and hence, $P(n)$) directly for large n, we instead estimate its value. As long as we carefully keep track of how good/bad the estimates can be, we can say something meaningful about $P(n)$ for large n without actually computing it directly. Very frequently, mathematicians use estimates like these when exact answers are difficult or impossible to obtain. These estimates are often as useful as exact answers.

Hint

(a) Drawing a picture (or a few pictures) could be helpful. Most of the work is in computing $f(n)$. One way to do this is to consider the various slopes that a line through three points in the grid can have. It might also be useful to do an internet search on *binomial coefficients*. A binomial coefficient, sometimes denoted $\binom{n}{k}$ (read "n choose k"), is equal to the number of ways to choose k objects from a set of n distinct objects. It can be computed as $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$.

(b) Consider breaking an $(n+1) \times (n+1)$ grid into two sets of points: the bottom-left $n \times n$ subgrid and the $2n+1$ points along the top and right. If three points lie on a line, how can they be distributed among these two sets? Remember, the goal is not to compute $f(n+1)$ exactly, but to bound it by $f(n) + 5n^4 + 5n^3 + 5n^2 + 5n$. In fact, you might be able to do better than this by showing, for example, that $f(n+1) < f(n) + 2n^4 + 2n^3 + 2n^2 + 2n$, or some other bound involving $f(n)$.

(c) Expand $(n+1)^5$.

(d) Use part (c). There are many values of c that will work. It will likely be helpful to show some other inequalities like $n^2 - 1 \leq \frac{1}{2}n^2$ for $n \geq 3$.

Problem of the Month

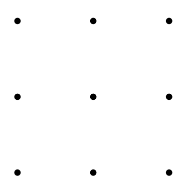
Solution to Problem 7: April 2021

As suggested in the hint, will use *binomial coefficients* throughout this solution. The symbol $\binom{m}{k}$ (in words, "m choose k") represents the number of ways to choose k objects from m distinct objects. We will use explicitly that $\binom{m}{2} = \frac{m(m-1)}{2}$ and $\binom{m}{3} = \frac{m(m-1)(m-2)}{6}$, but you may wish to explore this standard notation more generally.

We will also use the standard terminology that three points are *collinear* if they are on a common line.

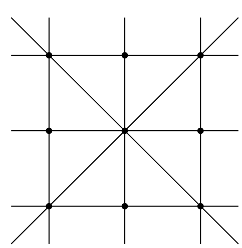
(a) Although it was not part of the problem, we will show that $f(3) = 76$ and $P(3) = \frac{19}{21}$.

Consider the 3×3 grid below:



We can draw three vertical lines and three horizontal lines such that each line passes through exactly three points. This accounts for a total of 6 ways to choose three distinct collinear points from the grid.

If three points are collinear but the line they define is neither vertical nor horizontal, then each point must be in a different row and a different column. The only sets of three collinear points with one point in each row and each column are the two "diagonals". This gives two more sets of three points, for a total of 8. Thus, $f(3) = 8$. The diagram below shows all eight of the lines that contain three points:



To compute $P(3)$, we will need the total number of ways to choose three distinct points from the $3 \times 3 = 9$ points in the grid. This is equal to $\binom{9}{3} = \frac{9 \times 8 \times 7}{6} = 84$.

There are 84 ways to choose three distinct points, and there are $f(3) = 8$ ways to choose three distinct points that are collinear. Each of the remaining $84 - 8 = 76$ sets of three distinct points are the vertices of a triangle of positive area, so $P(3) = \frac{76}{84} = \frac{19}{21}$.

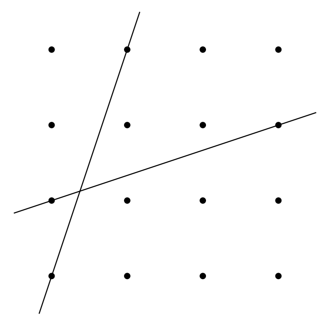
To compute $P(4)$, we will compute $f(4)$ and use that the number of ways to choose three distinct points from the $4 \times 4 = 16$ points in the grid is $\binom{16}{3} = 560$. Then we can compute

$P(4) = \dfrac{560 - f(4)}{560}$.

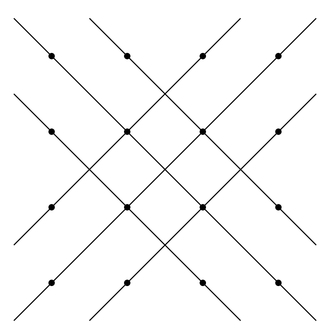
To count the sets of three collinear points, we will first find all lines that pass through at least two of the points. To make sure we do not miss any, we will examine the possible slopes of lines through at least two points, imagining that the bottom-left point is at the origin and the others are the points (a, b) where $0 \le a \le 3$ and $0 \le b \le 3$. If two points are chosen, then the slope can be computed as $\dfrac{\text{“rise”}}{\text{“run”}}$. Ignoring vertical and horizontal lines for now, the possible rises are -3, -2, -1, 1, 2, and 3 and the possible runs are the same set of values. Thus, the possible slopes of lines that are neither vertical nor horizontal are

$$\pm 3,\ \pm 2,\ \pm\frac{3}{2},\ \pm 1,\ \pm\frac{2}{3},\ \pm\frac{1}{2},\ \pm\frac{1}{3}$$

If a line that is neither horizontal nor vertical passes through three distinct points, then these three points must be in three different rows and three different columns. Suppose such a line has slope 3. Then the two points on this line that are farthest from each other must be a vertical distance of at least 2 apart, and hence, must be a vertical distance of at least 6 apart. There are only four rows of points, so it is impossible for a line of slope 3 to pass through three points in the grid. Similarly, a line of slope $\dfrac{1}{3}$ cannot pass through three points in the grid. The diagram below may help to illustrate this.



Similar arguments can be used to show that a line of any of the slopes above, except 1 and -1, can pass through at most two points in the grid. Thus, if a line passes through three or more points in the grid, it must have slope ± 1 or be horizontal or vertical. There are six lines of slope ± 1 that pass through at least three points: Two diagonals, one above each diagonal, and one below each diagonal. They are shown below:



The four lines other than the diagonals each pass through exactly three points in the grid. Thus, we get four sets of three collinear points.

Each diagonal passes through four points in the grid. From each such line, we can choose a set of three collinear points by ignoring one of the four points. Thus, each of these two lines contributes another four sets of three collinear points. So far, we have counted $4 + 4 + 4 = 12$ sets of three collinear points in the grid.

Each horizontal line passes through four points. By the same reasoning as in the previous paragraph, the horizontal lines each contribute four sets of three collinear points for a total of $4 \times 4 = 16$ more sets of collinear points. We similarly get 16 sets of three collinear points from the four vertical lines. In total, $f(4) = 12 + 16 + 16 = 44$. We can now compute $P(4)$ as

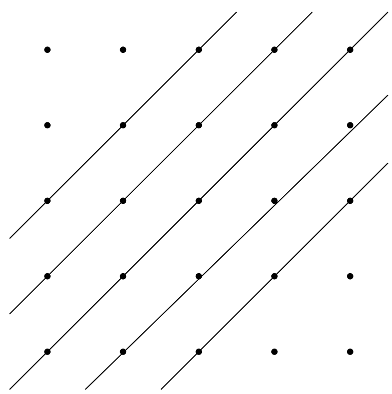
$$P(4) = \frac{560 - f(4)}{560} = \frac{560 - 44}{560} = \frac{516}{560} = \frac{129}{140}.$$

To compute $f(5)$, we will again examine the possible slopes of lines through at least two points in the grid. In a 5×5 grid, the possible rises of a line that is neither vertical nor horizontal are 1, 2, 3, and 4, and the possible runs are the same. The possible slopes coming from these rises and runs are

$$\pm 4\,, \pm 3\,, \pm 2\,, \pm\frac{3}{2}\,, \pm\frac{4}{3}\,, \pm 1\,, \frac{3}{4}\,, \pm\frac{2}{3}\,, \pm\frac{1}{2}\,, \pm\frac{1}{3}\,, \pm\frac{1}{4}$$

Using reasoning similar to the case for $n = 4$, it can be shown that of the slopes listed above, only a line with slope ± 1, $\pm\frac{1}{2}$, or ± 2 can pass through three points in a 5×5 grid. By examining the lines having these slopes, as well as the vertical and horizontal lines, we can compute $f(5)$.

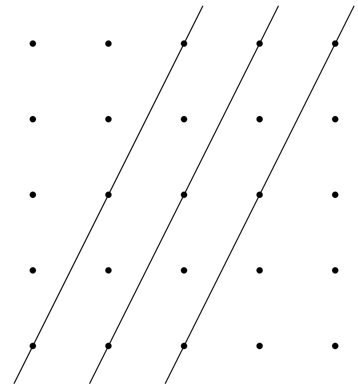
There are five lines of slope 1 that pass through at least three points: a diagonal, two lines above it, and two lines below it. These five lines are depicted below.



Of these five lines, there are two that pass through exactly three points in the grid. Thus, each of these two lines contributes one set of three collinear points. As argued in the $n = 4$ case, the two lines through four points each contribute four sets of three collinear points. The one line through five points contributes $\binom{5}{3} = 10$ sets of three collinear points. Thus, from the lines of slope 1, we get a total of $1 + 1 + 4 + 4 + 10 = 20$ sets of there collinear points. Similar reasoning can be used to show that there are 20 sets of three collinear points on the lines of slope -1.

The counts for each of the slopes $\pm\frac{1}{2}$ and ± 2 are essentially the same, so we will only explicitly examine the case when the slope is 2.

Suppose points A, B, and C are three collinear points in the grid so that the slope of the line through the points is 2. Also suppose that A and C are the two points that are farthest apart. Since the line is neither horizontal nor vertical, A, B, and C are in different columns and so the horizontal distance between A and C is at least 2. Since the slope is 2, this means the vertical distance between A and C is at least 4. However, 4 is the largest possible vertical distance between two points in the grid, so this means the vertical distance between A and C is exactly 4. Therefore, one of A and C must be in the bottom row. There are only five lines of slope 2 passing through one of the points in the bottom row, and only three of them pass through at least three points in the grid. In fact, each of these three lines passes through exactly three points in the grid, and they are depicted below:



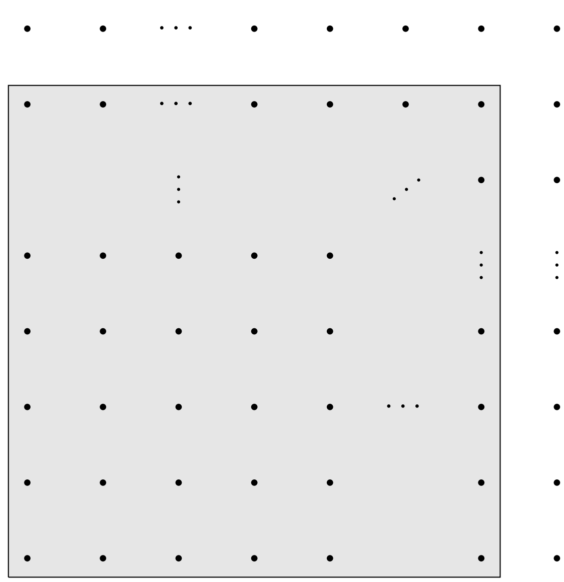
This means that we get an additional three sets of three collinear points from the lines of slope 2. Similarly, we get three sets from the lines of slopes -2, $\frac{1}{2}$, and $-\frac{1}{2}$.

Each vertical line and each horizontal line contains another $\binom{5}{3} = 10$ sets of three collinear points, so we get another $10 \times 10 = 100$ sets of collinear points. Combining with the earlier counts, have that $f(5) = 20 + 20 + 4 \times 3 + 100 = 152$.

The total number of sets of three points in a 5×5 grid is $\binom{25}{3} = 2300$, so

$$P(5) = \frac{2300 - 152}{2300} = \frac{2148}{2300} = \frac{537}{575}.$$

(b) We can consider an $(n+1) \times (n+1)$ grid as an $n \times n$ grid with $2n+1$ additional points:



We will denote by A the set of points in the highlighted $n \times n$ grid (the bottom left $n \times n$ grid) and by B the remaining $2n+1$ points. Within B, we will refer to the $n+1$ points in the rightmost column as the "vertical part" and the $n+1$ points in the top row as the "horizontal part". Note that the top right point is in both the vertical part and the horizontal part of B.

There are four possibilities for a set of three collinear points in the $(n+1) \times (n+1)$ grid: They are all in A, two are in A and one is in B, one is in A and two are in B, or they are all in B. $f(n+1)$ is the sum of the number of sets of collinear points in each case. Remember that our goal is not to compute $f(n+1)$ precisely, but to show that it is smaller than $f(n) + 5n^4 + 5n^3 + 5n^2 + 5n$.

Case 1: All three points are in A. By the definition of $f(n)$, there are exactly $f(n)$ sets of three collinear points in A.

Case 2: Two points are in A and one is in B. For any two points in A, the line defined by the two points intersects B at most once in the vertical part and at most once in the horizontal part. This means for any two distinct points in A, the line through them intersects B at most twice in total. Thus, the number of sets of three collinear points in this case is *no more than* two times the number of ways to choose two distinct points from A. There are n^2 points in A, so there are $\binom{n^2}{2} = \frac{n^2(n^2-1)}{2}$ pairs of distinct points in A. In this case, there are at most $2 \times \frac{n^2(n^2-1)}{2} = n^4 - n^2$ sets of three distinct points.

Case 3: One point is in A and two points are in B. The line defined by two points in the horizontal part of B does not contain any points in A. Similarly, the line defined by any two points in the vertical part of B does not contain any points in A. Therefore, for a set of three collinear points to fall in this case, we must have one of the two points in B in the vertical part and one of the two points in the horizontal part. Neither of these two points can be the top-right point since this would make the line either vertical or horizontal. Therefore, there are $n \times n$ possible ways to choose the two points from B. The line defined by these two points may contain no points from A or could contain several. Since we do not need to count precisely, it will be sufficient to observe that the line intersects each column at most once. Thus, for any of the n^2 pairs of points from B described above, there are at most n points from A on that line. Therefore, there are at most n^3 sets of three collinear points in this case.

Case 4: All three points are in B. Since there are three points, either two of the points are in the vertical part or two of the points are in the horizontal part. Two points define a line, so this means the line must be either horizontal or vertical. Thus, in fact, either all three points are in the vertical part or all three points are in the horizontal part.

There are $n+1$ points in the horizontal part and $n+1$ points in the vertical part (the top-right point is in both parts). Thus, there are $2\binom{n+1}{3} = 2\frac{(n+1)(n)(n-1)}{6} = \frac{n^3-n}{3}$ sets of three collinear points in this case.

We have shown that there are $f(n)$ sets of three collinear points in Case 1 and that there are $\frac{n^3-n}{3}$ sets of three collinear points in Case 4. As well, we showed that there are at most $n^4 - n^2$ sets of three collinear points in Case 2 and that there are at most n^3 sets of three collinear points in Case 3. In Cases 2 and 3, you might have noticed that we were not very careful about our counting. For instance, in Case 3, it is not difficult to show that the line defined by the two points in B intersects A at most $n-2$ times (a better bound than n) and in fact, likely intersects A even fewer times. However, these rather simple bounds will be sufficient.

Putting the information from Cases 1 through 4 together, we have

$$f(n+1) \le f(n) + (n^4 - n^2) + n^3 + \frac{n^3 - n}{3} = f(n) + n^4 + \frac{4}{3}n^3 - n^2 - \frac{1}{3}n.$$

Since n is a positive integer, $n^4 < 5n^4$, $\frac{4}{3}n^3 < 5n^3$, $-n^2 < 5n^2$, and $-\frac{1}{3}n < 5n$. We can apply these four inequalities to get

$$f(n+1) < f(n) + 5n^4 + 5n^3 + 5n^2 + 5n.$$

Note: The last step in the solution to (b) might seem like a strange thing to do since we already had a "better" bound on $f(n+1)$. It might seem like we threw away information, and in fact, we did. Roughly speaking, we have sacrificed some accuracy in order to get an expression that will be easier to work with. In general, when mathematicians use this kind of technique, it can be quite delicate trying to balance how much accuracy can be sacrificed while making the quantities and expressions involved easy enough to manage. It often involves going back to earlier parts of a solution several times to make an adjustment. Indeed, when we were writing this problem, part (c) was the last to be finalized. This is because, after writing a solution to part (d), we went back to adjust what we asked for in part (c). You may be able to solve part (d) by using different versions of parts (b) and (c).

(c) Continuing with our estimation in part (b), we have for $n \ge 3$ that $5n^2 \le 10n^2$ and $5n^3 \le 10n^3$. Thus, for all integers $n \ge 3$, we actually have that

$$\begin{aligned} f(n+1) &< f(n) + 5n^4 + 10n^3 + 10n^2 + 5n \\ &< f(n) + 5n^4 + 10n^3 + 10n^2 + 5n + 1 \end{aligned}$$

where the addition of 1 at the end will be used shortly.

Using the calculations in part (a), we get that $f(3) = 8 < 243 = 3^5$, $f(4) = 44 < 1024 = 4^5$, and $f(5) = 152 < 3125 = 5^5$, so $f(n) < n^5$ for each of $n = 3$, $n = 4$, and $n = 5$. We will now use induction to prove that $f(n) < n^5$ for all $n \ge 3$.

To do this, we will assume that $f(k) < k^5$ for some integer $k \ge 3$ and from this deduce that $f(k+1) < (k+1)^5$. Expanding $(k+1)^5$, we get $(k+1)^5 = k^5 + 5k^4 + 10k^3 + 10k^2 + 5k + 1$. Thus, using the inequality above and the assumption that $f(k) < k^5$, we have

$$\begin{aligned} f(k+1) &< f(k) + 5k^4 + 10k^3 + 10k^2 + 5k + 1 \\ &< k^5 + 5k^4 + 10k^3 + 10k^2 + 5k + 1 \\ &= (k+1)^5 \end{aligned}$$

Therefore, if the statement "$f(k) < k^5$" is true for some integer, then it is true for the next integer. By the principle of mathematical induction, $f(n) < n^5$ for all $n \ge 3$.

Note: It might be a little clearer now why we used such a "weak" bound in part (b). The calculation above was very easy because we bounded $f(n+1)$ by $f(n)$ plus part of the expression $(n+1)^5$.

(d) In an $n \times n$ grid, the number of ways to choose three points is $\binom{n^2}{3} = \dfrac{n^2(n^2-1)(n^2-2)}{6}$.

If we call this quantity $g(n)$, then

$$P(n) = \frac{g(n) - f(n)}{g(n)} = 1 - \frac{f(n)}{g(n)} = 1 - \frac{6f(n)}{n^2(n^2-1)(n^2-2)}.$$

We will find a constant c so that

$$\frac{6f(n)}{n^2(n^2-1)(n^2-2)} < \frac{c}{n}$$

and hence

$$-\frac{6f(n)}{n^2(n^2-1)(n^2-2)} > -\frac{c}{n}$$

which will mean that

$$P(n) = 1 - \frac{6f(n)}{n^2(n^2-1)(n^2-2)} > 1 - \frac{c}{n}$$

for $n \geq 3$.

From part (c), we have that $f(n) < n^5$ for all $n \geq 3$, so we get

$$\frac{6f(n)}{n^2(n^2-1)(n^2-2)} < \frac{6n^5}{n^2(n^2-1)(n^2-2)} = \frac{6n^3}{(n^2-1)(n^2-2)}. \qquad (*)$$

Now notice that for $n \geq 3$, we have $n^2 - 1 > \frac{1}{2}n^2$. To see this, observe that $n^2 > 2$ when $n \geq 3$, so $2n^2 - n^2 > 2$ which can be rearranged to $2n^2 - 2 > n^2$. Multiplying through by $\frac{1}{2}$ gives the result.

In a similar way, it can be argued that $n^2 - 2 > \frac{1}{2}n^2$ when $n \geq 3$.

Since $n^2 - 1 > \frac{1}{2}n^2$ and $n^2 - 2 > \frac{1}{2}n^2$, we have $\frac{1}{n^2-1} < \frac{2}{n^2}$ and $\frac{1}{n^2-2} < \frac{2}{n^2}$. Therefore,

$$\begin{aligned}
\frac{6n^3}{(n^2-1)(n^2-2)} &= 6n^3 \times \frac{1}{n^2-1} \times \frac{1}{n^2-2} \\
&< 6n^3 \times \frac{2}{n^2} \times \frac{2}{n^2} \\
&= \frac{24}{n}
\end{aligned}$$

Combining this with $(*)$, we have

$$\frac{6f(n)}{n^2(n^2-1)(n^2-2)} < \frac{24}{n}$$

for all $n \geq 3$. Thus,

$$P(n) = 1 - \frac{6f(n)}{n^2(n^2-1)(n^2-2)} > 1 - \frac{24}{n}.$$

This means we can take $c = 24$. In fact, any c larger than 24 will work.

Finally, we discuss what this means for large n. Suppose, for example, that $n = 24 \times 10^6$. Then the above inequality implies

$$P(n) > 1 - \frac{24}{24 \times 10^6} = 0.999999.$$

So for this value of n, the probability that the three chosen points form a triangle of positive area is *at least* 0.999999, which means it is very close to 1. As n gets even larger, the quantity $\dfrac{24}{n}$ gets even smaller, so $P(n)$ is forced to be even closer to 1. While this inequality never tells us the exact probability, it does give us a useful estimate.

It is worth pointing out that with $n = 3$, this inequality tells us that $P(n) > 1 - \dfrac{24}{3}$ so $P(n) > -7$. This is true, but it isn't very interesting since probabilities are always positive. This is typical of the type of argument that we have used. The result – that $P(n)$ is very close to 1 when n is large – is not meant to enlighten us for small values of n. The sacrifice of accuracy discussed at the end of the solution to (b) is apparent for small n, but insignificant for large n, as long as all we need to know is that for large n, the probability $P(n)$ is very close to 1. Of course, if we wanted to know something else about $P(n)$, this estimate might not be as useful.

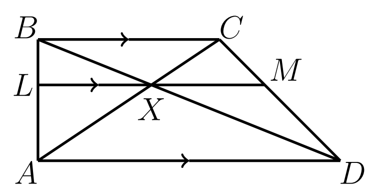
As you might expect, if we had been more careful with our estimates, we could have obtained a result that says more about $P(n)$ (and perhaps works better for some smaller values of n). This would likely come with a harder proof. For example, do you think it is possible to find a constant d so that $P(n) > 1 - \dfrac{d}{n^2}$ for all *sufficiently large* n? If so, it would give an even better understanding of how $P(n)$ behaves for large n, but would require much more careful estimates than those in part (b).

Problem of the Month

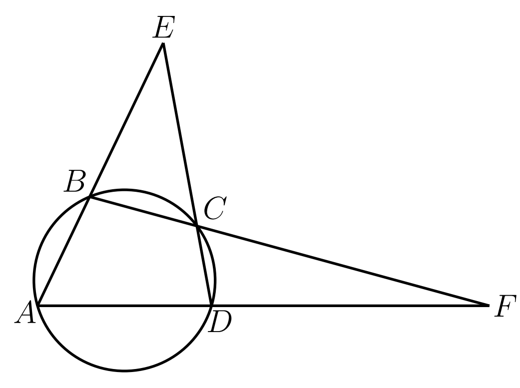
Problem 8: May 2021

Problem Try these three geometry problems! Problems (a), (b), and (c) are not intended to be related to each other. In each part, a diagram is provided to give an example of a figure satisfying the conditions to be explored in that part.

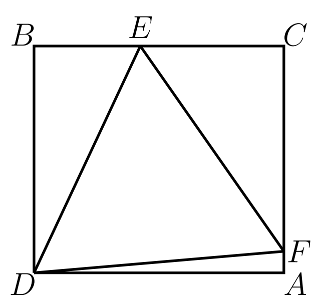
(a) In trapezoid $ABCD$, $\angle BAD = 90^\circ$ and BC is parallel to AD with $BC < AD$. The diagonals AC and BD intersect at point X. A line parallel to AD is drawn through X and intersects AB at L and CD and M. Determine the length of LM in terms of the lengths of BC and AD.



(b) Suppose quadrilateral $ABCD$ has no pair of parallel sides and is inscribed in a circle. AB and DC are extended to meet at point E and AD and BC are extended to meet at point F. The degree measures of $\angle AFB$, $\angle AED$, and $\angle EAF$ form an increasing arithmetic sequence in that order. The degree measure of each of these three angles is an integer. Find all possible values of $\angle AFB$.



(c) Rectangle $DBCA$ has E on BC and F on AC so that $\triangle DEF$ is equilateral. Find all possible values of $\dfrac{BD}{AD}$.



Hint

(a) The assumption that $\angle BAD = 90^\circ$ is not needed, but it might make a few calculations or observations easier. Try to find some similar triangles. Keep in mind that corresponding altitudes of similar triangles are in the same ratio as their sides.

(b) A quadrilateral that can be inscribed in a circle, such as $ABCD$, is called a *cyclic quadrilateral*. It might be useful to look up a few facts about cyclic quadrilaterals. Using these facts, try to shown that $\angle AFB + \angle AED + 2\angle BAD = 180^\circ$.

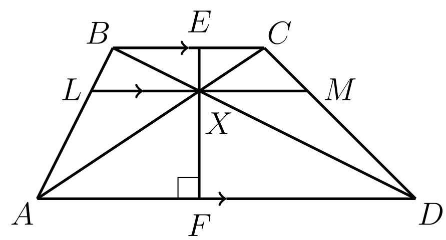
(c) $\triangle DEF$ being equilateral tells us that $\angle EDF = 60^\circ$ and $DE = DF$. Trigonometry might be useful in this problem. Since the question asks for the ratio $\dfrac{BD}{AD}$, you can assume that $AD = 1$ (or has some other fixed value) and explore the possible values of BD.

Problem of the Month

Solution to Problem 8: May 2021

(a) The length LM is equal to the *harmonic mean* of AD and BC, or $\dfrac{2(AD)(BC)}{AD+BC}$. As mentioned in the hint, the length of LM does not change even if $\angle BAD \neq 90^\circ$, so the solution that follows will not assume $\angle BAD = 90^\circ$. The solution given assumes that both $\angle BAD$ and $\angle CDA$ measure at most 90°. A very similar argument can be used to prove the result if one of these angles is obtuse.

Begin by drawing a line through X that is perpendicular to AD, intersecting BC at E and AD at F.



Since BC and AD are parallel, $\angle BCA = \angle DAC$ which means $\angle BCX = \angle DAX$. We also have that $\angle AXD = \angle CXB$ since they are opposite angles. Therefore, $\triangle BCX$ is similar to $\triangle DAX$ by angle-angle similarity.

Since $\triangle BCX$ is similar to $\triangle DAX$, the corresponding altitudes EX and FX are in the same ratio as corresponding sides. In particular, $\dfrac{EX}{FX} = \dfrac{BC}{AD}$ which can be rearranged to $FX = \dfrac{(AD)(EX)}{BC}$.

Since EF is perpendicular to AD and BC, it is perpendicular to LX. Therefore, the height of $\triangle BLX$ is the length of EX and the height of $\triangle BAD$ is the length of EF. From LX and AD being parallel, we also get that $\angle BLX = \angle BAD$ and $\angle BXL = \angle BDA$. Therefore, $\triangle BLX$ and $\triangle BAD$ are similar. This means $\dfrac{LX}{AD} = \dfrac{EX}{EF}$. Rearranging, we have $LX = \dfrac{(AD)(EX)}{EF}$.

Using that $EF = EX + FX$, we can now compute LX in terms of AD and BC:

$$\begin{aligned}
LX = \frac{(AD)(EX)}{EF} &= \frac{(AD)(EX)}{EX + \dfrac{(AD)(EX)}{BC}} \\
&= \frac{(AD)(EX)(BC)}{(EX)(BC) + (AD)(EX)} \\
&= \frac{EX}{EX} \cdot \frac{(AD)(BC)}{BC + AD} \\
&= \frac{(AD)(BC)}{AD + BC}
\end{aligned}$$

A very similar calculation shows that $XM = \dfrac{(AD)(BC)}{AD+BC}$. It is also possible to see this by showing directly that $LX = XM$. Therefore,

$$LM = LX + XM = \frac{2(AD)(BC)}{AD+BC}.$$

(b) Let $\angle AFB = x^\circ$, $\angle AED = y^\circ$, and $\angle EAF = z^\circ$. Since $ABCD$ is a cyclic quadrilateral, its opposite angles are supplementary. This means $\angle BCD + \angle BAD = 180^\circ$, and so $\angle BCD = 180^\circ - z^\circ$. It is also true that $\angle ABC + \angle ADC = 180^\circ$, but we will not use this directly.

Since BCF is a straight line, $\angle FCD = 180^\circ - \angle BCD = 180^\circ - (180^\circ - z^\circ) = z^\circ$. Since they are opposite angles, $\angle ECB = \angle FCD = z^\circ$ as well.

The angles in $\triangle CDF$ add up to 180°, which means $\angle CDF = 180^\circ - x^\circ - z^\circ$, and hence $\angle ADE = x^\circ + z^\circ$ since ADF is a straight line. Also, in $\triangle EAD$ we have

$$\begin{aligned} 180^\circ &= \angle EAD + \angle AED + \angle ADE \\ &= z^\circ + y^\circ + (x^\circ + z^\circ) \end{aligned}$$

which means $x + y + 2z = 180$.

We are also assuming that x, y, and z are integers and that x, y, z is an increasing arithmetic sequence. This means x is a positive integer and there is a positive integer d such that $y = x + d$ and $z = x + 2d$. Substituting these expressions into $x + y + 2z = 180$ gives $x + (x + d) + 2(x + 2d) = 180$ or $4x + 5d = 180$.

Rearranging this equation to $5d = 180 - 4x = 4(45 - x)$, it must be the case that $4(45 - x)$ is a multiple of 5, which means $45 - x$ is a multiple of 5 since 5 is prime and 4 is not a multiple of 5. It follows that x must be a multiple of 5. The possible values for x and the corresponding values of d are summarized in the table below:

x	5	10	15	20	25	30	35	40
d	32	28	24	20	16	12	8	4

Note that we cannot have $x \geq 45$ since this this would make $5d = 4(45 - x) \leq 0$ which cannot happen since $d > 0$.

So far, we have shown that if the figure in the problem statement satisfies the given conditions, then x must take one of the eight values in the table above. We will finish the argument by discussing the fact that each of the above values of x actually occurs.

For $x = 5$, we construct quadrilateral $ABCD$ so that $\angle ABC = 106^\circ$, $\angle BCD = 111^\circ$, $\angle CDA = 74^\circ$, and $\angle DAB = 69^\circ$. This quadrilateral is cyclic since

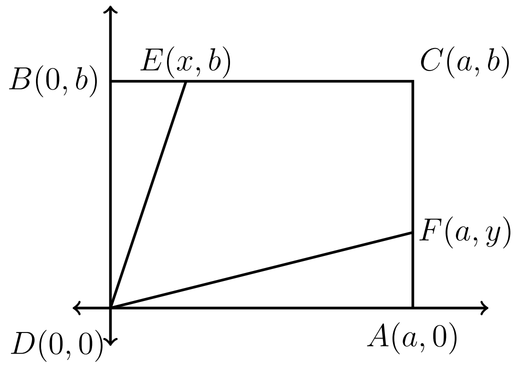
$$\angle ABC + \angle CDA = \angle BCD + \angle DAB = 180^\circ.$$

As well, if AB and DC are extended to meet at E, then $\triangle AED$ has angles measuring 69° and 74°, so the third angle $\angle AED = 180^\circ - 69^\circ - 74^\circ = 37^\circ$. Similarly, if AD and BC are extended to meet at F, then $\angle AFB = 180^\circ - 69^\circ - 106^\circ = 5^\circ$. Thus, with $ABCD$ constructed with the angles above, we get $\angle AFB = 5^\circ$, $\angle AED = 37^\circ$, as well as $\angle BAD = \angle EAF = 69^\circ$, so the degree measures of $\angle AFB$, $\angle AED$, and $\angle EAF$ form an increasing arithmetic sequence with integer values. Furthermore $\angle AFB = 5^\circ$, as desired.

Similar constructions can be used to show that the other seven values of x in the table above can be achieved.

It is also true that if x, y, and z are any positive numbers satisfying $x + y + 2z = 180$ (whether or not they are integers and form an arithmetic sequence), then it is possible to construct a quadrilateral $ABCD$ so that AB and DC can be extended to meet at E and AD and BC can be extended to meet at F in such a way that $\angle AFB = x^\circ$, $\angle AED = y^\circ$, and $\angle EAF = z^\circ$. You might want to try to prove this!

(c) Suppose the vertices of $DBCA$ are $D(0,0)$, $B(0,b)$, $C(a,b)$, and $A(a,0)$. Now let $E(x,b)$ be on BC and $F(a,y)$ be on AC.



Let $\alpha = \angle EDA$ and $\beta = \angle FDA$. We want to find conditions under which $\alpha - \beta = 60^\circ$ and $ED = FD$. This is because a triangle is equilateral exactly when it is isosceles with the equal sides meeting at a 60° angle.

Applying the Pythagorean theorem, we have that $ED = \sqrt{b^2 + x^2}$ and $DF = \sqrt{a^2 + y^2}$. Thus, we suppose $\sin(\alpha - \beta) = \sin 60^\circ$ and that $\sqrt{b^2 + x^2} = \sqrt{a^2 + y^2}$. Notice that BC and DA are parallel, so $\angle BED = \angle EDA$. Therefore, we can use right $\triangle BED$ to compute $\sin\alpha$ and $\cos\alpha$. Since $\sin 60^\circ = \frac{\sqrt{3}}{2}$, we have

$$\begin{aligned}
\frac{\sqrt{3}}{2} &= \sin(\alpha - \beta) \\
&= \sin\alpha\cos\beta - \cos\alpha\sin\beta \\
&= \frac{b}{\sqrt{b^2 + x^2}}\frac{a}{\sqrt{a^2 + y^2}} - \frac{x}{\sqrt{b^2 + x^2}}\frac{y}{\sqrt{a^2 + y^2}} \\
&= \frac{ab - xy}{b^2 + x^2}
\end{aligned}$$

where the final equality is because $\sqrt{b^2 + x^2} = \sqrt{a^2 + y^2}$. Rearranging this equation, we have $\sqrt{3}(b^2 + x^2) = 2(ab - xy)$. In a similar way, using that $\cos 60^\circ = \frac{1}{2}$, we get

$$\begin{aligned}
\frac{1}{2} &= \cos(\alpha - \beta) \\
&= \frac{ax + by}{b^2 + x^2}
\end{aligned}$$

which implies $b^2 + x^2 = 2(ax + by)$. We are interested in the quantity $\frac{b}{a}$, which we will denote by T. Similarly, set $X = \frac{x}{a}$ and $Y = \frac{y}{a}$. Dividing $\sqrt{3}(b^2 + x^2) = 2(ab - xy)$

by a^2 gives $\sqrt{3}(T^2+X^2) = 2(T-XY)$ and dividing $b^2+x^2 = 2(ax+by)$ by a^2 gives $T^2+X^2 = 2(X+TY)$. We now have two equations implied by the assumptions that $\angle EDF = \alpha - \beta = 60^\circ$ and $DE = DF$:

$$\sqrt{3}(T^2+X^2) = 2T - 2XY \tag{1}$$
$$T^2+X^2 = 2X + 2TY. \tag{2}$$

Multiplying Equation (1) by T and multiplying Equation (2) by X gives

$$\sqrt{3}T(T^2+X^2) = 2T^2 - 2TXY$$
$$X(T^2+X^2) = 2X^2 + 2TXY$$

then adding these equations leads to

$$(\sqrt{3}T + X)(T^2+X^2) = 2T^2 + 2X^2.$$

The quantities a and b are positive, so $T = \frac{b}{a}$ is positive, which means T^2+X^2 is positive, and hence, it is nonzero. Dividing by T^2+X^2, we get that $\sqrt{3}T + X = 2$, so $X = 2 - \sqrt{3}T$.

There are several ways to solve for Y in terms of T. We will return to the equations $\sqrt{3}(b^2+x^2) = 2(ab-xy)$ and $b^2+x^2 = 2(ax+by)$. Using that $b^2+x^2 = a^2+y^2$, we get $\sqrt{3}(a^2+y^2) = 2(ab-xy)$ and $a^2+y^2 = 2(ax+by)$ which can be divided by a^2 to get

$$\sqrt{3}(1+Y^2) = 2T - 2XY \tag{3}$$
$$1+Y^2 = 2X + 2TY \tag{4}$$

Multiplying Equation (4) by Y gives $Y(1+Y^2) = 2XY + 2TY^2$, and if we add this to Equation (3), we get

$$(\sqrt{3}+Y)(1+Y^2) = 2T(1+Y^2).$$

The quantity $1+Y^2$ is nonzero because it is positive, so we can divide both sides of the equation by $1+Y^2$ to get $\sqrt{3}+Y = 2T$ or $Y = 2T - \sqrt{3}$.

Observe that $0 < x < a$ and $0 < y < b$, so we can divide these two inequalities by a to get $0 < X < 1$ and $0 < Y < T$ respectively. From $0 < X < 1$ and $X = 2 - \sqrt{3}T$, we have $0 < 2 - \sqrt{3}T < 1$. Rearranging $0 < 2 - \sqrt{3}T$ gives $T < \frac{2}{\sqrt{3}}$. Rearranging $2 - \sqrt{3}T < 1$ gives $\frac{1}{\sqrt{3}} < T$, so we have

$$\frac{1}{\sqrt{3}} < T < \frac{2}{\sqrt{3}}.$$

Using $0 < Y < T$ and $Y = 2T - \sqrt{3}$, we have $0 < 2T - \sqrt{3} < T$. Rearranging $0 < 2T - \sqrt{3}$ gives $\frac{\sqrt{3}}{2} < T$, and rearranging $2T - \sqrt{3} < T$ gives $T < \sqrt{3}$. Therefore, we also have

$$\frac{\sqrt{3}}{2} < T < \sqrt{3}.$$

We know that $\frac{1}{\sqrt{3}} < T$ and $\frac{\sqrt{3}}{2} < T$. However, one can check that $\frac{1}{\sqrt{3}} < \frac{\sqrt{3}}{2}$, so the condition $\frac{1}{\sqrt{3}} < T$ is redundant. Similarly, $\frac{2}{\sqrt{3}} < \sqrt{3}$, so the condition $T < \sqrt{3}$ is redundant. We conclude that

$$\frac{\sqrt{3}}{2} < T < \frac{2}{\sqrt{3}}$$

and so we have that the quantity $T = \dfrac{BD}{AD}$ must lie in the interval $\left(\dfrac{\sqrt{3}}{2}, \dfrac{2}{\sqrt{3}}\right)$.

To finish the argument, we will show that if $\dfrac{BD}{AD}$ is in the interval $\left(\dfrac{\sqrt{3}}{2}, \dfrac{2}{\sqrt{3}}\right)$, then there are points E on BC and F on AC so that $\triangle DEF$ is equilateral. To that end, suppose $DBCA$ is a rectangle with $BD = b$ $AD = a$. Choose E on BC so that $BE = 2a - \sqrt{3}b$ and choose F on AC so that $AF = 2b - \sqrt{3}a$.

Notice that $2a - \sqrt{3}b > 0$ since this inequality can be obtained by rearranging $\dfrac{b}{a} < \dfrac{2}{\sqrt{3}}$ which is what we are assuming. Thus, BE is positive and can be chosen as a length. As well, $BC - BE = a - (2a - \sqrt{3}b) = \sqrt{3}b - a$ which is positive. This can be seen by rearranging $\dfrac{b}{a} > \dfrac{1}{\sqrt{3}}$, which is true since $\dfrac{b}{a} > \dfrac{\sqrt{3}}{2} > \dfrac{1}{\sqrt{3}}$.

Thus, it is indeed possible to choose E on BC so that $BE = 2a - \sqrt{3}b$. Similar arguments can be used to verify that it is possible to choose F on AC so that $AF = 2b - \sqrt{3}a$.

By the Pythagorean theorem, we have

$$\begin{aligned} DE &= \sqrt{b^2 + \left(2a - \sqrt{3}b\right)^2} \\ &= \sqrt{b^2 + 4a^2 - 4\sqrt{3}ab + 3b^2} \\ &= 2\sqrt{a^2 - \sqrt{3}ab + b^2} \end{aligned}$$

and similarly

$$\begin{aligned} DF &= \sqrt{a^2 + \left(2b - \sqrt{3}a\right)^2} \\ &= \sqrt{a^2 + 4b^2 - 4\sqrt{3}ab + 3a^2} \\ &= 2\sqrt{a^2 - \sqrt{3}ab + b^2} \\ &= DE \end{aligned}$$

which shows that $\triangle DEF$ is isosceles. To show that $\angle EDF = 60^\circ$, we will show that $\cos \angle EDF = \dfrac{1}{2}$. This is sufficient since $\angle EDF$ is an angle in a triangle, and 60° is the only angle between 0° and 180° with its cosine equal to $\dfrac{1}{2}$. Using right-angle trigonometry as well as the formula for cosine of a difference, we have

$$\begin{aligned}
\cos\angle EDF &= \cos(\angle EDA - \angle FDA)\\
&= \cos\angle EDA \cos\angle FDA + \sin\angle EDA \sin\angle FDA\\
&= \left(\frac{BE}{DE}\right)\left(\frac{DA}{DF}\right) + \left(\frac{BD}{DE}\right)\left(\frac{AF}{DF}\right)\\
&= \frac{(2a-\sqrt{3}b)(a)}{(DE)(DF)} + \frac{b(2b-\sqrt{3}a)}{(DE)(DF)}\\
&= \frac{2a^2-\sqrt{3}ab+2b^2-\sqrt{3}ab}{4(a^2-\sqrt{3}ab+b^2)}\\
&= \frac{2(a^2-\sqrt{3}ab+b^2)}{4(a^2-\sqrt{3}ab+b^2)}\\
&= \frac{1}{2}.
\end{aligned}$$

The cancellation in the last line is allowed because the denominator is the product of two side lengths of a triangle and is hence nonzero.

We have now shown that if $\dfrac{BD}{AD}$ is between $\dfrac{\sqrt{3}}{2}$ and $\dfrac{2}{\sqrt{3}}$, then there is a way to choose points E and F on BC and AC, respectively, so that $\triangle DEF$ is equilateral.

Remark: If you assumed that $AD = 1$, then the correct answer to this problem is that $\dfrac{\sqrt{3}}{2} < BD < \dfrac{2}{\sqrt{3}}$. As well, we have assumed in this solution that E does not coincide with B or C and that F does not coincide with A or C. You might want to think about what happens if this is allowed and how it fits with what has been shown.