The PRIMES 2015 Math Problem Set

Dear PRIMES applicant!

This is the PRIMES 2015 Math Problem Set. Please send us your solutions as part of your PRIMES application by December 1, 2015. For complete rules, see http://web.mit.edu/primes/apply.shtml

Note that this set contains two parts: "General Math problems" and "Advanced Math." Please solve as many problems as you can in both parts.

You can type the solutions or write them up by hand and then scan them. Please attach your solutions to the application as a PDF (preferred), DOC, or JPG file. The name of the attached file must start with your last name, for example, "smith-solutions." Include your full name in the heading of the file.

Please write not only answers, but also proofs (and partial solutions/results/ideas if you cannot completely solve the problem). Besides the admission process, your solutions will be used to decide which projects would be most suitable for you if you are accepted to PRIMES.

You are allowed to use any resources to solve these problems, *except other people's help*. This means that you can use calculators, computers, books, and the Internet. However, if you consult books or Internet sites, please give us a reference.

Note that posting these problems on problem-solving websites before the application deadline is not allowed. Applicants who do so will be disqualified, and their parents and recommenders will be notified.

Note that some of these problems are tricky. We recommend that you do not leave them for the last day. Instead, think about them, on and off, over some time, perhaps several days. We encourage you to apply if you can solve at least 50% of the problems. [1]

Enjoy!

[1] We note, however, that there will be many factors in the admission decision besides your solutions of these problems.

General math problems

Problem G1. You roll three dice trying to get all three to be equal. If at some throw two are equal, you keep rolling the third one in the hope of it turning equal to the other two. What is the chance that you *do not* succeed if you are allowed to roll the dice two times? n times?

Solution. Let's compute the chance p that you will not succeed. The chance that there will be no two equals on a throw is 5/9. So the chance of no two equals all the way is $(5/9)^n$. Now, the chance to get a pair on a throw but not all equals is $90/216 = 5/12$. If this happens on the k-th throw, the chance that we have no luck in the $n-k$ remaining ones is $(5/6)^{n-k}$. So we have

$$p = \left(\frac{5}{9}\right)^n + \frac{5}{12}\sum_{k=1}^{n}\left(\frac{5}{9}\right)^{k-1}\left(\frac{5}{6}\right)^{n-k} = \frac{3 \cdot 15^n - 10^n}{2 \cdot 18^n}.$$

Example: for $n = 1$ this is 35/36 and for $n = 2$ it is 575/648.

Problem G2. John lives 2 miles north from a road, which is separated from John's house by a grove. If he walks from his house to the road along any straight line, the last mile of his walk is through the grove. Find the shape of the grove (describe its northern boundary by an equation).

Solution. Let the road be the line $x = 0$ and John's house be at the point (0,2). Assume that he walks on the line $y = 2 - 2x/a$, for some real a (for $a = 0$ this is the line $x = 0$). Its x-intercept is a. The point where the grove begins has coordinates

$$(a - \frac{a}{\sqrt{a^2+4}}, \frac{2}{\sqrt{a^2+4}})$$

which gives a parametric equation of the curve. We can solve for a in terms of y:

$$a = 2y^{-1}\sqrt{1-y^2}$$

for $y > 0$. Thus

$$x = (2y^{-1} - 1)\sqrt{1-y^2}$$

for $y > 0$. So the equation is

$$x^2y^2 = (2-y)^2(1-y^2).$$

(the component for $y > 0$).

Problem G3. Two white and two black rooks are placed at random on a standard 8-by-8 chessboard. What is the chance that NO white rook attacks a black rook? (recall that a rook attacks along the vertical and horizontal line it stands on). Give the answer as a fraction in lowest terms.

Solution. The number of rook placements is $64 \cdot 63 \cdot 62 \cdot 61$. To count non-attacking placements, let's place white rooks first. For the first one there are 64 positions, and for the second one there are 49 positions so that they don't attack each other and 14 positions if they do. In the first case they leave a 6-by-6 board for the black rooks (so $36 \cdot 35$ positions), and in the second case a 6-by-7 board (so $42 \cdot 41$ positions). Thus the answer is

$$p = \frac{49 \cdot 36 \cdot 35 + 14 \cdot 42 \cdot 41}{63 \cdot 62 \cdot 61} = \frac{2044}{5673}.$$

Problem G4. The number 1 is written on a blackboard. John plays the following game with himself: he chooses a number on the blackboard, multiplies it by 2 or 3, adds 1, and puts the result on the blackboard (if it does not appear there already).

a) How many pairs of consecutive numbers can appear?

b) Can it happen that three consecutive numbers appear on the blackboard?

c) Can it happen that the numbers $n, n+1, n+3, n+4$ appear for some n?

Solution. 1) Infinitely many. If we have a number x, we can also have $3(2x+1)+1 = 6x+4$ and $2(3x+1)+1 = 6x+3$.

2) No. A number of the form $3k+2$ can never appear. Indeed, take the smallest number of this form that appears. Then $3k+2 = 2m+1$, where m also appears, so $2m = 3k+1$ and hence $m = 3p+2$ for $p < k$, contradiction.

3) No. A number divisible by 6 obviously cannot appear, which together with (b) implies the statement (indeed, n and $n+3$ would have to be divisible by 3 so one of them is divisible by 6).

Problem G5. Prove that for every real $C > 0$, there is some finite set $A \subset \mathbb{Z}$ such that $|A+A| \geq C|A-A|$ (where $|X|$ is the number of elements in a set X). Here

$$A + A := \{a + b : a \in A, b \in A\}$$

and

$$A - A := \{a - b : a \in A, b \in A\}$$

are the set of pairwise sums and the set of pairwise differences of A, respectively.

For example, when $A = \{0, 2, 3, 4, 7, 11, 12, 14\}$. We have $A + A = [0, 28] \setminus \{1, 20, 27\}$ and $A - A = [-14, 14] \setminus \{\pm 6, \pm 13\}$. (Here $[m, n]$ denotes $\{m, m+1, \ldots, n\}$). Thus $|A+A| = 26$ and $|A-A| = 25$, so the example proves the statement for $C \leq 26/25$.

Solution. Let $A_1 = \{0, 2, 3, 4, 7, 11, 12, 14\}$ as in the example, so that $|A_1 + A_1| = 26$ and $|A_1 - A_1| = 25$.

For each $n \geq 2$, let A_n denote the 8^n-element set

$$A_n = \{a_0 + a_1 100 + a_2 100^2 + \cdots + a_{n-1} 100^{n-1} : a_0, \ldots, a_{n-1} \in A_1\}.$$

We see that

$$A_n + A_n = \{b_0 + b_1 100 + b_2 100^2 + \cdots + b_{n-1} 100^{n-1} : b_0, \ldots, b_{n-1} \in A_1 + A_1\}$$

so that $|A_n + A_n| = 26^n$. Similarly, $|A_n - A_n| = 25^n$. Thus $|A + A| \geq C|A - A|$ can be achieved by taking $A = A_n$ with n sufficiently large.

Remark. Typically one expect $A - A$ to be larger than $A + A$ since two elements of A generate two differences yet only one sum. Sets A with $|A + A| > |A - A|$ have been called *more sums than differences sets*. The first example of such a set was found by Conway in the 1960's (his set is the one given in the problem). These sets turn out to be more common than one might initially guess. It is expected (from experimental data and also proved to some extent) that about .045% of all subsets of $\{1, 2, \ldots, n\}$ satisfy $|A + A| > |A - A|$ when n is large. For more information, see the following references:

M. B. Nathanson, *Sets with more sums than differences*, Integers 7 (2007), A5, 24 pp. (electronic)

G. Martin and K. O'Bryant, *Many sets have more sums than differences*, Additive combinatorics, CRM Proc. Lecture Notes, vol. 43, Amer. Math. Soc., Providence, RI, 2007, 287305. arXiv:0608131

Y. Zhao, *Sets characterized by missing sums and differences*, J. Number Theory 131 (2011), 2107–2134. arXiv:0911.2292.

Given the "tensor power" construction as in the solution, a more satisfying problem is to determine the maximum and minimum possible values of $\log|A + A|/\log|A - A|$. This is open.

Problem G6. Let $a_n, n \geq 1$ be the sequence determined recursively by the rule

$$a_1 = 1,\ a_{n+1} = \frac{n+2}{n+1} a_n + \frac{n^3 + 3n^2 + 2n - 2}{n(n+1)}.$$

Find a formula for a_n.

Hint. Compute the first few values and try to guess the answer.

Solution. By computing the first few terms and interpolating, we guess that $a_n = (n^3 - n + 1)/n$. This is easily proved by induction.

Problem G7. Let M_n denote the $n \times n$ matrix whose entries are 0 below the main diagonal and 1 on and above the main diagonal. E.g.,

$$M_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Determine M_n^r for any $r \geq 1$.

Solution. We have $M_n = 1 + J + ... + J^{n-1}$, where J has ones right above the main diagonal and zeros elsewhere. This equals $(1-J)^{-1}$ since $J^n = 0$. So $M_n = (1-J)^{-r} = \sum_{m \geq 0} \binom{m+r-1}{r-1} J^m$. The ij-th element in question is the coefficient of J^{j-i}, which is $\binom{j-i+r-1}{r-1}$.

Advanced math problems

Problem M1. There are $2k$ points on the plane. No three of them are co-linear. We build the following graph based on this configuration. Each point is a vertex. Two vertices are connected if and only if the line passing through them divides the other points into two equal sets: there are $(k-1)$ points on each side of the line. We call this graph the middle-road graph based on the given configuration of points.

a) The complete bipartite graph $K_{n,m}$ has vertices $1, ..., n+m$, and i, j are connected if and only if $i \le n$, $j > n$, or $i > n$, $j \le n$. For which pairs n, m, can the complete bipartite graph $K_{n,m}$ be the middle-road graph of some configuration?

b) The path graph P_n on n vertices is the graph with vertices $1, ..., n$ such that i, j are connected if and only if $|i-j| = 1$. For which n, can the path graph P_n be the middle-road graph of some configuration?

If the graph exists explain how to construct it. If not, prove that it doesn't exist.

Solution.

a) The middle-road graph is called a halving-edges graph. By continuously rotating a line around a point we can see that the halving edges graph doesn't have isolated vertices. Moreover, the points on the convex hull have degree 1. That means the halving graph can be a complete bipartite graph only if at least one of n, m is 1. In addition, the total number of vertices is $2k$. So it could only be $K_{1,2k-1}$. Such a configuration can be constructed by placing $2k-1$ vertices on the convex hull of a regular $(2k-1)$-gon and the last vertex at its center.

b) A more subtle rotation argument can show that all vertices of a halving graph have an odd degree. That means the only way it can be a path graph is for $n = 2$. This configuration can be constructed by using any two points on a plane.

Problem M2. Let $n \ge 2$ be an integer. You choose n points $x_1, ..., x_n$ uniformly at random on a circle of length 2. Let $0 < t \le 1$. What is the chance that all the points belong to an arc of length t?

Solution. Let $P_n(t)$ be the probability that the shortest arc containing all the points has length $\le t$. Suppose this happens, and the smallest length of an arc containing $x_1, ..., x_{n-1}$ is $s \le t$. Then the probability that after adding x_n the minimal arc will have length $\le t$ is $t - s/2$. Thus,

$$P_n(t) = \int_0^t P'_{n-1}(s)(t - s/2)ds, \ n \ge 3.$$

Also, $P_2(t) = t$. From this it is clear that $P_n(t) = C_n t^{n-1}$ for some C_n. So we get from the above equation

$$C_{n+1} = C_n(n-1)(\frac{1}{n-1} - \frac{1}{2n}) = \frac{n+1}{2n}C_{n-1}.$$

This implies that $C_n = \frac{n}{2^{n-1}}$, so $P_n(t) = n(t/2)^{n-1}$. So we get: $P_2 = t$, $P_3 = \frac{3}{4}t^2$, $P_4 = \frac{1}{2}t^3$, etc.

Problem M3. (i) Let $x_1, ..., x_k$ be variables, and $n_1, \dots, n_k$ be positive integers. Let A be the matrix with entries $a_{ij} := x_i^{n_j-1}$, and $P(x_1, ..., x_k) = \det A$. Show that

$$\frac{P(x_1, ..., x_n)}{\prod_{i<j}(x_i - x_j)}$$

is a polynomial with integer coefficients.

(ii) Show that

$$\prod_{i<j} \frac{q^{n_i} - q^{n_j}}{q^i - q^j}$$

is a polynomial in q with integer coefficients. Deduce that

$$\prod_{i<j} \frac{n_i - n_j}{i - j}$$

is an integer.

Solution. (i) This follows since the numerator is antisymmetric.

(ii) This follows from (i) if we plug in $x_i = q^{i-1}$ and use the Vandermonde determinant formula for the numerator. The second statement follows from the first one by computing the limit $q \to 1$.

Problem M4. Let f be a continuous function on the plane. In any rectangle $ABCD$ so that AB is parallel to the x-axis and B has greater y-coordinate than C, we have

$$\int_{ABC} f = \int_{CDA} f.$$

Prove that f is constant.

Solution. Let $x_1 < x_2 < x_3$, $y_1 < y_2 < y_3$. Let $P_{ij} = (x_i, y_j)$. Let $R_1 = P_{11}P_{12}P_{22}P_{21}$. $R_2 = P_{12}P_{13}P_{23}P_{22}$, $R_3 = P_{21}P_{22}P_{32}P_{31}$, $R_4 = P_{22}P_{23}P_{33}P_{32}$. Let $R = P_{11}P_{13}P_{33}P_{31}$. For a region Q, let $I(Q)$ be the integral of f over Q. For a rectangle Q, let Q^+, Q^- be the upper and lower triangle of Q, as in the formulation of the problem (so $I(Q^+) = I(Q^-)$). Applying the condition on f to R, we get

$$I(R_1) + I(R_2^-) + I(R_3^-) = I(R_4) + I(R_2^+) + I(R_3^+).$$

Thus, $I(R_1) = I(R_4)$, which implies that $I(Q) = I(P)$ for any two rectangles Q, P with sides parallel to axis, provided that Q, P intersect by a vertex at the right upper, respectively left lower corner and have the same area. Tending the width of Q and height of P to zero, we find that

$$\int_L f/\ell(L) = \int_M f/\ell(M),$$

where L is a vertical interval and M a horizontal interval forming the letter Γ, and ℓ is the length. Tending the length of L to zero, we get

$$\int_M f = \ell(M) f(M_-),$$

where M_- is the left end of M. Differentiating by the length of M, we get $f(M_+) = f(M_-)$ where M_+ is the right end of M. Similarly, tending the length of M to zero and then varying L, we get $f(L_+) = f(L_-)$ for any vertical interval. This implies that f is constant.

Problem M5. Two players, A and B, are playing a game with a special coin which, when tossed, is heads up with probability 2/3 and tails up with probability 1/3. The rules are as follows:

- First A chooses one pattern from the four patterns HH, HT, TH and TT.
- Next B chooses one pattern different from the one A chose.
- Then they toss the coin until one of their chosen patterns appear. That person will win the game.

For example, if A chose HH and B chose TT, and by tossing the coin they got HTHTHTT, then the winner would be B.

Now suppose A and B are both smart enough. Then what is the probability that A wins the game? What pattern should A choose? What pattern should B choose?

Solution. We compute probabilities that one of the possible 4 outcomes appears earlier than another.

First compute $P(HH < HT)$. This happens iff the beginning is HH, or THH, or TTHH, etc. So the answer is

$$\frac{4}{9}(1 + \frac{1}{3} + \frac{1}{9} + ...) = \frac{2}{3}.$$

Now compute $P(HH < TH)$. This happens iff the beginning is HH, so the answer is 4/9.

Now compute $P(HH < TT)$. This happens iff the beginning is HH, or HTHH, or HTHTHH,... or THH, THTHH, THTHTHH, etc. This gives

$$(1+\frac{1}{3})(1+\frac{2}{9}+\frac{2^2}{9^2}+...)\frac{4}{9}=\frac{16}{21}.$$

Now compute $P(HT < TT)$. This happens iff the beginning is H, or TH, so the probability is $2/3 + 2/9 = 8/9$.

Now compute $P(TH < TT)$. This happens iff the beginning is TH or HTH or HHTH or HHHTH, etc. So the answer is

$$(1+\frac{2}{3}+\frac{2^2}{3^2}+...)\frac{2}{9}=\frac{2}{3}.$$

Finally, compute $P(HT < TH)$. This happens iff the beginning is H, so the probability is 2/3.

Thus, A should select HH and B should select TH, and the probability of A winning will be 4/9.

Problem M6.

Let V be a subset of a finite field L which is closed under addition (i.e, a subgroup of the additive group of L).

(a) Show that the polynomial $\prod_{v\in V}(X+v)\in L[X]$ is a p-polynomial (that is, an L-linear combination of the monomials $X^{p^0}=X, X^{p^1}=X^p, X^{p^2}, ...$).

(b) Let $t\in L\setminus V$. Prove that

$$\sum_{v\in V}\frac{1}{t+v}=\left(\prod_{v\in V}\frac{1}{t+v}\right)\cdot\left(\prod_{v\in V\setminus 0}v\right).$$

Remark

Part **(a)** is a known fact (e.g., it immediately follows from [Conrad14, Theorem A.1 2) and Corollary A.3]). Part **(b)** is a lemma from unfinished work of Darij Grinberg and James Borger on Carlitz polynomials.

Solution sketch

We WLOG assume that $V\neq 0$ (since otherwise, the statements are evident). Let $p=\operatorname{char}L$. Since V is finite, every element of V has finite order. This, combined with $V\neq 0$, easily shows that p is positive, so that p is a prime. Then, L is a field extension of $\mathbb{F}_p$, and V is a finite-dimensional $\mathbb{F}_p$-vector subspace.

Let W be the polynomial $\prod_{v\in V}(X+v)\in L[X]$.

(a) We need to prove that W is a p-polynomial.

For every finite-dimensional $\mathbb{F}_p$-vector subspace A of L, let W_A denote the polynomial $\prod_{v\in A}(X+v)\in L[X]$. Then, $W=W_V$. Hence, it is enough to show that W_A is a p-polynomial for every finite-dimensional $\mathbb{F}_p$-vector subspace A of L.

It is easy to show that if A is a finite-dimensional $\mathbb{F}_p$-vector subspace of L, and if B and C are two $\mathbb{F}_p$-vector subspaces of A satisfying $A = B \oplus C$, then $W_A = W_B \circ W_C$, where the sign $\circ$ denotes composition of polynomials (i.e., for two univariate polynomials P and Q, we set $P \circ Q = P(Q)$). Hence, in this situation, in order to prove that W_A is a p-polynomial, it is enough to show that W_B and W_C are p-polynomials (since the composition of two p-polynomials over a field of characteristic p is a p-polynomial again). Hence, we can reduce our goal to smaller subgoals, until at the end it remains to prove that W_A is a p-polynomial whenever A is an $\mathbb{F}_p$-vector subspace A of L of dimension ≤ 1 (since every finite-dimensional $\mathbb{F}_p$-vector space is a direct sum of finitely many such subspaces). But this is easy to check: If A is of dimension 0, then $W_A = X$; and if A is of dimension 1, then $A = \mathbb{F}_p w$ for some nonzero vector $w \in L$, and therefore

$$\begin{aligned} W_A &= W_{\mathbb{F}_p w} = \prod_{v \in \mathbb{F}_p w} (X+v) = \prod_{x \in \mathbb{F}_p} (X + xw) = (X+0w)(X+1w)\dots(X+(p-1)w) \\ &= X^p - w^{p-1}X \qquad \left(\begin{array}{c} \text{by the homogenization of the polynomial identity} \\ (X+0)(X+1)\dots(X+(p-1)) = X^p - X \text{ in } \mathbb{F}_p[X] \end{array} \right), \end{aligned}$$

which is a p-polynomial. Thus, Problem 2 **(a)** is solved.

(b) Part **(a)** yields that the polynomial W is a p-polynomial. Hence, its derivative equals its coefficient in front of X^1 (because the derivative of any p-polynomial in characteristic p equals its coefficient in front of X^1). But this coefficient is $\prod_{v \in V \setminus 0} v$. Thus, we know that the derivative of W equals $\prod_{v \in V \setminus 0} v$. Hence, $W'(t) = \prod_{v \in V \setminus 0} v$.

On the other hand, since $W = \prod_{v \in V} (X+v)$, the Leibniz formula yields

$$\begin{aligned} W' &= \sum_{w \in V} \underbrace{(X+w)'}_{=1} \cdot \prod_{\substack{v \in V; \\ v \neq w}} (X+v) = \sum_{w \in V} \prod_{\substack{v \in V; \\ v \neq w}} (X+v) = \sum_{w \in V} \frac{\prod_{v \in V} (X+v)}{X+w} \\ &= \left(\prod_{v \in V} (X+v) \right) \cdot \left(\sum_{w \in V} \frac{1}{X+w} \right). \end{aligned}$$

Applying this to $X = t$, we obtain

$$W'(t) = \left(\prod_{v \in V} (t + v) \right) \cdot \left(\sum_{w \in V} \frac{1}{t + w} \right),$$

so that

$$\sum_{w \in V} \frac{1}{t + w} = \underbrace{\frac{1}{\prod\limits_{v \in V} (t + v)}}_{= \prod_{v \in V} \frac{1}{t+v}} \cdot \underbrace{W'(t)}_{= \prod_{v \in V \setminus 0} v} = \left(\prod_{v \in V} \frac{1}{t + v} \right) \cdot \left(\prod_{v \in V \setminus 0} v \right).$$

Rename the index w as v and obtain the claim of Problem 2 **(b)**.

References

[Stanley11] Richard Stanley, *Enumerative Combinatorics, volume 1*, Cambridge University Press, 2011.

[Conrad14] Keith Conrad, *Carlitz extensions*