PUMaC

Combinatorics A Solutions

1. In the country of PUMaC-land, there are 5 villages and 3 cities. Vedant is building roads between the 8 settlements according to the following rules:

   a) There is at most one road between any two settlements;

   b) Any city has exactly three roads connected to it;

   c) Any village has exactly one road connected to it;

   d) Any two settlements are connected by a path of roads.

   In how many ways can Vedant build the roads?

   Proposed by Sunay Joshi

   Answer: $\boxed{90}$

   If a village is connected to another village, then neither village is connected to the rest of the settlements, so this cannot be possible. Thus, every village is connected to a city. If some city is connected to three villages, then these four settlements cannot be connected to the other four, which means this is impossible. Thus, each city is connected to at most two villages, which is only possible if two cities are connected to two villages and one city is connected to one village. The only possible such configuration has the one-village city connected to both other cities. There are 3 ways to choose which city is the one-village city, then 5 ways to choose which village is connected to this city. Finally, there are $\binom{4}{2} = 6$ ways to choose the two villages connected to one of the other cities. Thus, our total number of possibilities is $3 \cdot 5 \cdot 6 = 90$.

2. Ten evenly spaced vertical lines in the plane are labeled $\ell_1, \ell_2, \ldots, \ell_{10}$ from left to right. A set $\{a, b, c, d\}$ of four distinct integers $a, b, c, d \in \{1, 2, \ldots, 10\}$ is *squarish* if some square has one vertex on each of the lines ℓ_a, ℓ_b, ℓ_c, and ℓ_d. Find the number of squarish sets.

   Proposed by Ben Zenker

   Answer: $\boxed{50}$

   Without loss of generality, assume that $a < b < c < d$. Then, it is easy to see that $\{a, b, c, d\}$ is squarish if and only if the distance between ℓ_a and ℓ_b equals the distance between ℓ_c and ℓ_d. In other words, we must count the number of subsets $\{a, b, c, d\}$ of $\{1, 2, \ldots, 10\}$ with $d - c = b - a$.

   To do this, we proceed by casework on $k = d - c$.

   - If $k = 1$, we find that for each value of d, the maximum possible value for a is $d - 3$, so there are $d - 3$ possible combinations of the four numbers. Then, d ranges from 4 to 10 inclusive, for a total of $7 + 6 + 5 + 4 + 3 + 2 + 1 = 28$ combinations.
   - If $k = 2$, each value of d gives $d - 5$ possible combinations. Then, d ranges from 6 to 10 inclusive, for a total of $5 + 4 + 3 + 2 + 1 = 15$ combinations.
   - If $k = 3$, each value of d gives $d - 7$ possible combinations. Then, d ranges from 8 to 10 inclusive, for a total of $3 + 2 + 1 = 6$ combinations.
   - If $k = 4$, there is only 1 combination $a = 1, b = 5, c = 6, d = 10$.

   Summing yields a total of $28 + 15 + 6 + 1 = 50$ sets $\{a, b, c, d\}$.

3. Randy has a deck of 29 distinct cards. He chooses one of the 29! permutations of the deck and then repeatedly rearranges the deck using that permutation until the deck returns to its original order for the first time. What is the maximum number of times Randy may need to rearrange the deck?

Proposed by Aditya Gollapudi and Owen Yang

Answer: $\boxed{2520}$

Every permutation can be decomposed into disjoint cycles, so the number of times Randy shuffle the deck for a given permutation is equal to the least common multiple of the lengths of these cycles. Thus, we want to maximize the LCM of these lengths under the constraint that the lengths sum to 29. Since length 1 cycles do not increase the LCM, we may instead assume that the lengths are greater than one and have sum at most 29 (which we can compensate for by creating many cycles of length 1). We may also assume that these lengths are relatively prime, since removing a common factor from one of the lengths does not change the LCM and decreases the total sum.

If we have three cycle lengths that are not equal to 1, say a, b, c, then by AM-GM we have $\text{lcm}(a, b, c) = abc \le (\frac{a+b+c}{3})^3 < 10^3 = 1000$. Similar proofs show that we cannot have only one or two cycle lengths. On the other hand, if we have five cycle lengths not equal to 1, then the set of 5 relatively prime numbers with smallest sum is $2, 3, 5, 7, 11$ which has sum 28, which has LCM 2310. Any other set of 5 relatively prime numbers has a sum larger than 29. Furthermore, the smallest sum of 6 or more relatively prime numbers is more than 29.

Thus, we need only consider sets of four cycle lengths; call them a, b, c, d. Note that $5+7+8+9 = 29$ and these four numbers have LCM 2520. Since $a \neq b \neq c \neq d$ and each number is as close to the mean $\frac{29}{4}$ as possible, the only other possible maximum is at $\{a, b, c, d\} = \{5, 6, 8, 10\}$, which gives a smaller LCM. Thus, the answer is 2520.

4. Let C_n denote the n-dimensional unit cube, consisting of the 2^n points

$$\{(x_1, x_2, \ldots, x_n) | x_i \in \{0, 1\} \text{ for all } 1 \le i \le n\}.$$

A tetrahedron is *equilateral* if all six side lengths are equal. Find the smallest positive integer n for which there are four distinct points in C_n that form a non-equilateral tetrahedron with integer side lengths.

Proposed by Ryan Alweiss

Answer: $\boxed{11}$

Note that the square of the Euclidean distance between any two points in C_n equals the Hamming distance d_H between the points, which is defined as $d_H(x, y) = |\{i \mid 1 \le i \le n, x_i \neq y_i\}|$. Note that $d_H(x, y) + d_H(y, z) \ge d_H(x, z)$ for all $x, y, z \in C_n$.

I claim that if A and B are vertices of the tetrahedron, then $d_H(A, B) > 1$. If not, then given a third vertex z of the tetrahedron, $d_H(z, x)$ and $d_H(z, y)$ are two squares that differ by 1. The only such squares are 0 and 1, which implies that either $z = x$ or $z = y$, which is impossible. Therefore, we have $d_H(x, y) \ge 4$. We cannot have $n < 9$, because otherwise the only possible integer side length would be 2 and the tetrahedron would be equilateral.

For a construction when $n = 11$, consider the vertices given by $(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$, $(0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1)$, $(1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1)$, $(1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0)$.

It suffices to rule out any possibilities for $n = 9$ and $n = 10$. Since one of the side lengths of the tetrahedron must be 3, we must have that $d_H(A, B) = 9$ for some points $A, B \in C_n$. Thus, for any other vertex C, we cannot have $d_H(A, C) = d_H(B, C) = 4$ because $d_H(A, C) + d_H(C, B) \ge d_H(A, B) = 9$. Since the only other possible side length is 9, we can assume without of loss of generality that $d_H(A, C) = 9$. If $n = 9$, then there is exactly one point with Hamming distance 9 away from A, which implies that $B = C$, a contradiction. If $n = 10$, then B and C must have at least $9 + 9 - 10 = 8$ coordinates in common, which means they have Hamming distance at most 2. But we know that $d_H(B, C) \ge 4$, which is a contradiction.

It follows that $n = 11$ is minimal.

5. An *n-folding process* on a rectangular piece of paper with sides aligned vertically and horizontally consists of repeating the following process n times:

   - Take the piece of paper and fold it in half vertically (choosing to either fold the right side over the left, or the left side over the right).
   - Rotate the paper 90° degrees clockwise.

   A 10-folding process is performed on a piece of paper, resulting in a 1-by-1 square base consisting of many stacked layers of paper. Let $d(i, j)$ be the Euclidean distance between the center of the ith square from the top and the center of the jth square from the top before the paper was folded. Determine the maximum possible value of $\sum_{i=1}^{1023} d(i, i+1)$.

   Proposed by Frank Lu

   Answer: $\boxed{14043}$

   We will determine the answer by inducting on n, the index of the folding process that results in a 1-by-1 square. Let S_n be the answer for n, so that our answer is S_{10}. Note that $S_1 = 1$ since we have two adjacent squares, whose centers are clearly 1 apart.

   It is easy to see that $S_1 = 1$, as we only have one pair of consecutive labels, and they are adjacent. Now, consider a piece of paper P such that, after performing a $k+1$-folding process, we get a 1-by-1 square. Let Q be another piece of paper with half the size of P, so that Q is identical to the top of P after the first step in the folding process. Perform the same remaining k steps in the folding process on both P and Q, label the ith square from the top with i, then unfold P and Q k times, so that Q is completely unfolded and P is still folded once.

   By preserving the orientations of P and Q and placing Q directly over P, each square labeled a in Q is directly over the squares labeled $2a-1$ and $2a$ in P. Furthermore, considering consecutive squares a and $a+1$ on Q, if we flip P so that squares $2a-1$ and $2a$ are on opposite halves of the fold, and squares $2a$ and $2a+1$ are on the same half of the fold. Thus, the distance between the centers of squares $2a$ and $2a+1$ in P is the same as the distance between the centers of squares a and $a+1$ in Q, and the distance between the centers of squares $2a-1$ and $2a$ is twice the distance from the center of square a in Q to the edge corresponding to the fold.

   Note that $S_{k+1} = \sum_{i=1}^{2^{k+1}-1} d(i, i+1)$ can be split up into two sums, one for all terms where i is odd and one for all terms where i is even. The sum where i is even is just S_k, and the sum where i is odd is twice the sum of the distances from all centers of squares to the edge of the fold. Since Q has dimensions $2^{\lfloor \frac{k}{2} \rfloor}$ by $2^{\lceil \frac{k}{2} \rceil}$, and the fold is the edge of Q with length $2^{\lceil \frac{k}{2} \rceil}$, we have that this sum is $2 \cdot 2^{\lceil \frac{k}{2} \rceil} \sum_{j=1}^{2^{\lfloor \frac{k}{2} \rfloor}} \frac{2j-1}{2} = 2^{\lceil \frac{k}{2} \rceil}(2^{\lfloor \frac{k}{2} \rfloor})^2 = 2^{\lceil \frac{k}{2} \rceil + 2\lfloor \frac{k}{2} \rfloor} = 2^{k+\lfloor \frac{k}{2} \rfloor}$. Thus, $S_{k+1} = S_k + 2^{k+\lfloor \frac{k}{2} \rfloor}$.

   Starting from $S_1 = 1$, we see that $S_{10} = 1 + \sum_{j=1}^{9} 2^{j+\lfloor \frac{j}{2} \rfloor} = 3 + \sum_{j=2}^{9} 2^{j+\lfloor \frac{j}{2} \rfloor} = 3 + \sum_{l=1}^{4} 2^{3l} + 2^{3l+1} = 3 + 3\sum_{l=1}^{4} 2^{3l} = 3\sum_{l=0}^{4} 2^{3l}$. This is $3 \cdot \frac{2^{15}-1}{2^3-1} = 3 \cdot \frac{32767}{7} = 3 \cdot 4681 = 14043$. (In particular, it doesn't matter which sides were folded over at each step, the sum is always the same!)

6. Fine Hall has a broken elevator. Every second, it goes up a floor, goes down a floor, or stays still. You enter the elevator on the lowest floor, and after 8 seconds, you are again on the lowest floor. If every possible such path is equally likely to occur, the probability you experience no stops is $\frac{a}{b}$, where a, b are relatively prime positive integers. Find $a + b$.

   Proposed by Adam Huang

   Answer: $\boxed{337}$

Suppose there are u ups, d downs, and s seconds at which the elevator stays still. Since the elevator returns to its original height, $u = d$. Since 8 seconds elapse, $u + d + s = 2u + s = 8$. It is clear that s is even, so $s \in \{0, 2, 4, 6, 8\}$, and $u = d = \frac{8-s}{2}$. We do casework based on the value of s.

Given a path with s stops, delete all seconds at which the elevator stays still to obtain a "reduced" path. The resulting path corresponds to a walk in the u-d plane by sending each up step to $(1, 0)$ and each down step to $(0, 1)$, such that the walk goes from $(0, 0)$ to $(\frac{8-s}{2}, \frac{8-s}{2})$ without crossing the line $u = d$. The number of such paths is the Catalan number $C_{\frac{8-s}{2}}$. Now, we count how many ways there are to insert the stops into the path. Suppose we insert x_i stops before the i-th move in the reduced path, as well as x_{9-s} after the last move. We must count the number of solutions to $x_1 + \ldots + x_{9-s} = s$ over the nonnegative integers. By stars and bars, this is $\binom{8}{8-s}$.

We can explicitly compute that $C_0 = 1, C_1 = 1, C_2 = 2, C_3 = 5, C_4 = 14$. Then, the total number of possible paths is $\sum_{s\in\{0,2,4,6,8\}} \binom{8}{8-s} C_{\frac{8-s}{2}} = 323$. The number of paths with no stops is simply the term corresponding to $s = 0$, namely $\binom{8}{8} C_4 = 14$. It follows that the desired probability is $\frac{14}{323}$, so our answer is $323 + 14 = 337$.

7. Kelvin has a set of eight vertices. For each pair of distinct vertices, Kelvin independently draws an edge between them with probability $p \in (0, 1)$. A set S of four distinct vertices is called *good* if there exists an edge between v and w for all $v, w \in S$ with $v \neq w$. The variance of the number of good sets can be expressed as a polynomial $f(p)$ in the variable p. Find the sum of the absolute values of the coefficients of $f(p)$.

   (The *variance* of random variable X is defined as $\mathbb{E}[X^2] - \mathbb{E}[X]^2$.)

   Proposed by Sunay Joshi

   Answer: $\boxed{7420}$

   For convenience, let $n = 8$ and let X denote the number of good subsets. Note that $X = \sum_{|S|=4} I_S$, where I_S denotes the indicator random variable that S is a good subset, and where the sum runs over all subsets of size 4. Writing $\text{Var}(X) = \text{Cov}(X, X)$ and expanding using bilinearity, we find

$$\text{Var}(X) = \sum_{|S|=4} \text{Var}(I_S) + 2 \sum_{|S|,|T|=4, S\neq T} \text{Cov}(I_S, I_T),$$

   where the second sum runs over all pairs of distinct sets S, T of size 4. Since $\text{Var}(I_S) = \mathbb{P}(\text{S good}) - \mathbb{P}(\text{S good})^2$ and since $\mathbb{P}(\text{S good}) = p^6$, we have $\text{Var}(I_S) = p^6 - p^{12}$. We now consider the second sum. Note that $\text{Cov}(I_S, I_T) = \mathbb{P}(\text{S and T good}) - \mathbb{P}(\text{S good})\mathbb{P}(\text{T good})$. From the above, this reduces to $\mathbb{P}(\text{S and T good}) - p^{12}$. If $|S \cap T| = 2$, then $\mathbb{P}(\text{S and T good}) = p^{11}$, and there are $\binom{n}{2}\binom{n-2}{2}\binom{n-4}{2} \cdot \frac{1}{2}$ such pairs. If $|S \cap T| = 3$, then $\mathbb{P}(\text{S and T good}) = p^9$, and there are $\binom{n}{3}\binom{n-3}{2}$ such pairs. Plugging this into our sum, we find

$$\text{Var}(X) = \binom{n}{4}(p^6 - p^{12}) + 2\left[\binom{n}{2}\binom{n-2}{2}\binom{n-4}{2} \cdot \frac{1}{2}(p^{11} - p^{12}) + \binom{n}{3}\binom{n-3}{2}(p^9 - p^{12})\right]$$

   Rewriting, we find the polynomial

$$f(p) = -p^{12}\left(\binom{n}{4} + \binom{n}{2}\binom{n-2}{2}\binom{n-4}{2} + 2\binom{n}{3}\binom{n-3}{2}\right) + p^{11}\left(\binom{n}{2}\binom{n-2}{2}\binom{n-4}{2}\right) + p^9\left(2\binom{n}{3}\binom{n-3}{2}\right) + p^6\binom{n}{4}$$

   Plugging in $n = 8$ and summing the absolute values of the coefficients, we find 7420, our answer.

8. A permutation $\pi : \{1, 2, \ldots, N\} \to \{1, 2, \ldots, N\}$ is *very odd* if the smallest positive integer k such that $\pi^k(a) = a$ for all $1 \le a \le N$ is odd, where π^k denotes π composed with itself k times. Let $X_0 = 1$, and for $i \ge 1$, let X_i be the fraction of all permutations of $\{1, 2, \ldots, i\}$ that are very odd. Let $\mathcal{S}$ denote the set of all ordered 4-tuples (A, B, C, D) of nonnegative integers such that $A + B + C + D = 2023$. Find the last three digits of the integer

$$2023 \sum_{(A,B,C,D)\in\mathcal{S}} X_A X_B X_C X_D.$$

Proposed by Daniel Carter

Answer: $\boxed{116}$

A permutation is very odd if and only if its cycle decomposition consists only of odd cycles. Letting c_k be the number of cycles of length k, the number of very odd permutations for a given n is therefore equal the sum over all solutions to the equation $n = \sum_{i=1}^{\infty}(2i-1)c_{2i-1}$ of the number of ways there are to create a permutation with those cycle lengths. Given any fixed c_k, there are

$$\frac{1}{c_k!}\prod_{i=0}^{c_k-1}\binom{n-ik}{k} = \frac{1}{c_k!}\cdot\frac{n!}{(n-kc_k)!(k!)^{c_k}}$$

ways to choose the numbers in c_k cycles of length k where the cycles are indistinguishable, and $(k-1)!$ ways to order each cycle given the k numbers in a cycle, for a total of

$$\frac{((k-1)!)^{c_k}}{c_k!}\cdot\frac{n!}{(n-kc_k)!(k!)^{c_k}} = \frac{n!}{(n-kc_k)!c_k!k^{c_k}}$$

ways to form the c_k cycles of length k Now, we can repeat the process on c_{k+2} using the new $n' = n - kc_k$. Multiplying this over all k, the product

$$\frac{n!}{(n-k_1)!}\cdot\frac{(n-k_1)!}{(n-3k_3)!}\cdot\frac{(n-3k_3)!}{(n-5k_5)!}\cdots$$

telescopes to $\frac{n!}{0!} = n!$, which gives that there are $\frac{n!}{\prod_{i=1}^{\infty} c_{2i-1}!(2i-1)^{c_{2i-1}}}$ ways to create a permutation with c_{2i-1} cycles of length $2i-1$ for all $i \in \mathbb{N}$. Therefore, the generating function of the sequence X_n is given by

$$\sum_{n=0}^{\infty} X_n t^n = \prod_{n=0}^{\infty}\left(\sum_{m=0}^{\infty}\frac{1}{m!(2n+1)^m}t^{m(2n+1)}\right) = \prod_{n=0}^{\infty}\exp\left(\frac{t^{2n+1}}{2n+1}\right) = \exp(t + \frac{t^3}{3} + \frac{t^5}{5} + \ldots)$$

Let $\mathcal{S}_n$ denote the set of all ordered 4-tuples of nonnegative integers satisfying $a+b+c+d = n$. Consider the sequence $Y_n = \sum_{(a,b,c,d)\in\mathcal{S}_n} X_a X_b X_c X_d$. Our answer is $2023Y_{2023}$. But the generating function of Y_n is the fourth power of the generating function of X_n, so since the generating function of $\log(1+x)$ is $\sum_{n=1}^{\infty}\frac{(-1)^{n+1}x^n}{n}$, we have

$$\sum_{n=0}^{\infty} Y_n t^n = \exp(4t + \frac{4t^3}{3} + \frac{4t^5}{5} + \ldots) = \exp(2\log(1+t) - 2\log(1-t)) = \left(\frac{1+t}{1-t}\right)^2$$

$$= \left(\frac{2}{1-t} - 1\right)^2 = 1 - \frac{4}{1-t} + \frac{4}{(1-t)^2} = 1 + \frac{4t}{1-t}$$

Since the generating function for $\frac{1}{1-t}$ is $\sum_{n=0}^{\infty} t^n$, the generating function for this is equal to $1 + 4\sum_{n=1}^{\infty} nt^n$. The coefficient of t^{2023} is $Y_{2023} = 4 \cdot 2023$, so our answer is $4 \cdot 2023^2$, the last three digits of which are 116.