WISCONSIN MATHEMATICS, SCIENCE & ENGINEERING TALENT SEARCH

SOLUTIONS TO PROBLEM SET IV (2021-2022)

1. We put a green and a yellow marble in a bag. We choose one of them uniformly at random, and return it into the bag with an additional marble that has the same color as the chosen one. We repeat this procedure five more times, after which we have 8 marbles in the bag, some of them green and some of them yellow. What is the probability that there are four green and four yellow marbles in the bag at this point?

SOLUTION. We will show that the answer is $1/7$.

Let us call the act of choosing a marble randomly from the bag, and returning it together with an additional marble of the same color a *step*. After each step the number of marbles in the bag increases by one: so after step k we have $k+2$ marbles. If there are g green and y yellow marbles in the bag, then after the next step we have either $g+1$ green and y yellow or g green and $y+1$ yellow, depending on the color of the chosen marble. After k steps the number of green marbles in the bag can be any number between 1 and $k+1$ (since there are $k+2$ marbles all together, and we have to have at least one green and one yellow). We will show that these possibilities all have the same probability, in other words the probability of having $1 \le g \le k+1$ green marbles in the bag after k steps is always $\frac{1}{k+1}$. This will imply that after 6 steps the probability of having the same number of green and yellow marbles is $\frac{1}{7}$.

We will prove this statement by induction on the number of steps. Imagine that we label the initial green and yellow marbles by 1 and 2, respectively, and we label the additional marble added after step k by $k+2$. This way each time we pick a marble we just have to pick a number from the available labels: at step k these are the numbers from 1 to $k+1$. If we record the picked labels, then we also know the colors of the picked marbles: the first picked label tells us which color we picked first, and what the color of marble #3 is, then the second label tells us the color of the second picked marble and the color of marble #4, and so on. For example, the sequence $(1,3)$ means that we picked the green marble in the first step, hence the marble labeled with 3 is also green, and then we picked marble #3, which is green, so marble #4 is also green. After these series of picks we have three green marbles (labeled 1, 3, 4) and a yellow marble (labeled 2) in the bag.

The result of the first k steps is a sequence of k numbers where the ℓth number is between 1 and $\ell+1$. There are $2 \cdot 3 \cdot \ldots \cdot (k+1) = (k+1)!$ such sequences (since there are always 2 possible choices in the first step, 3 in the second step, and so on). Since in each step we choose a label uniformly at random from the available numbers, the sequence encoding the first k steps is also uniformly chosen from all possible allowed length k sequences. This means that any particular allowable sequence of length k has a probability of $\frac{1}{(k+1)!}$ to appear. This setup allows us to compute probabilities by counting outcomes: if we want to know the probability of seeing g green and y yellow marbles after k steps, then we need to count how many length k sequences produce that many green and yellow marbles, and divide this number by $(k+1)!$.

Let us return to the proof of the statement above. In the first step we choose a green marble with probability $1/2$ so the probability of having 2 green balls in the bag after the first step is also $1/2$. This also shows that the probability of having 1 green ball in the bag after the first step is also $1/2$. Hence the statement is true for $k=1$. Now assume that it is true for some $k \ge 1$, we will show that it will also be true for $k+1$. By our induction hypothesis after k steps the probability of seeing $1 \le g \le k+1$ green marbles is $\frac{1}{k+1}$, which means that for each particular g there are $\frac{(k+1)!}{k+1} = k!$ equally likely sequences of length k producing exactly g green marbles. In order to get

g green marbles after $k+1$ steps, we had to have $g-1$ or g green marbles in the bag after k steps, and then had to choose the marble with the appropriate color in the $(k+1)$st step. We could
– have $g-1$ green marbles in the bag after k steps and then pick one of the $g-1$ green marbles in the $(k+1)$st step (which we can do $k! \cdot (g-1)$ different ways if $g-1 \geq 1$), or
– have g green marbles in the bag after k steps and then pick one of the $k+2-g$ yellow marbles in the $(k+1)$st step (which we can do $k! \cdot (k+2-g)$ different ways if $g \leq k+1$.

This means that
– if $2 \leq g \leq k+1$, then there are $k! \cdot (g-1) + k! \cdot (k+2-g) = k! \cdot (k+1) = (k+1)!$ different sequences of length $k+1$ resulting in g green marbles in the bag after step $k+1$,
– if $g = 1$, then there are $k! \cdot (k+2-1) = (k+1)!$ different sequences of length $k+1$ resulting in g green marbles, and
– if $g = k+2$, then there are $k! \cdot (k+2-1) = (k+1)!$ different sequences of length $k+1$ resulting in g green marbles.

This shows that after $k+1$ steps for each $1 \leq g \leq k+2$, the probability of getting exactly g green marbles in the bag is $\frac{(k+1)!}{(k+2)!} = \frac{1}{k+2}$, which proves the induction step and our statement.

ALTERNATE SOLUTION. In order to see 4 green and 4 yellow marbles in the bag, we need to pick 3 green and 3 yellow marbles during the six steps in some order. One possibility is picking 3 green marbles in steps 1, 2, and 3, and then 3 yellow marbles in steps 4, 5, and 6. Let's count how many of the $2 \cdot 3 \cdots 6 = 6!$ equally likely sequences produce these colors. For the first pick we have to choose the green marble (label 1). For the second pick we can choose the original green marble, or the new one (labels 1 or 3), and the third pick we can choose any of the three green marbles in the bag (labels 1, 3, or 4). After that we have to choose the only yellow marble in the bag (label 2), for the fifth pick we can choose on of the two yellows (labels 2 or 5), and for the sixth pick on of the three yellows (labels 2, 5, or 6). When we are counting these sequences, the number of choices for a given pick is not influenced by the previous picks, so we can just multiply these possibilities together to get $1 \cdot 2 \cdot 3 \cdot 1 \cdot 2 \cdot 3 = (3!)^2$. This is the number of ways we can get the picks g-g-g-y-y-y.

We claim that any other specific ordering of three green and three yellow would correspond to exactly $(3!)^2$ sequences of picks as well. For any fixed ordering of colors (e.g. g-y-g-g-y-y) we can look at the number of possible choices that produce the appropriate color for each pick. When we have to pick the first green marble, then we will have only one green marble in the bag, so there is only one way to do that. When we pick the second green marble, then we already picked green once, so we will have exactly two green marbles in the bag: we will have two possible choices. Similarly, for the third green marble we will have 3 possible choices. The same way we can see that for the first, second, and third yellow marble we will have one, two, and three possible choices, respectively. This shows that for a given ordering of three green and three yellow there are $(3!)^2$ sequences of picks producing that order.

There are $\frac{6 \cdot 5 \cdot 4}{3!}$ ways to order three green and three yellow marbles. One way to see this is by noting that we can put down the first green marble into 6 possible positions, then the second one into 5, the third into 4. This gives $6 \cdot 5 \cdot 4$ ways of placing the three green into one of the 6 positions, but since the three green marbles are indistinguishable, we overcounted each configuration 3! times (this is the number of ways we can reorder the three green marbles once their three positions are fixed). Thus, there are $\frac{6 \cdot 5 \cdot 4}{3!}$ orderings of three green and three yellow, each corresponding to $(3!)^2$ possible sequences of picks, which means that there are $\frac{6 \cdot 5 \cdot 4}{3!} \cdot (3!)^2 = 6 \cdot 5 \cdots 1 = 6!$ sequences of 6 picks resulting in an equal number of green and yellow marbles. Since there are $(6+1)! = 7!$ sequences all together, and each are equally likely, the probability of seeing four green and four yellow marbles in the end is $\frac{6!}{7!} = \frac{1}{7}$.

2. Prove that a circle centered at the point $(\sqrt{2}, \sqrt{3})$ cannot contain more than one point for which both coordinates are integers.

SOLUTION.
Suppose that the circle contains two such points with integer coordinates: (k, ℓ) and (m, n). Then

$$(\sqrt{2} - k)^2 + (\sqrt{3} - \ell)^2 = (\sqrt{2} - m)^2 + (\sqrt{3} - n)^2,$$

because both sides are equal to the square of the radius of the circle. Expanding and moving all squares to the right-hand side and all double products to the left-hand side we get

$$2\sqrt{2}(k - m) + 2\sqrt{3}(\ell - n) = k^2 + \ell^2 - m^2 - n^2,$$

from which we see that the left-hand side must be an integer. Denote that integer by c. Also let $a = k - m$ and $b = \ell - n$. Using the new notations the last equation becomes

$$2\sqrt{2}a + 2\sqrt{3}b = c,$$

where a, b and c are integers. Squaring both sides and moving all the squares to the right-hand side we get

$$8\sqrt{6}ab = c^2 - 12a^2 - 8c^2,$$

which implies that either $ab = 0$ or $\sqrt{6} = (c^2 - 12a^2 - 8c^2)/8ab$ is a rational number (a number is called rational if it can be represented as a ratio of an integer and a positive integer).

If $ab = 0$, then either $a = k - m = 0$ or $b = m - \ell = 0$. In either case, the very first equation implies that both $k = m$ and $\ell = n$. It remains to show that $\sqrt{6}$ is irrational, that is, it cannot be represented as $\sqrt{6} = p/q$, where p and q are positive integers.

Indeed, let $\sqrt{6} = p/q$ where p, q are positive integers. We can assume that p and q do not have common positive integer divisors other than 1 because otherwise we could cancel that divisor out in the fraction. Let us square the equation and rewrite it as $6q^2 = p^2$. Notice that the left-hand side is an even integer and, therefore, p^2 is an even integer. Since a square of any odd integer is odd, it implies that p must be even, $p = 2r$ for some positive integer r. Hence, $6q^2 = 4r^2$, which implies that $6q^2$ is divisible by 4. Since 6 is only divisible by 2 and not by 4, it follows that q^2 must be divisible by 2, which, as before, implies that q is even. But if p and q are both even, then they have a common divisor 2, and we obtain a contradiction.

(The argument in the previous paragraph is the standard proof of irrationality of $\sqrt{2}$, which is a historic mathematical result allegedly obtained in ancient Greece by Hippasus of Metapontum.)

3. How many integers between 1 and 10^{20} can be represented in the form $\lfloor x^2 \rfloor + \lfloor x \rfloor$ for some positive real number x? (Here $\lfloor x \rfloor$ refers to the greatest integer less than or equal to x, see Problem 2 of Problem Set 2.)

SOLUTION. The answer is $10^{20} - 10^{10}$. First note that whenever $0 < a < b$ that $\lfloor a^2 \rfloor + \lfloor a \rfloor \leq \lfloor b^2 \rfloor + \lfloor b \rfloor$. If $0 < x < 1$, then $\lfloor x^2 \rfloor + \lfloor x \rfloor = 0$. When $x = 1$, the value is $\lfloor x^2 \rfloor + \lfloor x \rfloor = 2$, which shows that for $x > 1$, the value is $\lfloor x^2 \rfloor + \lfloor x \rfloor \geq 2$. This means that $\lfloor x^2 \rfloor + \lfloor x \rfloor$ is never equal to 1.

If k is a positive integer, then $\lfloor k^2 \rfloor + \lfloor k \rfloor = k^2 + k$, and for x in the interval $[k, k+1)$, $\lfloor x \rfloor$ remains constant equal to k, but $\lfloor x^2 \rfloor$ increases from k^2 to $(k+1)^2 - 1$, taking on all the integer values in between, showing that $\lfloor x^2 \rfloor + \lfloor x \rfloor$ takes on all the integer values from $k^2 + k$ to $(k+1)^2 - 1 + k = k^2 + 3k$. Then $\lfloor (k+1)^2 \rfloor + \lfloor k+1 \rfloor = (k+1)^2 + (k+1) = k^2 + 3k + 2$ showing that for no positive real number x is $\lfloor x^2 \rfloor + \lfloor x \rfloor$ equal to $k^2 + 3k + 1$. Therefore, the positive integer values that are 'missed' by the possible values of $\lfloor x^2 \rfloor + \lfloor x \rfloor$ are exactly the numbers of the form $k^2 + 3k + 1$, where k is a nonnegative integer. For example, $\lfloor x^2 \rfloor + \lfloor x \rfloor$ is never equal to $0^2 + 3 \cdot 0 + 1 = 1$, $1^2 + 3 \cdot 1 + 1 = 5$, $2^2 + 3 \cdot 2 + 1 = 11$, and $3^2 + 3 \cdot 3 + 1 = 19$. Thus, the answer is the number of the integers between 1 and 10^{20} are not of the form of $k^2 + 3k + 1$. For each nonnegative integer k

$$(k+1)^2 = k^2 + 2k + 1 \leq k^2 + 3k + 1 < k^2 + 4k + 4 = (k+2)^2,$$

and for $k \geq 1$ the inequality on the left is strict.

Since $10^{20} = (10^{10})^2$, we have the following integers of the form $k^2 + 3k + 1$ up to 10^{20}: the number $0^2 + 3 \cdot 0 + 1 = 1$, and exactly one integer in each of the intervals $(2^2, 3^2)$, $(3^2, 4^2)$, $\ldots$, $((10^{10} - 1)^2, (10^{10})^2)$. This gives $10^{10} - 1$ integers all together that are missed by the possible values of $\lfloor x^2 \rfloor + \lfloor x \rfloor$ between 1 and 10^{20}. Hence, the number of positive integers that are obtained is $10^{20} - 10^{10} + 1$.

4. Two 2021-gons, $P_1P_2 \ldots P_{2021}$ and $Q_1Q_2 \ldots Q_{2021}$, are inscribed in the same circle. It turns out that their corresponding sides are parallel: $P_1P_2 \parallel Q_1Q_2$, $P_2P_3 \parallel Q_2Q_3$, and so on; finally, $P_{2021}P_1 \parallel Q_{2021}Q_1$. Show that the corresponding sides must have equal length: $|P_1P_2| = |Q_1Q_2|, \ldots, |P_{2021}P_1| = |Q_{2021}Q_1|$.

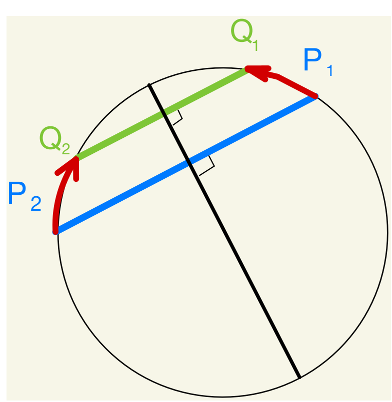
SOLUTION. Consider two points P_1, Q_1 on the circle. There are three possibilities:

Case 1. The two points P_1 and Q_1 coincide. Then, since $P_2P_1 \parallel Q_2Q_1$, we see that P_2 and Q_2 must coincide as well. Repeating the argument, we see that the two polygons are actually the same.

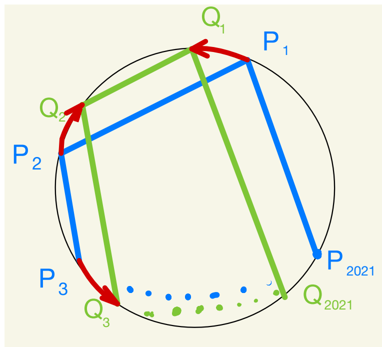
Case 2. The two points P_1 and Q_1 are diametrically opposed (that is, P_1Q_1 is a diameter of the circle). Again, using that $P_2P_1 \parallel Q_2Q_1$, we see that P_2 and Q_2 must be diametrically opposed as well. One way to justify this is to rotate points Q_2 and Q_1 around the center of the circle 180°: if we denote the resulting points Q_2' and Q_1', then Q_1' coincides with P_1, and since $Q_2'Q_1' \parallel Q_2Q_1 \parallel P_2P_1$, we see that Q_2' coincides with P_2 as in the previous case. Hence, $|P_1P_2| = |Q_1Q_2|$. Repeating the argument, we obtain the required statement. (In fact, we see that the two polygons differ by a 180° rotation.)

Case 3. The two points P_1 and Q_1 are neither the same nor diametrically opposed. Let us show that this is actually impossible.

Note first that since P_1 and Q_1 are neither diametrically opposed nor the same, they separate the circle into two non-equal arcs. Denote by $\widehat{P_1Q_1}$ the minor arc (the shorter of the two). Note that the arc has *direction*: Q_1 is either clockwise or counterclockwise from P_1 on $\widehat{P_1Q_1}$.



The key observation is that, when $P_1P_2 \parallel Q_1Q_2$, the direction of $\overparen{P_2Q_2}$ is opposite to that of $\overparen{P_1Q_1}$. (Part of the claim is that P_2 is distinct from Q_2 and not diametrically opposed to it, so that we can talk about the direction of $\overparen{P_2Q_2}$.) The easiest way to see this is to draw the diameter of the circle that is perpendicular to the segment P_1P_2 (and, therefore, also to Q_1Q_2): we then see that P_2 and Q_2 are the mirror reflections of P_1 and Q_1 in this diameter, and therefore, $\overparen{P_2Q_2}$ is the mirror reflection of $\overparen{P_1Q_1}$. Hence, their directions are opposite. (We also see that the two arcs $\overparen{P_1Q_1}$ and $\overparen{P_2Q_2}$ have equal length, but we do not use this fact.)



By the same argument, we see that the direction of $\overparen{P_3Q_3}$ is opposite to the direction of $\overparen{P_2Q_2}$ (and hence $\overparen{P_1Q_1}$ and $\overparen{P_3Q_3}$ have the same direction) and so on. Iterating, we see that all of the 'odd-numbered' arcs $\overparen{P_1Q_1}, \overparen{P_3Q_3}, \ldots, \overparen{P_{2021}Q_{2021}}$ have the same direction, while all of the 'even-numbered' arcs $\overparen{P_2Q_2}, \ldots, \overparen{P_{2020}Q_{2020}}$ have the opposite direction. After one last iteration, we get a contradiction: the direction of $\overparen{P_1Q_1}$ is opposite to itself.

5. There is an invisible virus in one of the cells of a 2022×2022 square grid. John has a scanner that can scan the cells of any 50×50 square inside the grid. He can use the scanner once every minute, and each time he can choose where to scan. If the virus is in the scanned area at the time of a scan, then the virus is caught. However, if the virus is not in the scanned area, then it moves to one of the cells that shares a side with its current cell. (This move happens after each unsuccessful scan, before the next scan is made.) Prove that John can find the virus after finitely many scans, and show how this can be done.

SOLUTION. Let us present the algorithm for John first, and show that it is works afterwards. We index the cells of the square with pairs of numbers (A, B), $1 \le A, B \le 2022$ from left to right and from top to bottom. We will say that John *aims* at the cell (A, B) when he scans the square with corner cells (A, B) and $(A + 49, B + 49)$.

We will proceed by "covering the strips". Let John aim his scanner at the following series of cells $(1, 1)$, $(50, 1)$, $(99, 1)$, $\ldots$, $(1 + 49k, 1)$, $\ldots$, $(1961, 1)$, $(1973, 1)$. Here, the last scan is adjusted from the formula value of $(2010, 1)$ in order to make sure that the entire scanned area fits inside the 2022×2022 square, since it is not explicitly stated in the problem that John is allowed to scan outside of the square (perhaps the scans can be dangerous to the environment outside the square). This requires 42 scans in total.

Next, let John repeat this horizontal series of scans, shifting it down (that is, increasing B) by 7. Thus, John aims at the series of cells $(1, 8)$, $(50, 8)$, $(99, 8)$, $\ldots$, $(1 + 49k, 8)$, $\ldots$, $(1961, 8)$, $(1973, 8)$. He then continues with another series shifted further by 7 (with the value $B = 15$), then with $B = 22$, and so on. The value of B for the series number m will be $7m - 6$.

Finally, after $B = 1968$ (in the 282nd series), let John do another series for $B = 1973$. Once again, we are adjusting the value of B to avoid scanning outside the square. Thus, the entire process requires 283 series of 42 scans each, for $11,886$ scans total (which would take John more than eight days).

Now we show that John will always find the virus using this strategy. We will need to keep track of the possible positions of the virus at different stages of the process, assuming the virus has not been located yet. It is convenient to divide the square into 42 vertical strips of widths 49, ..., 49, 12, 50, in this order. That is to say, the first strip covers the squares with A between 1 and 49, the second with A between 50 and 99, and so on, the forty-first strip is for A between 1961 and 1972, and the last forty-second strip is for A between 1973 and 2022.

Before the first scan, the virus can be anywhere in the square. After John's first scan and the virus's first move, we can be certain that the top 49 rows in the first strip are virus-free (assuming, of course, that the virus was not found). After the second scan (and the second move), the top 48 rows in the first strip and the top 49 rows in the second strip are virus-free, and so on. After the first series of 42 scans, the number of virus-free upper rows in each of the strips is 8, 9, 10, ..., 49, respectively.

After the first scan of the second series (the forty-third scan in total), the numbers of virus-free upper rows are 56, 8, 9, ..., 48, then 55, 56, 8, 9, ..., 47, and so on. After the entire second series of scans, there are 15, 16, ..., 56 virus-free rows in each of the strips, respectively.

Proceeding in this manner, we see that after the second-to-last (that is, 282nd) series, the numbers of free rows by the strip are as follows: 1975, 1976, ..., 2016. Now in the last series, the first scan ensures that the first strip is virus-free (that is, the numbers of virus-free rows now stand at 2022, 1975, ..., 2015), the second ensures that the first two strips are virus-free (giving the numbers 2022, 2022, 1975, ..., 2014) and so on. Finally, after the last scan, the entire square would be virus-free, which means that John must catch the virus by that last scan.