HMMT November 2017

November 11, 2017

Theme Round

1. **T**wo ordered pairs (a, b) and (c, d), where a, b, c, d are real numbers, form a basis of the coordinate plane if $ad \neq bc$. Determine the number of ordered quadruples (a, b, c, d) of integers between 1 and 3 inclusive for which (a, b) and (c, d) form a basis for the coordinate plane.

   Proposed by: Ashwin Sah

   Answer: $\boxed{66}$

   Any pair of distinct points will form a basis except when (a, b) and (c, d) are both from $\{(1, 1), (2, 2), (3, 3)\}$, so the answer is $9 \cdot 8 - 3 \cdot 2 = 66$.

2. **H**orizontal parallel segments $AB = 10$ and $CD = 15$ are the bases of trapezoid $ABCD$. Circle γ of radius 6 has center within the trapezoid and is tangent to sides AB, BC, and DA. If side CD cuts out an arc of γ measuring 120°, find the area of $ABCD$.

   Proposed by: Michael Tang

   Answer: $\boxed{\frac{225}{2}}$

   Suppose that the center of the circle is O and the circle intersects CD at X and Y. Since $\angle XOY = 120^\circ$ and triangle XOY is isosceles, the distance from O to XY is $6 \cdot \sin(30^\circ) = 3$. On the other hand, the distance from O to AB is 6 as the circle is tangent to AB, and O is between AB and CD, so the height of the trapezoid is $6 + 3 = 9$ and its area is $\frac{9 \cdot (10+15)}{2} = \frac{225}{2}$.

3. **E**milia wishes to create a basic solution with 7% hydroxide (OH) ions. She has three solutions of different bases available: 10% rubidium hydroxide (Rb(OH)), 8% cesium hydroxide (Cs(OH)), and 5% francium hydroxide (Fr(OH)). (The Rb(OH) solution has both 10% Rb ions and 10% OH ions, and similar for the other solutions.) Since francium is highly radioactive, its concentration in the final solution should not exceed 2%. What is the highest possible concentration of rubidium in her solution?

   Proposed by: Yuan Yao

   Answer: $\boxed{1\%}$

   Suppose that Emilia uses R liters of Rb(OH), C liters of Cs(OH), and F liters of Fr(OH), then we have

   $$\frac{10\% \cdot R + 8\% \cdot C + 5\% \cdot F}{R + C + F} = 7\% \text{ and } \frac{5\% \cdot F}{R + C + F} \leq 2\%.$$

   The equations simplify to $3R + C = 2F$ and $3F \leq 2R + 2C$, which gives

   $$\frac{9R + 3C}{2} \leq 2R + 2C \Rightarrow 5R \leq C.$$

   Therefore the concentration of rubidium is maximized when $5R = C$, so $F = 4R$, and the concentration of rubidium is

   $$\frac{10\% \cdot R}{R + C + F} = 1\%.$$

4. **M**ary has a sequence $m_2, m_3, m_4, \ldots$, such that for each $b \geq 2$, m_b is the least positive integer m for which none of the base-b logarithms $\log_b(m), \log_b(m+1), \ldots, \log_b(m+2017)$ are integers. Find the largest number in her sequence.

   Proposed by: Michael Tang

   Answer: $\boxed{2188}$

   It is not difficult to see that for all of the logarithms to be non-integers, they must lie strictly between n and $n+1$ for some integer n. Therefore, we require $b^{n+1} - b^n > 2018$, and so $m_b = b^n + 1$ where n is the smallest integer that satisfies the inequality. In particular, this means that $b^n - b^{n-1} \leq 2018$.

Note that $m_2 = 2^{11} + 1 = 2049$ (since $2^{12} - 2^{11} = 2048 > 2018$) and $m_3 = 3^7 + 1 = 2188$ (since $3^8 - 3^7 = 4374 > 2018$). we now show that 2188 is the maximum possible value for m_b.

If $n = 0$, then $m_b = 1 + 1 = 2$.

If $n = 1$, then $b - 1 \leq 2018$ and thus $m_b = b + 1 \leq 2020$.

If $n = 2$, then $b^2 - b \leq 2018$, which gives $b \leq 45$, and thus $m_b = b^2 + 1 \leq 2018 + b + 1 \leq 2065$.

If $n = 3$, then $b^3 - b^2 \leq 2018$, which gives $b \leq 12$, and thus $m_b = b^3 + 1 \leq 12^3 + 1 = 1729$.

If $n = 4$, then $b^4 - b^3 \leq 2018$, which gives $b \leq 6$, and thus $m_b = b^4 + 1 \leq 6^4 + 1 = 1297$.

It then remains to check the value of m_4 and m_5. Indeed, $m_4 = 4^5 + 1 = 1025$ and $m_5 = 5^4 + 1 = 626$, so no values of m_b exceeds 2188.

5. **E**ach of the integers $1, 2, \ldots, 729$ is written in its base-3 representation without leading zeroes. The numbers are then joined together in that order to form a continuous string of digits: $12101112202122\ldots$. How many times in this string does the substring 012 appear?

   Proposed by: Michael Tang

   Answer: $\boxed{148}$

   Ignore $729 = 3^6 = 1000000_3$ since it will not contribute to a 012 substring. Break into cases on how 012 appears: (i) when an individual integer contains the string 012; (ii) when 01 are the last two digits of an integer and 2 is the first digit of the next integer; and (iii) when 0 is the last digit of an integer and 12 are the first two digits of the next integer.

   For case (i), we want to find the total number of appearances of the string 012 in $1, 2, 3, \ldots, N$. Since each number has at most six digits, 012 appears at most once per number. If such a number has d digits, $4 \leq d \leq 6$, then there are $d - 3$ possible positions for the substring 012 and $2 \cdot 3^{d-4}$ possible choices for the remaining digits (since the leftmost digit must be nonzero). Thus there are

$$\sum_{d=4}^{6} (d - 3) \cdot \left(2 \cdot 3^{d-4}\right) = 1 \cdot 2 + 2 \cdot 6 + 3 \cdot 18 = 68$$

   appearances of 012 in case (i).

   For case (ii), we have an integer n for which n ends in 01 and $n + 1$ starts with 2. Then n must also start with 2. Hence it suffices to count the number of integers $1 \leq n < N$ which start with 2 and end with 01 in base three. If n has d digits, $3 \leq d \leq 6$, then there are 3^{d-3} possibilities for the middle digits, so we have

$$\sum_{d=3}^{6} 3^{d-3} = 1 + 3 + 9 + 27 = 40$$

   appearances of 012 in case (ii).

   For case (ii), we have an integer n for which n ends in 0 and $n + 1$ starts with 12. Then n must also start with 12, so we want to count the number of $1 \leq n < N$ starting with 12 and ending in 0. Like case (ii), there are also 40 appearances of 012 in case (iii).

   In total, the substring 012 appears $68 + 40 + 40 = 148$ times.

6. **R**thea, a distant planet, is home to creatures whose DNA consists of two (distinguishable) strands of bases with a fixed orientation. Each base is one of the letters H, M, N, T, and each strand consists of a sequence of five bases, thus forming five pairs. Due to the chemical properties of the bases, each pair must consist of distinct bases. Also, the bases H and M cannot appear next to each other on the same strand; the same is true for N and T. How many possible DNA sequences are there on Rthea?

   Proposed by: Yuan Yao

   Answer: $\boxed{12 \cdot 7^4 \text{ or } 28812}$

   There are $4 \cdot 3 = 12$ ways to choose the first base pairs, and regardless of which base pair it is, there are 3 possibilities for the next base on one strand and 3 possibilities for the next base on the other strand. Among these possibilities, exactly 2 of them have identical bases forming a base pair (using one of the base not in the previous base pair if the previous pair is H-M or N-T, or one of the base in the previous pair otherwise), which is not allowed. Therefore there are $3 \cdot 3 - 2 = 7$ ways to choose each of the following base pairs. Thus in total there are $12 \cdot 7^4 = 28812$ possible DNA (which is also the maximum number of species).

7. **O**n a blackboard a stranger writes the values of $s_7(n)^2$ for $n = 0, 1, \ldots, 7^{20} - 1$, where $s_7(n)$ denotes the sum of digits of n in base 7. Compute the average value of all the numbers on the board.

   Proposed by: Ashwin Sah

   Answer: $\boxed{3680}$

   Solution 1: We solve for 0 to $b^n - 1$ and $s_b(n)^2$ (i.e. base b).

   Let $n = d_1 \ldots d_n$ in base b, where there may be leading zeros. Then $s_b(n) = d_1 + \cdots + d_n$, regardless of the leading zeros.

$$\mathbb{E}\left[s_d(n)^2\right] = \mathbb{E}\left[(d_1 + \cdots + d_n)^2\right] = \sum_{1 \le i \le n} \mathbb{E}\left[d_i^2\right] + 2 \sum_{1 \le i < j \le n} \mathbb{E}[d_i d_j],$$

   and now notice that we can treat choosing n uniformly as choosing the d_i uniformly independently from $\{0, \ldots, b-1\}$. Thus this simplifies to

$$\mathbb{E}\left[s_d(n)^2\right] = n\mathbb{E}\left[d_1^2\right] + n(n-1)\mathbb{E}\left[d_1\right]^2.$$

   Now

$$\mathbb{E}\left[d_1^2\right] = \frac{0^2 + \cdots + (b-1)^2}{b} = \frac{(b-1)(2b-1)}{6},$$
$$\mathbb{E}[d_1] = \frac{0 + \cdots + (b-1)}{b} = \frac{b-1}{2},$$

   so the answer is

$$n \cdot \frac{(b-1)(2b-1)}{6} + n(n-1) \cdot \left(\frac{b-1}{2}\right)^2.$$

   Plugging in $b = 7, n = 20$ yields the result.

   Solution 2: There are two theorems we will cite regarding variance and expected value. The first is that, for any variable X,

$$Var(X) = E[X^2] - E[X]^2.$$

   The second is that, for two independent variables X and Y,

$$Var(X + Y) = Var(X) + Var(Y).$$

   Let X be the sum of all of the digits. We want to find $E[X^2]$. The expected of a single digit is $\frac{1}{7}(0+1+2+3+4+5+6) = 3$. Thus, the expected value of the sum of the digits is $E[X] = 20 \times 3 = 60$, so $E[X]^2 = 3600$. The variance of a single digit is $\frac{1}{7}[(0-3)^2 + (1-3)^2 + \ldots + (6-3)^2] = \frac{9+4+1+0+1+4+9}{7} = 4$. Since the digits are independent, their variances add by the second theorem above. Therefore, the variance of the sum of all of the digits is $Var(X) = 20 \times 4 = 80$. Finally, using the first theorem, we have $E[X^2] = E[X]^2 + Var(X) = 3680$.

8. **U**ndecillion years ago in a galaxy far, far away, there were four space stations in the three-dimensional space, each pair spaced 1 light year away from each other. Admiral Ackbar wanted to establish a base somewhere in space such that the sum of squares of the distances from the base to each of the stations does not exceed 15 square light years. (The sizes of the space stations and the base are negligible.) Determine the volume, in cubic light years, of the set of all possible locations for the Admiral's base.

   Proposed by: Yuan Yao

   Answer: $\boxed{\frac{27\sqrt{6}}{8}\pi}$

Solution 1: Set up a coordinate system where the coordinates of the stations are $(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}})$, $(-\frac{1}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}})$, $(\frac{1}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}})$, and $(-\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}})$. The sum of squares of the distances is then

$$\sum_{\text{sym}} 2(x - \frac{1}{2\sqrt{2}})^2 + 2(x + \frac{1}{2\sqrt{2}})^2 = 4(x^2 + y^2 + z^2) + \frac{3}{2} = 4r^2 + \frac{3}{2},$$

where r is the distance from the center of the tetrahedron. It follows from $4r^2 + \frac{3}{2} \leq 15$ that $r \leq \frac{3\sqrt{6}}{4}$, so the set is a ball with radius $R = \frac{3\sqrt{6}}{4}$, and the volume is $\frac{4\pi}{3}R^3 = \frac{27\sqrt{6}}{8}\pi$.

Solution 2: Let P be the location of the base; S_1, S_2, S_3, S_4 be the stations; and G be the center of the tetrahedron. We have:

$$\sum_{i=1}^{4} PS_i^2 = \sum_{i=1}^{4} \vec{PS_i} \cdot \vec{PS_i}$$

$$\sum_{i=1}^{4} PS_i^2 = \sum_{i=1}^{4} \left(\vec{PG} + \vec{GS_i}\right) \cdot \left(\vec{PG} + \vec{GS_i}\right)$$

$$\sum_{i=1}^{4} PS_i^2 = \sum_{i=1}^{4} \left(\vec{PG} \cdot \vec{PG} + 2\vec{PG} \cdot \vec{GS_i} + \vec{GS_i} \cdot \vec{GS_i}\right)$$

Since $GS_1 = GS_2 = GS_3 = GS_4$, we have:

$$\sum_{i=1}^{4} PS_i^2 = 4PG^2 + 4GS_1^2 + 2\sum_{i=1}^{4} \left(\vec{PG} \cdot \vec{GS_i}\right)$$

$$\sum_{i=1}^{4} PS_i^2 = 4PG^2 + 4GS_1^2 + 2\vec{PG} \cdot \left(\sum_{i=1}^{4} \vec{GS_i}\right)$$

Since G is the center of the tetrahedron, $\vec{GS_1} + \vec{GS_2} + \vec{GS_3} + \vec{GS_4} = \vec{0}$. Thus:

$$\sum_{i=1}^{4} PS_i^2 = 4PG^2 + 4GS_1^2$$

Since $GS_1^2 = \frac{3}{8}$, we have $PG^2 \leq \frac{27}{8}$. Thus, the locus of all good points is a ball centered at G with radius $r = \sqrt{\frac{27}{8}}$. Then, the volume is $V = \frac{4}{3}\pi r^3 = \frac{27\sqrt{6}\pi}{8}$.

9. **N**ew this year at HMNT: the exciting game of *RNG baseball*! In RNG baseball, a team of infinitely many people play on a square field, with a base at each vertex; in particular, one of the bases is called the *home base*. Every turn, a new player stands at home base and chooses a number n uniformly at random from $\{0, 1, 2, 3, 4\}$. Then, the following occurs:

- If $n > 0$, then the player and everyone else currently on the field moves (counterclockwise) around the square by n bases. However, if in doing so a player returns to or moves past the home base, he/she leaves the field immediately and the team scores one point.
- If $n = 0$ (a strikeout), then the game ends immediately; the team does not score any more points.

What is the expected number of points that a given team will score in this game?

Proposed by: Yuan Yao

Answer: $\boxed{\frac{409}{125}}$

For $i = 0, 1, 2, 3$, let P_i be the probability that a player on the i-th base scores a point before strikeout (with zeroth base being the home base). We have the following equations:

$$
\begin{aligned}
P_0 &= \frac{1}{5}(P_1 + P_2 + P_3 + 1) \\
P_1 &= \frac{1}{5}(P_2 + P_3 + 1 + 1) \\
P_2 &= \frac{1}{5}(P_3 + 1 + 1 + 1) \\
P_3 &= \frac{1}{5}(1 + 1 + 1 + 1)
\end{aligned}
$$

Solving the system of equations gives $P_3 = \frac{4}{5}, P_2 = \frac{19}{25}, P_1 = \frac{89}{125}, P_0 = \frac{409}{625}$, so the probability that a batter scores a point himself is $\frac{409}{625}$, given that he is able to enter the game before the game is over. Since the probability that the nth player will be able to stand on the home base is $(\frac{4}{5})^{n-1}$ (none of the previous $n-1$ players receive a strikeout), the expected value is $\frac{409}{625}(1+\frac{4}{5}+(\frac{4}{5})^2+\cdots) = \frac{409}{625} \cdot \frac{1}{1-\frac{4}{5}} = \frac{409}{125}$.

10. **D**enote $\phi = \frac{1+\sqrt{5}}{2}$ and consider the set of all finite binary strings without leading zeroes. Each string S has a "base-ϕ" value $p(S)$. For example, $p(1101) = \phi^3 + \phi^2 + 1$. For any positive integer n, let $f(n)$ be the number of such strings S that satisfy $p(S) = \frac{\phi^{48n}-1}{\phi^{48}-1}$. The sequence of fractions $\frac{f(n+1)}{f(n)}$ approaches a real number c as n goes to infinity. Determine the value of c.

Proposed by: Ashwin Sah

Answer: $\boxed{\frac{25+3\sqrt{69}}{2}}$

We write everything in base ϕ. Notice that

$$\frac{\phi^{48n}-1}{\phi^{48}-1} = 10\ldots010\ldots01\ldots10\ldots01,$$

where there are $n-1$ blocks of 47 zeros each. We can prove that every valid base-ϕ representation comes from replacing a consecutive string 100 with a 011 repeatedly. Using this, we can easily classify what base-ϕ representations are counted by $f(n)$.

Notice that $10000000 = 01100000 = 01011000 = 01010110$ and similar, so that in each block of zeros we can choose how many times to perform a replacement. It turns out that we can do anywhere from 0 to 24 such replacements, but that if we choose to do 24 then the next block cannot have chosen 0 replacements. (An analogy with lower numbers is $10001000 = 01101000 = 01100110 = 01011110$, with the first block "replaced twice," which was only allowed since the second block had "replaced once," opening up the slot which was filled by the last 1 in the final replacement 011).

Thus we have a bijection from $f(n)$ to sequences in $\{0,\ldots,24\}^{n-1}$ such that (a) the sequence does not end in 24 and (b) the sequence never has a 24 followed by a 0.

We let a_n denote the number of length-n sequences starting with a 0, b_n for the number of such sequences starting with any of 1 to 23, and c_n for the number of such sequences starting with 24. We know $a_1 = 1, b_1 = 23, c_0 = 0$ and that $f(n) = a_{n-1} + b_{n-1} + c_{n-1}$.

Now,

$$
\begin{aligned}
a_n &= a_{n-1} + b_{n-1} + c_{n-1} \\
b_n &= 23(a_{n-1} + b_{n-1} + c_{n-1}) \\
c_n &= b_{n-1} + c_{n-1}
\end{aligned}
$$

so $b_n = 23a_n$ for all n. Substituting gives $a_n = 24a_{n-1} + c_{n-1}, c_n = 23a_{n-1} + c_{n-1}$. Solving for $c_n = a_{n+1} - 24a_{n-1}$ and plugging in gives

$$a_{n+1} - 24a_n = a_n - a_{n-1},$$

which gives a characteristic polynomial of $\lambda^2 - 25\lambda + 1 = 0$. We easily find that a_n grows as λ^n (where λ is the larger solution to the quadratic equation) and thus b_n, c_n do as well, implying that $f(n)$ grows as λ^n, where

$$\lambda = \frac{25 + \sqrt{25^2 - 4}}{2} = \frac{25 + 3\sqrt{69}}{2},$$

which is our answer.