Solution to IMO 2019 shortlisted problems.

Anzo Teh

December 31, 2020

Algebra

A1. (IMO 1) Let $\mathbb{Z}$ be the set of integers. Determine all functions $f : \mathbb{Z} \to \mathbb{Z}$ such that, for all integers a and b,

$$f(2a) + 2f(b) = f(f(a+b)).$$

Answer. $f \equiv 0$ (the zero function) or $f(n) = 2n + c$ for all $n \in \mathbb{Z}$ with c an arbitrary integer constant. It's east to verify that these functions work.

Solution. We consider the following for each $n \in \mathbb{Z}$, substituting $(a, b) = (0, n)$ and $(1, n-1)$:

$$f(2(0)) + 2f(n) = f(f(n)) = f(2(1)) + 2f(n-1)$$

which means $f(n) - f(n-1) = \frac{f(2)-f(0)}{2}$, which is fixed. As f is defined only on $\mathbb{Z}$, we can dedue right away that f is linear.

Now write $f(n) = mn + c$ with m being the gradient and c being the intercept of linear equation, the desired equation now becomes:

$$m(2a) + c + 2(mb + c) = f(m(a+b) + c) = m(m(a+b) + c) + c$$

or rather, $2m(a+b) + 3c = m^2(a+b) + (m+1)c$ for all a and b. Pushing everything to one side we get

$$(2m - m^2)(a+b) + (3 - m - 1)c = 0$$

Since this must be true for all integers a and b, $m(2-m) = 2m - m^2 = 0$ which means $m = 0$ or $m = 2$. If $m = 0$ then $0 = (3-1)c = 2c$ so $c = 0$. If $m = 2$ then there's no additional restriction on c. This completes the proof.

A2. Let $u_1, u_2, \ldots, u_{2019}$ be real numbers satisfying

$$u_1 + u_2 + \cdots + u_{2019} = 0 \quad \text{and} \quad u_1^2 + u_2^2 + \cdots + u_{2019}^2 = 1.$$

Let $a = \min(u_1, u_2, \ldots, u_{2019})$ and $b = \max(u_1, u_2, \ldots, u_{2019})$. Prove that

$$ab \leqslant -\frac{1}{2019}.$$

Solution. W.l.o.g. let $u_1 \le \cdots \le u_{2019}$, and let k be such that $u_k \le 0$ but $u_{k+1} > 0$. Also let:

$$b_i = u_{k+i}, i = 1, \cdots, 2019 - k \qquad a_i = |u_{k+1-i}|, i = 1, \cdots, k$$

then we have the relation

$$\sum a_i = \sum b_i \qquad \sum a_i^2 + \sum b_i^2 = 1$$

We thus have $a = -\max(a_i)$ and $b = \max(b_i)$ Denote $S = \sum a_i$, and $\ell = 2019 - k$. Then $S = \sum a_i \le |a|k$ and $S = \sum b_i \le b(2019 - k) = b\ell$. This then gives

$$2019 = k + \ell \ge \frac{S}{|a|} + \frac{S}{b} = S(\frac{|a| + b}{|a|b})$$

Also, given that $a_i \le |a|$ for each a_i and $b_i \le b$ for each b_i,

$$\sum a_i^2 \le |a| \sum a_i = |a|S \qquad \sum b_i^2 \le b \sum b_i = bS$$

so $1 \le S(|a| + b)$. This gives

$$|a|b = S(|a| + b) \div S(\frac{|a| + b}{|a|b}) \ge \frac{1}{2019}$$

and since $a < 0$ (given that $u_i \le 0$ for all $i \le k$), we have $ab \le -\frac{1}{2019}$ as desired.

A3. Let $n \geqslant 3$ be a positive integer and let $(a_1, a_2, \ldots, a_n)$ be a strictly increasing sequence of n positive real numbers with sum equal to 2. Let X be a subset of $\{1, 2, \ldots, n\}$ such that the value of

$$\left|1 - \sum_{i \in X} a_i\right|$$

is minimised. Prove that there exists a strictly increasing sequence of n positive real numbers $(b_1, b_2, \ldots, b_n)$ with sum equal to 2 such that

$$\sum_{i \in X} b_i = 1.$$

Solution. Denote $Y = \{1, 2, \cdots, n\} \backslash X$. By multiplying the numbers by a constant factor, we may assume that we're given the condition such that

$$\left|\sum_{i=1}^{n} a_i/2 - \sum_{i \in X} a_i\right|$$

is minimized. This is also the same as minimizing

$$\left|\sum_{i \in Y}^{n} a_i - \sum_{i \in X} a_i\right|$$

(which has double the magnitude of the previous expression). The sequence $(b_1, \cdots, b_n)$ must satisfy

$$\sum_{i \in X} b_i = \sum_{i=1}^{n} b_i/2$$

which is equivalent to saying

$$\sum_{i \in X} b_i = \sum_{i \in Y} b_i$$

We say that $Y > X$ if, whenever $(a_1, \cdots, a_n)$ are positive reals and strictly increasing, we have

$$\sum_{i \in Y} a_i > \sum_{i \in X} a_i$$

and $X > Y$ is defined similarly. We also define $X \sim Y$ if neither $X > Y$ nor $Y > X$ holds.

We first claim that, if $X \sim Y$, then there exists a sequence of positive reals, strictly increasing, that

$$\sum_{i \in Y} a_i = \sum_{i \in X} a_i$$

By the relation $X \sim Y$, we know that there exist sequences $(a_1, \cdots, a_n)$ and $(b_1, \cdots, b_n)$ such that

$$\sum_{i \in Y} a_i \geq \sum_{i \in X} a_i \quad \sum_{i \in Y} b_i \leq \sum_{i \in X} b_i$$

The idea is that, we want to adjust a_i in a continuous manner such that $a_i = b_i$ in the end but $\sum_{i \in Y} a_i - \sum_{i \in X} a_i$ changes sign (from nonnegative to nonpositive), so the claim follows from intermediate value theorem.

We'll choose i in the following manner: whenever exist, choose the minimum i such that $a_i > b_i$. If $i = 1$, we can just decrease a_i to become b_i, easily (while maitning that a_i is strictly increasing). Otherwise, $a_i > b_i > b_{i-1} \geq a_{i-1}$ by the minimality of i, so we can still safely decrease a_i to become b_i. Once we're done with this, we do the opposite: choose the maximum i with $a_i < b_i$ and increase a_i. This will ensure that our moves are continuous without breaking the constraint that it's increasing.

Next, we need a characterization of $X > Y$. We claim that $X > Y$ if the following are true: arrange the numbers in X in the following order: $x_1 > x_2 > \cdots > x_{|X|}$ and Y in the following order: $y_1 > \cdots > y_{|Y|}$. Then $X > Y$ if and only if the following are satisfied:

- $|X| \geq |Y|$
- For all $i = 1, \cdots, |Y|$, $x_i > y_i$.

First, notice that if this is true, then:

$$\sum_{i \in X} a_i \geq \sum_{i=1}^{|Y|} a_{x_i} > \sum_{i=1}^{|Y|} a_{y_i} = \sum_{i \in Y} a_i$$

Now, we show that if $X > Y$ then the condition above must hold true. If $|X| < |Y|$ then we can consider the sequence $a_i = a + i\epsilon$ for some $\epsilon > 0$ and fixed a, then there will be constants x and y such that $\sum_{i \in X} a_i = |X|a + x\epsilon$ and $\sum_{i \in Y} a_i = |Y|a + y\epsilon$. By choosing ϵ sufficiently small, and using $|X| < |Y|$, we can make $\sum_{i \in X} a_i < \sum_{i \in Y} a_i$.

Otherwise, if there exists k with $x_k < y_k$ (we notice that $x_i \neq y_i$ since X and Y are disjoint) then we can say that there exists k numbers in Y that's at least y_i but at most $k-1$ in X (or even, exactly $k-1$ if we take the minimum such k) that's at least y_i. Consider now the following structure: $a_i = \begin{cases} \epsilon i : i < y_i \\ a + \epsilon i : i \geq y_i \end{cases}$ then there exist constants x, y, k' with $k' \leq k-1$ such that $\sum_{i \in X} a_i = x\epsilon + k'a$ and $\sum_{i \in Y} a_i = y\epsilon + ka$ and by getting a sufficiently large, we have $\sum_{i \in Y} a_i > \sum_{i \in X} a_i$, so $X > Y$ cannot hold.

Finally, we show that, if the value of $|\sum\limits_{i \in Y} a_i - \sum\limits_{i \in X} a_i|$ is minimized, then $X \sim Y$ must hold. Suppose on the contrary that $Y > X$. Then $\sum\limits_{i \in Y} a_i - \sum\limits_{i \in X} a_i$ is positive. If X is empty, consider $Y' = \{2, \cdots, n\}$ and $X' = \{1\}$, then

$$\sum_{i \in Y} a_i - \sum_{i \in X} a_i > \sum_{i \in Y'} a_i - \sum_{i \in X'} a_i > 0$$

which contradicts the minimality of our set. Now suppose that X and Y are both nonempty. Then we can arrange $x_1 > x_2 > \cdots x_|X| \in X$ and $y_1 > y_2 > \cdots y_|Y| \in Y$ and exchange x_1 and y_1 to yield X' and Y', respectively. Then since $y_1 > x_1$, we have

$$\sum_{i \in Y} a_i - \sum_{i \in X} a_i > \sum_{i \in Y'} a_i - \sum_{i \in X'} a_i$$

if the second one is positive, then we get our desired contradiction. Otherwise, $\sum\limits_{i \in Y'} a_i - \sum\limits_{i \in X'} a_i \leq 0$ and so $|\sum\limits_{i \in Y'} a_i - \sum\limits_{i \in X'} a_i| = -(\sum\limits_{i \in Y'} a_i - \sum\limits_{i \in X'} a_i)$. Therefore,

$$|\sum_{i \in Y} a_i - \sum_{i \in X} a_i| - |\sum_{i \in Y'} a_i - \sum_{i \in X'} a_i| = (\sum_{i \in Y} a_i - \sum_{i \in X} a_i) + (\sum_{i \in Y'} a_i - \sum_{i \in X'} a_i) = 2(\sum_{i \in Y \setminus \{y_1\}} a_i - \sum_{i \in X \setminus \{x_1\}} a_i)$$

but since $n \geq 3$, we have $|Y| \geq 2$ and therefore we have $\sum\limits_{i \in Y \setminus \{y_1\}} a_i - \sum\limits_{i \in X \setminus \{x_1\}} a_i > 0$. This still contradicts the minimality of the difference in sum of sets.

Therefore we have $X \sim Y$ and the required b_i can be found such that

$$\sum_{i \in Y} b_i = \sum_{i \in X} b_i$$

A4. Let $n \geqslant 2$ be a positive integer and $a_1, a_2, \ldots, a_n$ be real numbers such that

$$a_1 + a_2 + \cdots + a_n = 0.$$

Define the set A by

$$A = \{(i,j) \,|\, 1 \leqslant i < j \leqslant n, |a_i - a_j| \geqslant 1\}$$

Prove that, if A is not empty, then

$$\sum_{(i,j)\in A} a_i a_j < 0.$$

Solution. Denote $B = \{1, \cdots, n\}^2 \backslash A =$

$$\{(i,j) \,|\, 1 \leqslant i < j \leqslant n, |a_i - a_j| < 1\}$$

We first notice that

$$\sum_{(i,j)\in A\cup B} a_i a_j = (\sum_{i=1}^{n} a_i)^2 = 0$$

so it suffices to prove that $\sum_{(i,j)\in B} a_i a_j > 0$ given that A is nonempty.

Denote, also, $B_0 = \{(i,j) \in B : \max\{|a_i|, |a_j|\} \geq 1\}$ and $B_1 = \{(i,j) \in B : \max\{|a_i|, |a_j|\} < 1\}$, then $B = B_0 \dot\cup B_1$ (disjoint union). If $|a_i| \geq 1$ and $|a_i - a_j| < 1$, then by triangle inequality a_i and a_j must have the same sign, and therefore $a_i a_j > 0$. This gives $\sum_{(i,j)\in B_0} a_i a_j \geq 0$ as each term is positive (with equality if and only if B_0 is empty). It remains to show that $\sum_{(i,j)\in B_1} a_i a_j \geq 0$.

Consider, now, the subset of numbers $C = \{i : 1 \leq i \leq n, |a_i| < 1\}$. We now have $C^2 = B_1 \dot\cup \{(i,j) \notin B : \max\{|a_i|, |a_j|\} < 1\}$. The second set $\{(i,j) \notin B : \max\{|a_i|, |a_j|\} < 1\}$ (namely B_2) is the same as $\{(i,j) : \max\{|a_i|, |a_j|\} < 1, |a_i - a_j| \geq 1\}$. We notice that:

$$\sum_{(i,j)\in C^2} a_i a_j = \sum_{(i,j)\in B_1} a_i a_j + \sum_{(i,j)\in B_2} a_i a_j$$

We notice that $\sum_{(i,j)\in C^2} a_i a_j = (\sum_{i\in C} a_i)^2 \geq 0$. In B_2, $|a_i|, |a_j| < 1$ but $|a_i - a_j| \geq 1$ means both are different signs (and nonzero), hence $a_i a_j < 0$. Therefore, $\sum_{(i,j)\in B_2} a_i a_j \leq 0$ with equality if and only if B_2 is empty. This therefore gives $\sum_{(i,j)\in B_1} a_i a_j \geq 0$, with equality iff $\sum_{i\in C} a_i = 0$ and B_2 empty.

We therefore have $\sum_{(i,j)\in B} a_i a_j = \sum_{(i,j)\in B_0} a_i a_j + \sum_{(i,j)\in B_1} a_i a_j \geq 0$ since both sums are nonnegative. Equality holds iff both sums are 0. This happens only when each of the following are satisfied:

- $B_0 = \{(i,j) \in B : \max\{|a_i|, |a_j|\} \geq 1\} = \emptyset$
- $B_2 = \{(i,j) : \max\{|a_i|, |a_j|\} < 1, |a_i - a_j| \geq 1\} = \emptyset$
- $\sum_{|a_i|<1} a_i = 0$

If $|a_i| \geq 1$, then $(i,i) \in B$ anf therefore $(i,i) \in B_0$. For B_0 to be empty, we need $|a_i| < 1$ for all i. This will mean $B_2 = A$. But since A is nonempty, B_2 cannot be nonempty. This means $\sum_{(i,j)\in B} a_i a_j > 0$, i.e. inequality must be strict.

Combinatorics

C1. The infinite sequence $a_0, a_1, a_2, \dots$ of (not necessarily distinct) integers has the following properties: $0 \le a_i \le i$ for all integers $i \ge 0$, and

$$\binom{k}{a_0} + \binom{k}{a_1} + \cdots + \binom{k}{a_k} = 2^k$$

for all integers $k \ge 0$. Prove that all integers $N \ge 0$ occur in the sequence (that is, for all $N \ge 0$, there exists $i \ge 0$ with $a_i = N$).

Solution. We show that the following invariants are maintained for each k: among $a_0, \cdots, a_n$, we have:

- For each i, exactly two of them in $\{i, n-i\}$ if $2i \neq n$, and exactly one of them is i if $2i = n$.
- If there exists $x \neq y$ and p with $a_x = a_y = p$ then $\max\{x, y\} \ge 2p + 1$.
- More generally, if p occurs 2 times in $a_0, \cdots, a_n$, then so does $p - 1$ (and hence, iteratively, $p-2, p-3, \cdots, 0$).

If we have these, then for each positive integer N, exactly one of $a_0, \cdots, a_{2N}$ is equal to N, proving the problem conclusion.

When $k = 0$, we must have $a_0 = 0$ so both invariants are maintained. Now, suppose that for some n what we've written above is true for $a_0, \cdots, a_k$ for all $k = 0, 1, \cdots, n$. We now consider the possible choices for a_{n+1}.

We first show that for each i, at most two of $a_0, \cdots, a_n$ in $\{i, n+i-1\}$ if $2i \neq n+1$, and at most one of them is i if $2i = n+1$. Suppose otherwise, then for some i, we have among $a_0, \cdots, a_n$:

- $2i \neq n + 1$ and (at least) three of them in $\{i, n+1-i\}$
- $2i = n + 1$ and (at least) two of them equal to i.

The second case can't occur since otherwise, the second invariant says that one of the indices must be at least $2i + 1 \neq n + 2$ but we are considering $a_0, \cdots, a_n$. If the first case were to occur, then either i or $n + 1 - i$ occurs two times. Without loss of generality we assume that i occurs two times (it can't occur three times otherwise the first invariant would have been broken by $a_0, \cdots, a_n$). This means that i occurs two times and $n+1-i$ occurs (at least) one time. By the third invariant, we get $0, 1, \cdots, i-1$ each occurs two times in the sequence $a_0, \cdots, a_n$. Therefore since $i-1$ occurs two times, $n-(i-1) = n+1-i$ cannot occur due to first invariant on $a_0, \cdots, a_n$, contradiction.

This means we can prove the first invariant for $a_0, \cdots, n+1$. To see why, we have

- $2i \neq n + 1$ and at most two of them in $\{i, n+1-i\}$
- $2i = n + 1$ and at most one of them equal to i.

Given that we're operating in $a_0, \cdots, a_n$ (instead of a_{n+1}), all but one of the above will be equality (with the odd one out being one less than equality). Let i_{n+1} be that quantity (i.e. $2i_{n+1} \neq n+1$ and one of $a_0, \cdots, a_n$ in $\{i_{n+1}, n+1-i_{n+1}\}$, or $2i_{n+1} = n+1$ and this i is not part of the sequence). By considering the binomial sum, we need $\binom{n+1}{a_{n+1}} = \binom{n+1}{i_{n+1}}$ so $a_{n+1} \in \{i_{n+1}, n+1-i_{n+1}\}$, as desired, establishing the first invariant.

The third invariant is simply a generalization of the second one (well if $a_x = a_y = p$ then each of $0, \cdots, p$ occur two times so $\max\{x, y\} \ge 2(p+1) - 1 = 2p + 1$), so we'll straight prove this third invariant. If the resulting a_{n+1} has $a_i \neq a_{n+1}$ for all i we are done.

Otherwise, suppose that $p = a_i = a_{n+1}$ (i.e. p occurs two times in this example). As the first invariant has been maintained for $n+1$, we cannot have $2p = n+1$. If $n+1 < 2p$, then $n+1-p$ is not in $a_0, \cdots, a_{n+1}$, and $2(n+1-p) < n+1$. Considering $a_0, \cdots, a_{2(n+1-p)}$, where all the invariants are maintained till this stage, we have $n+1-p$ occuring exactly once, which is a contradiction. Thus $n+1 > 2p$.

Finally to show the third invariant, we recall that $n+1-p \notin a_0, \cdots, a_{n+1}$. Now that this invariant is maintained on $a_0, \cdots, a_{n-i}$ for each $i \geq 0$, we know that:

- Either $2(n+1-p) = n-i$ and $n+1-p$ occurs once;
- or $2(n+1-p) \neq n-i$ and $n-i-(n+1-p) = p-i-1$ occurs twice.

The first case cannot happen as $n+1-p \notin a_0, \cdots, a_{n+1}$, so $p-i-1$ does occur twice, finishing the third invariant.

C2. You are given a set of n blocks, each weighing at least 1; their total weight is $2n$. Prove that for every real number r with $0 \leq r \leq 2n-2$ you can choose a subset of the blocks whose total weight is at least r but at most $r+2$.

Solution. We relax the condition a bit to show that the conclusion holds when we have n blocks of weight at least 1 each, with total weight at most $2n$ (instead of exactly $2n$). We will approach the problem via induction on n, where $n = 1$ is clear.

Inductive step: Suppose we have $n+1$ blocks $a_1 \leq \cdots \leq a_{n+1}$, our conclusion holds for $a_1, \cdots, a_n$ and if $S = a_1 + \cdots a_n$ then by induction hypothesis the claim would hold for $0 \leq r \leq S-2$ (where we can even extend it to $0 \leq r \leq S$ as we can simply take $\{a_1, \cdots, a_n\}$ for a total weight of S.

Therefore all it remains to show the same for $S \leq r \leq S + a_{n+1} - 2$. To determine if it's possible to choose a subset of blocks with total weight in $[r, r+2]$ where $r > S$, by choosing the element a_{n+1}, we're left with choosing a subset from $\{a_1, \cdots, a_n\}$ with sum in $[r - a_{n+1}, r - a_{n+1} + 2]$. Considering the empty set (weight sum 0), and by hypothesis, it then suffices to show that $r - a_{n+1} + 2 \geq 0$, and that $r - a_{n+1} \leq S$. Thus $a_{n+1} \leq n+2$.

The fact that $r - a_{n+1} \leq S$ follows from $r \leq S + a_{n+1}$. In addition, given that $S \geq n$ and $S + a_{n+1} \leq 2(n+1)$, we have $r - a_{n+1} + 2 \geq S + 1 - a_{n+1} + 2 \geq n + 3 - (n+2) = 1$, as desired.

C3. (IMO 5) The Bank of Bath issues coins with an H on one side and a T on the other. Harry has n of these coins arranged in a line from left to right. He repeatedly performs the following operation: if there are exactly $k > 0$ coins showing H, then he turns over the kth coin from the left; otherwise, all coins show T and he stops. For example, if $n = 3$ the process starting with the configuration THT would be $THT \to HHT \to HTT \to TTT$, which stops after three operations.

(a) Show that, for each initial configuration, Harry stops after a finite number of operations.

(b) For each initial configuration C, let $L(C)$ be the number of operations before Harry stops. For example, $L(THT) = 3$ and $L(TTT) = 0$. Determine the average value of $L(C)$ over all 2^n possible initial configurations C.

Answer. The expected value, $E(L(C)) = \frac{n(n+1)}{4}$.
Solution. Let's induce on n and denote $f(n)$ as the answer for n. It's obvious that when $n = 1$, $E(C) = \frac{0+1}{2} = \frac{1}{2}$, corresponding to the configuration T and H, respectively. We'll show that $f(n+1) = f(n) + \frac{n+1}{2}$ which completes the solution. For clarity sake, denote $op(C)$ as the configuration by applying the operation on C.

Let A_0 be the set of sequence of $n+1$ coins ending on a T, and A_1 be the set of sequence of $n+1$ coins ending on a H. Also, for each set S of sequences of coins let $L(S)$ be the expected value of number of operations needed (that is, $L(S) = \frac{1}{|S|}\sum_{C\in S} L(C)$). We have the relation $f(n+1) = \frac{L(A_0)+L(A_1)}{2}$ since A_0 and A_1 have the same size, each being 2^n.

Let B_0 be the set of sequence of n coins. We first define a mapping γ from A_0 to B_0 as follows: for each $C \in A_0$, let C' be the sequence in B_0 that is C with the last T removed. Then $\gamma(C) = C'$. This mapping is clearly bijective. We claim that $\gamma(op(C)) = op(\gamma(C))$ if C has at least a head. Notice that C and C' each has the same number of heads, say k, so in the operation, the k-th coin from the left in each C and $C' = \gamma(C)$ are flipped. Since the first n coins of C and $\gamma(C)$ are the same, so are the first n coins of $op(C)$ and $op(\gamma(C))$. Since $k \le n$, the last tail of C is not flipped in getting into $op(C)$, hence $\gamma(C) = op(\gamma(C))$. Continuing this iteration, we find that $L(C) = L(C')$ and therefore aggregating this for all configurations $C \in A_0$ we have $L(A_0) = L(B_0) = f(n)$.

Computing $L(A_1)$ is trickier. Again, define a mapping σ from A_1 to B_0 such that for each C in A_1, $\sigma(C)$ is obtainly by the following algorithm: take C, drop the last H, invert the coin sequence and flip all the coins (formally, $\sigma(s_0 s_1 \cdots s_{n-1} H) = \overline{s_{n-1} \cdots s_1 s_0}$ where $\overline{H} = T$ and $\overline{T} = H$). Again this mapping is bijective. We claim that if C has at least a tail, then $\sigma(op(C)) = op(\sigma(C))$. Let $1 \le k \le n$ be the number of heads in C, and let $C = s_0 s_1 \cdots s_{n-1} H$. Then $op(C) = s_0 \cdots \overline{s_{k-1}} s_k \cdots s_{n-1} H$. Excluding the last head, C actually has $k-1$ heads and $n-k+1$ tails, which means that $\sigma(C) = \overline{s_{n-1} \cdots s_1 s_0}$ has $n-k+1$ heads and $k-1$ tails, which means that $op(\sigma(C)) = \overline{s_{n-1} \cdots \overline{s_{k-1}} \cdots s_1 s_0} = \overline{s_{n-1} \cdots} s_{k-1} \overline{\cdots s_1 s_0}$, which establishes that $\sigma(op(C)) = op(\sigma(C))$. If m is the number operations needed for $\sigma(C)$ to reach all T, $op^m(\sigma(C)) = TT \cdots T$ and from above we can deduce that $op^m(\sigma(C)) = \sigma(op^m(C)) = TT \cdots T$, we have $op^m(C) = HH \cdots H$ (i.e. $n+1$ H's). It's now not hard to see that the subsequent moves on $op^m(C)$ are to flip the rightmost possible H, and there are $n+1$ of them, hence $L(C) = m + n + 1 = L(\sigma(C)) + n + 1$. Aggregating this over all C in A_1 we get $L(A_1) = L(B_0) + n + 1 = f(n) + n + 1$.

Summarizing above we have $f(n+1) = \frac{L(A_0)+L(A_1)}{2} = \frac{f(n)+f(n)+n+1}{2} = f(n) + \frac{n+1}{2}$, as desired.

C4. On a flat plane in Camelot, King Arthur builds a labyrinth $\mathfrak{L}$ consisting of n walls, each of which is an infinite straight line. No two walls are parallel, and no three walls have a common point. Merlin then paints one side of each wall entirely red and the other side entirely blue.

At the intersection of two walls there are four corners: two diagonally opposite corners where a red side and a blue side meet, one corner where two red sides meet, and one corner where two blue sides meet. At each such intersection, there is a two-way door connecting the two diagonally opposite corners at which sides of different colours meet.

After Merlin paints the walls, Morgana then places some knights in the labyrinth. The knights can walk through doors, but cannot walk through walls.

Let $k(\mathfrak{L})$ be the largest number k such that, no matter how Merlin paints the labyrinth $\mathfrak{L}$, Morgana can always place at least k knights such that no two of them can ever meet. For each n, what are all possible values for $k(\mathfrak{L})$, where $\mathfrak{L}$ is a labyrinth with n walls?

Answer. The only possible answer is $k(\mathfrak{L}) = n + 1$.

Solution. Given n lines on the plane with no two parallel and no three concurrent, we can model the configuration as regions on the plane. The number of such regions (which can be established via induction on n) is $\frac{n(n+1)}{2} + 1$ (adding one line to the existing configuration of n lines will split $n+1$ regions into two). If we model this configuration as an undirected graph with regions as vertices and doors as edges, then the maximum number of knights that can be placed where no two knights can ever meet is the number of connected components in the graph.

We first notice that there's exactly one door corresponding to each pair of lines, giving $\binom{n}{2} = \frac{n(n-1)}{2}$ edges. Therefore the number of connected components is at least

$$|V| - |E| = \frac{n(n+1)}{2} + 1 - \frac{n(n-1)}{2} = n + 1$$

with equality if and only if no cycle exists.

It now remains to show that no matter how the n lines are placed in the plane, Merlin can always colour the walls such that exactly $n+1$ connected components remain. Equivalently, the colouring should result in no cycle (when seen as a graph).

We establish this colouring before showing that it will work (motivation: try out $n = 3$). Since there are $\binom{n}{2}$ intersection points (i.e. finitely many of them), there exists a circle (say, $\mathcal{C}$) such that all these intersection points lie inside the circle (e.g. choose a reference point, say origin, and consider M as the furthest of the $\binom{n}{2}$ points from the circle. Then make the radius of the circle $M + 1$). These n lines intersect the circle forming n chords. Since all the n chords intersect each other within the circle, we can label the endpoints as $a_1, a_2, \cdots, a_n, b_1, b_2, \cdots, b_n$ such that:

- $a_1, \cdots, a_n, b_1, \cdots, b_n$ are clockwise in that order (and b_n meets a_1 again).
- $a_i b_i$ is the chord corresponding to line i.

Going clockwise, we colour all a_i such that red comes before blue in clockwise order (and therefore for b_i blue comes before red in clockwise order).

Now, each region may have finite or infinite area; the former corresponds to convex polygons; the latter corresponds to regions that need to be bounded by $\mathcal{C}$. for each region R, denote $f(R)$ as the number of lines such that this region is in the *blue* side of the line (in other words we make red $= 0$ and blue $= 1$). Therefore, if $R_{(a_i, a_{i+1})}$ is the region facing the arc (i.e. infinite if not bounded by the arc) that's 'adjacent' to a_i and a_{i+1} (and define similarly for $R_{(b_i, b_{i+1})}$), then:

- $f(R_{(b_n, a_1)}) = 0$
- $f(R_{(a_i, a_{i+1})}) = i$
- $f(R_{(a_n, b_1)}) = n$
- $f(R_{(b_i, b_{i+1})}) = n - i$

Moreover, any two regions R_1, R_2 connected by a door must have $f(R_1) = f(R_2)$.

We see that $R_{(b_n, a_1)}$ and $R_{(a_n, b_1)}$ is not connected to each door because to have $f(\cdot) = 0$ means it's on the red side of all lines and there's only one possible region that can satisfy this; similar case for regions with $f(\cdot) = n$. We now show that:

- for each $1 \le i \le n - 1$, $R_{(a_i, a_{i+1})}$ and $R_{(b_i, b_{i+1})}$ each connected to 1 door
- for all other regions (i.e. those with finite areas) they all each connected to 2 doors.

We first establish the lower bounds that they have *at least* 1 or 2 doors. For $R_{(a_i, a_{i+1})}$, it's adjacent to the blue side of a_i and red side of a_{i+1}. Considering all the sides of that region, (excluding the circular arc), the fact that it's at different side of a_i and a_{i+1} means there must be two adjacent sides of that region that are of different colours. This gives a door to this region.

For the finite region, recall that it's a convex polygon P. We consider the colour of the sides that enclose the region (that is, the internal part). As long as it consists of both red and blue sides, by considering one cycle clockwise there will be instances that goes from red to blue and then blue to red, thereby giving the two doors. It then suffices to show that the region cannot be all red or all blue.

We first show that we can choose three lines from the sides of P that will define a triangle that convers the whole P. If P is already a triangle then we are done. Otherwise, choose two non-adjacent sides of P. These two sides partition the plane into four regions; we consider the one region that contains both the segments of P (and therefore covering the whole P). These two sides of P also partitions the other sides into two consecutive sections; choosing a side from one of the to complete the triangle will result in P not covered by the triangle at all (except adjacent sides); choosing one from the other side will cover P completely. This gives us the required triangle T that covers P completely.

Notice that T has three lines x, y, z with endpoints $a_x, a_y, a_z, b_x, b_y, b_z$. By considering the function f with respect to this T, the region with $f(\cdot) = 0$ and 3 have both been taken by the infinite regions, so $f(T)$ is either 1 or 2, thus cannot have inner walls being all blue or all red. Since the inner walls of T is also a subset of inner walls of P (since T covers P), it follows that the inner walls of P cannot be all red or all blue, as claimed.

Finally, to turn inequality into equality, we have $2n$ regions of infinite area, and $\frac{n(n+1)}{2} + 1 - 2n = \frac{n(n-3)}{2} + 1$ regions with finite area (for $n \geq 1$). $2n - 2$ of the inifinite region have doors, so the number of correspondence to doors is at least

$$2n - 2 + 2(\frac{n(n-3)}{2} + 1) = 2n - 2 + n(n-3) + 2 = n(n-1)$$

Each door corresponds to 2 regions and there are $\frac{n(n-1)}{2}$ doors, and therefore equality must hold. This means (in the language of graph theory) that 2 regions have degree 0, $2n - 2$ regions have degree 1, and the rest have degree 2. This gives rise of $n - 1$ simple paths (correspond to those with $f(\cdot) = 1, 2, \cdots, n-1$) and several other cycles, each disjoint from each other. We also know that each path must start and end with unbounded regions (with degree 1). Notice that the paths and cycles (if any) are also disjoint in the 'literal' sense: for each finite region we can draw a line (or anything within the polygonal region itself) between the two vertices connected to a door, and for infinite region we can draw a line from the door to anywhere on the arc on P that still corresponds to the region (i.e. don't cross over the region).

Suppose a cycle does exist. Since the cycles and paths are disjoint, we can define an ordering based on whether one cycle encloses the other. Since the number of cycles must be finite, we can choose the 'minimal' according to these ordering, i.e. a cycle that does not enclose any other cycle. If there's any region enclosed by that cycle, that region must be a finite area region and therefore be part of a path that ends in an unbounded region, which must then be enclosed by the cycle (which is contradiction that this region that's end of path is unbounded). Hence no region can be enclosed by the cycle at all.

Consider, now, any door that joins two opposite regions, and the two side regions. One side region is outside the cycle; the other is inside, contradicting that no region can be enclosed by the cycle. Therefore there's no cycle to start with, as desired.

C5. (IMO 3) A social network has 2019 users, some pairs of whom are friends. Whenever user A is friends with user B, user B is also friends with user A. Events of the following kind may happen repeatedly, one at a time: Three users A, B, and C such that A is friends with both B and C, but B and C are not friends, change their friendship statuses such that B and C are now friends, but A is no longer friends with B, and no longer friends with C. All other friendship statuses are unchanged. Initially, 1010 users have 1009 friends each, and 1009 users have 1010 friends each. Prove that there exists a sequence of such events after which each user is friends with at most one other user.

Solution. (Reproduced from AoPS) For a connected component C of a graph, call it good if either $|C| \leq 2$, or $|C|$ is not a clique and not all vertices have even degree. One can easily prove that the initial graph has a single connected component (otherwise, take the smaller component which has at most 1009 vertices and therefore having degree at most 1008), and obviously not a clique. For each good connected component of size at least 3, we show that we can make a move as described in the problem statement such that either the connected component is preserved, or the move splits it into two good components (in the latter case, the resulting component will still be good because as Swistak mentioned, the each move preserves the degree of each vertex, and the total number of edges decreases by 1, hence cannot be a clique).

Consider any good component C of size at least 3. For each two vertices (u, v), denote the distance $d(u, v)$ as the length of the shortest path from u and v. Choose two of them, A_0, A_n such that $d(u, v) = n$ and is the maximal possible within that component. Moreover, let $A_0, A_1, \cdots, A_n$ to be a path of length n. By the definition of distance, A_i and A_j is not connected by an edge if $|i - j| > 1$. In addition, $n \geq 2$ since C is not a clique.

Now, one possible candidate (call this "candidate move") is to consider edges A_0A_1 and A_1A_2, and remove them, resulting in a new edge A_0A_2. If A_1 is still connected to A_0, or A_1 has no other vertices connected to itself, then we are good. Otherwise, A_1 will have another neighbour, B.

We first claim that in this "supposedly new" component, A_1 has distance 1 with all other vertices, Otherwise, choose B and C such that there are edges A_1B and BC but not A_1C. Now these B and C are isolated from $A_0, A_2, \cdots, A_n$. In the original graph, any path from A_n to C must pass through A_1 (length $n - 1$), and from A_1 it takes exactly two steps to reach C. Thus $d(A_n, C)$ in the pre-move configuration is $n + 1$, contradicting the maximality of n.

Back to the original configuration; we know that originally, the neighbours of A_1 are A_0, A_2, and a bunch of $B_1, \cdots, B_k$ $(k \geq 1)$ which will be isolated from $A_0, A_2, \cdots, A_n$ should we use the candidate move. In addition, by the immediate previous claim, B_i has no other neighbours other than B_j $(j \neq i)$ and A_1. There are three cases:

Case 1. $k = 1$. This means the degree of A_1 is 3. We now disconnect A_1B_1 and A_1A_0, resulting in new edge B_1A_0. If this were to split things up into two components, then B_1, A_0 are in the same component, while $A_1, A_2, \cdots, A_n$ are in the other. Both B_1 and A_1 now have degree 1, so both components are guaranteed to be good.

Case 2. $k \geq 2$ and each B_i has degree 1. The previous algorithm will still work: disconnect A_1B_1 and A_1A_0 with new edge B_1A_0. Again if this were to split things up into two components then B_1 and A_1 are in different components. Now B_1 has degree 1; B_2 has degree 1 too and is in A_1's component, so both components are good too.

Case 3. Some B_i has degree more than 1. WLOG let B_1 to have degree more than 1, and by before's observation, B_1 has no other neighbour other than some other B_j's and A_1. Therefore there's another B_j that's B_1's neighbour, say, B_2. Same algorithm, cut A_1B_1, A_1A_0, add B_1A_0. Now $A_0B_1B_2A_1$ is a path, so the original component remains connected.

Finally, notice that we always perform such a move above until we run out of connected components of size ≥ 3. This will eventually happen since each move decreases the number of edges in the graph by 1, thus cannot happen forever. At this stage, each vertex will have degree at most 1, as desired.

C6. Let $n > 1$ be an integer. Suppose we are given $2n$ points in the plane such that no three of them are collinear. The points are to be labelled $A_1, A_2, \ldots, A_{2n}$ in some order. We then consider the $2n$ angles $\angle A_1A_2A_3, \angle A_2A_3A_4, \ldots, \angle A_{2n-2}A_{2n-1}A_{2n}, \angle A_{2n-1}A_{2n}A_1, \angle A_{2n}A_1A_2$.

We measure each angle in the way that gives the smallest positive value (i.e. between $0°$ and $180°$). Prove that there exists an ordering of the given points such that the resulting $2n$ angles can be separated into two groups with the sum of one group of angles equal to the sum of the other group.

Solution. We consider the "pivoted directed angle" $\theta(A_i, A_{i+1}, A_{i+2})$ as the angle that needs to rotate line A_iA_{i+1} to $A_{i+1}A_{i+2}$ under the pivot A_{i+1} clockwise, and also restrict this value $\theta(A_i, A_{i+1}, A_{i+2})$ to the range $(-180°, 180°)$. Let's explore certain properties of this θ function:

Lemma 1. $\sum_{i=1}^{2n} \theta(A_i, A_{i+1}, A_{i+2})$ is always a multiple of $360°$.

Proof: we consider moving line A_1A_2 to A_2A_3, from A_2A_3 to A_3A_4, etc. We also fix point A at A_1 and point B at A_2 and rotate accordingly. Then we will notice that when A_i is a pivot, this pivot will be B if i is even, and A if i is odd. Thus, $\sum_{i=1}^{2n} \theta(A_i, A_{i+1}, A_{i+2})$ will be the angle needed to rotate A_1A_2 all the way back to A_1A_2 itself. Since A will always lie on odd-indiced vertices and B on even-indiced vertices, A will be on A_1 and B will be on A_2 as before. This shows that $\sum_{i=1}^{2n} \theta(A_i, A_{i+1}, A_{i+2})$ is a multiple of full turn (if A and B flip position, then that will be congruent to $180°$ modulo $360°$ but that's not the case here).

Lemma 2. By swapping A_i and A_{i+1} for some i, the quantity $\sum_{i=1}^{2n} \theta(A_i, A_{i+1}, A_{i+2})$ changes by at most $360°$.

Proof: W.L.O.G. we assume that A_3 and A_4 are swapped, and we want to consider how does

$$\theta(A_1, A_2, A_3) + \theta(A_2, A_3, A_4) + \theta(A_3, A_4, A_5) + \theta(A_4, A_5, A_6)$$

compare to

$$\theta(A_1, A_2, A_4) + \theta(A_2, A_4, A_3) + \theta(A_4, A_3, A_5) + \theta(A_3, A_5, A_6)$$

We first compare $\theta(A_1, A_2, A_3) + \theta(A_2, A_3, A_4)$ and $\theta(A_1, A_2, A_4) + \theta(A_2, A_4, A_3)$. We claim that they differ (in either direction) by $180°$.

Assume, w.l.o.g. that A_2, A_3, A_4 lie in that order, clockwise. Then $\theta(A_2, A_3, A_4)$ moves from A_2A_3 to A_3A_4 which is anticlockwise, while $\theta(A_2, A_4, A_3)$ moves from A_2A_4 to A_3A_4, which is clockwise. Thus $\theta(A_2, A_4, A_3) - \theta(A_2, A_3, A_4)$ is positive and is equal to sum of angles $\angle A_2A_3A_4 + \angle A_2A_4A_3 = 180° - \angle A_3A_2A_4$. As for $\theta(A_1, A_2, A_3)$ vs $\theta(A_1, A_2, A_4)$, we have a few cases:

- If A_1A_2 lies in the angle domain of $\angle A_3A_2A_4$, then the rotation from A_1A_2 to A_2A_3 is anticlockwise and A_1A_2 to A_2A_4 is clockwise, so $\theta(A_1, A_2, A_4) - \theta(A_1, A_2, A_3) = \angle A_3A_2A_4$ here (i.e. sum of $\angle A_3A_2A_1 + \angle A_1A_2A_4$).
- If A_1A_2 lies on the side (not opposite) of angle domain of $\angle A_3A_2A_4$ (the lines A_2A_3 and A_2A_4, when extended, divde the plane into 4 regions and segment A_1A_2 lies on the region that's by the side of $A_3A_2A_4$), then either both the rotation from A_1A_2 to A_2A_3 and A_1A_2 to A_2A_4 are clockwise (if A_1A_2 is closer to A_2A_3 compared to A_2A_4), or both anticlockwise (if A_1A_2 is closer to A_2A_4 compared to A_2A_3). In both cases we have $\theta(A_1, A_2, A_4) - \theta(A_1, A_2, A_3) = \angle A_3A_2A_4$.
- If A_1A_2 lies at the opposite of angle domain of $\angle A_3A_2A_4$, then A_1A_2 to A_2A_3 is clockwise while to A_2A_4 is anticlockwise. Thus $\theta(A_1, A_2, A_4) - \theta(A_1, A_2, A_3) = -(360° - \angle A_3A_2A_4)$.

Thus $\theta(A_1, A_2, A_4) - \theta(A_1, A_2, A_3)$ is either $\angle A_3A_2A_4$ or $-(360° - \angle A_3A_2A_4)$. This gives

$$\theta(A_1,A_2,A_4)+\theta(A_2,A_4,A_3)-(\theta(A_1,A_2,A_3)+\theta(A_2,A_3,A_4)) = \begin{cases} 180° - \angle A_3A_2A_4 + \angle A_3A_2A_4 \\ 180° - \angle A_3A_2A_4 - (360° - \angle A_3A_2A_4) \end{cases}$$

i.e. is always $\pm 180°$.

The same thing can be said to $\theta(A_4,A_3,A_5)+\theta(A_3,A_5,A_6)-(\theta(A_3,A_4,A_5)+\theta(A_4,A_5,A_6))$, therefore the difference of

$$\theta(A_1, A_2, A_3) + \theta(A_2, A_3, A_4) + \theta(A_3, A_4, A_5) + \theta(A_4, A_5, A_6)$$

compare to

$$\theta(A_1, A_2, A_4) + \theta(A_2, A_4, A_3) + \theta(A_4, A_3, A_5) + \theta(A_3, A_5, A_6)$$

is either $180° + 180°, 180° - 180°, -180° - 180°$, therefore either 0 or $\pm 360°$.

Lemma 3. There exists an ordering of vertices such that $\sum_{i=1}^{2n} \theta(A_i, A_{i+1}, A_{i+2}) = 0$.

Solution. Let's start with an arbitrary ordering. Suppose that $\sum_{i=1}^{2n} \theta(A_i, A_{i+1}, A_{i+2}) \neq 0$, and w.l.o.g. we assume that it's positive, i.e. > 0. Reversing the order (e.g. $A_i \to A_{2n-i+1}$) makes $\sum_{i=1}^{2n} \theta(A_i, A_{i+1}, A_{i+2})$ become $\sum_{i=1}^{2n} \theta(A_{i+2}, A_{i+1}, A_i)$, i.e. instead of moving A_iA_{i+1} to $A_{i+1}A_{i+2}$, we make $A_{i+1}A_{i+2}$ to A_iA_{i+1}. This essentially negates each term $\theta(A_i, A_{i+1}, A_{i+2})$, and therefore $\sum_{i=1}^{2n} \theta(A_{i+2}, A_{i+1}, A_i)$, the new $\sum_{i=1}^{2n} \theta(A_i, A_{i+1}, A_{i+2})$ becomes negative.

We recall that this summation $\sum_{i=1}^{2n} \theta(A_i, A_{i+1}, A_{i+2})$ is a multiple of 360°. To reverse it, we simply have to swap A_i and A_{i+1} iteratively (think of it as doing bubble sort). Along the way, we monitor $\sum_{i=1}^{2n} \theta(A_i, A_{i+1}, A_{i+2})$. Since we start at positive multiple of 360°, end at negative, and move by either exactly 0 or 360° at each time, there must be a point where we hit exactly 0, completing the proof.

Now we can complete the solution. Label the vertices such that $\sum_{i=1}^{2n} \theta(A_i, A_{i+1}, A_{i+2}) = 0$. Notice that $\theta(A_i, A_{i+1}, A_{i+2}) = \pm\angle A_iA_{i+1}A_{i+2}$, so we simply have to split these angles into two groups such that one group has $\theta(A_i, A_{i+1}, A_{i+2}) > 0$ and the other group has $\theta(A_i, A_{i+1}, A_{i+2}) < 0$.

Geometry

G1. Let ABC be a triangle. Circle Γ passes through A, meets segments AB and AC again at points D and E respectively, and intersects segment BC at F and G such that F lies between B and G. The tangent to circle BDF at F and the tangent to circle CEG at G meet at point T. Suppose that points A and T are distinct. Prove that line AT is parallel to BC.

Solution. We have

$$\angle TFG = \angle BDF = \angle AGF$$

the first equality due to TF tangent to BDF and the second due to A, D, F, G concyclic. Similarly we have

$$\angle TGF = \angle CEG = \angle AFG$$

and therefore triangles AFG and TGF are similar and with the common side FG, they are in fact congruent. Since A and T are on the same side of FG, we have $ATGF$ isoceles trapezoid and therefore $AT \parallel FG$ (and therefore $AT \parallel BC$ since FG and BC are the same line).

G2. Let ABC be an acute-angled triangle and let D, E, and F be the feet of altitudes from A, B, and C to sides BC, CA, and AB, respectively. Denote by ω_B and ω_C the incircles of triangles BDF and CDE, and let these circles be tangent to segments DF and DE at M and N, respectively. Let line MN meet circles ω_B and ω_C again at $P \neq M$ and $Q \neq N$, respectively. Prove that $MP = NQ$.

Solution. Let r_B and r_C be the inradius of BDF and CDE respectively, and θ_B and θ_C be the angles subtended by MP and NQ in ω_B and ω_C, respectively. Then $MP = r_B \sin\theta_B$ and $NQ = r_C \sin\theta_C$.

Notice, first, that since DM is tangent to ω_B, θ_B is the angle between MP and DM, which is the angle $\angle NMD$. Similarly $\theta_C = \angle MND$. Therefore,

$$\frac{MP}{NQ} = \frac{r_B \sin\theta_B}{r_C \sin\theta_C} = \frac{r_B \sin\angle NMD}{r_C \sin\angle MND} = \frac{r_B \cdot ND}{r_C \cdot MD}$$

where the last equality is due to sine rule on triangle NMD.

Finally, we have triangles DBF and DEC similar. This means that $\frac{MD}{ND}$, the tangent from D to each circle, is equal to $\frac{r_B}{r_C}$. Therefore,

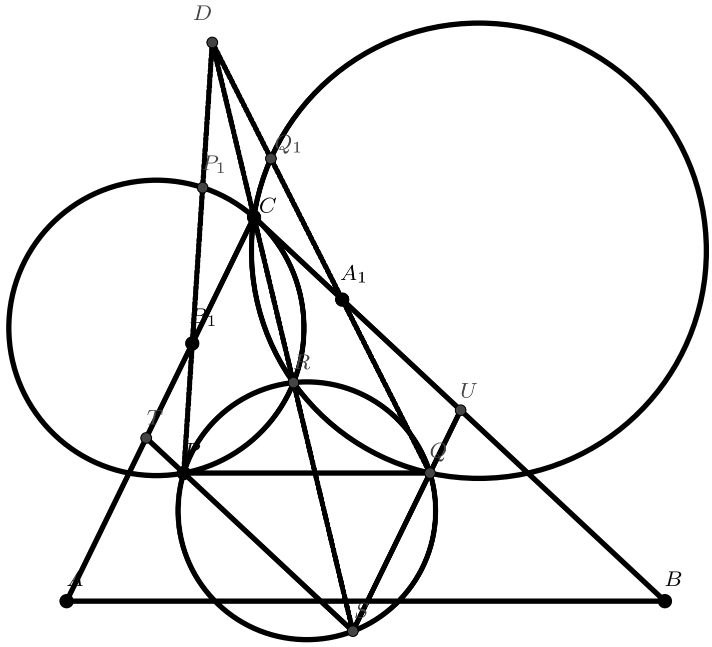
$$\frac{r_B \cdot ND}{r_C \cdot MD} = \frac{r_B \cdot r_C}{r_C \cdot r_B} = 1$$

as desired.

G3. (IMO 2) In triangle ABC, point A_1 lies on side BC and point B_1 lies on side AC. Let P and Q be points on segments AA_1 and BB_1, respectively, such that PQ is parallel to AB. Let P_1 be a point on line PB_1, such that B_1 lies strictly between P and P_1, and $\angle PP_1C = \angle BAC$. Similarly, let Q_1 be the point on line QA_1, such that A_1 lies strictly between Q and Q_1, and $\angle CQ_1Q = \angle CBA$.

Prove that points P, Q, P_1, and Q_1 are concyclic.

Solution. An equivalent thing would be proving that lines PP_1, QQ_1 and the radical axis of circumcircles of CPP_1 and CQQ_1 are either concurrent or parallel. Let R be the second intersection of circles CPP_1 and CQQ_1; CR is the radical axis of these two circles.



Now, let S be the second intersection of CR and circle PQR. We claim that $SP \parallel BC$ and $SQ \parallel AC$. For this part only, let's use directed angles for the sake of clarity. Observe that:

$$\angle(CA, AB) = \angle(CP_1, P_1P) = \angle(CR, RP) = \angle(RS, RP) = \angle(QS, QP) = \angle(QS, AB)$$

with the first equality following from $\angle PP_1C = \angle BAC$ and the last equality from $QP \parallel AB$. Thus CA and QS are parallel, and similarly so are lines SP and BC. Thus we will focus on proving that PP_1, QQ_1, CS are either concurrent or parallel.

Let SP intersect AC at T, and QP intersect BC at U. THis means $CTSU$ is a parallelogram. The fact that PP_1, QQ_1, CS are parallel is actually the same as the below equality:

$$\frac{SP}{PT} \cdot \frac{TB_1}{B_1C} \cdot \frac{CA_1}{A_1U} \cdot \frac{UQ}{QS} = 1$$

notice that we need to be careful of the sign convention, though it's quite clear that P is between T and S and Q is between U and S, so the first and the last ratio can be easily made positive. We now let $\frac{TB_1}{B_1C}$ be postive if B_1 lies between T and C and negative otherwise, similarly $\frac{CA_1}{A_1U}$ be positve if A_1 between C and U and negative otherwise. This allows us to adopt the convention that $TB_1 = TC - B_1C$ and $UA_1 = UC - A_1C$.

Since $PQ \parallel AB$, $SP \parallel BC$ and $SQ \parallel AC$, the triangles SPQ and CBA are in fact similar. Thus $\frac{SP}{SQ} = \frac{BC}{AC}$. Since $CU = TS$ and $CT = US$, we in fact have $\frac{SP}{ST} \div \frac{SQ}{SU} = \frac{BC}{CU} \div \frac{AC}{CT}$. In addition, we have the following identity:

$$\frac{PT}{A_1C} = \frac{AT}{AC} \qquad \frac{UQ}{B_1C} = \frac{BU}{BC}$$

thus allowing us to write

$$A_1C = PT \cdot \frac{AC}{AT} = PT \cdot \frac{AC}{AC - TC} \qquad B_1C = UQ \cdot \frac{BC}{BU} = UQ \cdot \frac{BC}{BC - CU}$$

and therefore

$$\begin{aligned}
\frac{SP}{PT}\cdot\frac{TB_1}{B_1C}\cdot\frac{CA_1}{A_1U}\cdot\frac{UQ}{QS} &= \frac{SP}{QS}\cdot\frac{CT-B_1C}{\frac{BC}{BC-CU}}\cdot\frac{\frac{AC}{AC-TC}}{CU-CA_1}\\
&= \frac{SP}{QS}\cdot\frac{CT(1-B_1C/CT)}{\frac{BC}{BC-CU}}\cdot\frac{\frac{AC}{AC-TC}}{CU(1-CA_1/CU)}\\
&= \frac{BC}{AC}\cdot\frac{CT(1-\frac{UQ}{US}\cdot\frac{BC}{BC-CU})}{\frac{BC}{BC-CU}}\cdot\frac{\frac{AC}{AC-TC}}{CU(1-\frac{PT}{ST}\cdot\frac{AC}{AC-TC})}\\
&= (\frac{BC}{CU}\div\frac{AC}{CT})\cdot\frac{1-(1-\frac{SQ}{SU})\cdot\frac{BC}{BC-CU}}{\frac{BC}{BC-CU}}\cdot\frac{\frac{AC}{AC-TC}}{1-(1-\frac{SP}{ST})\cdot\frac{AC}{AC-TC}}\\
&= \frac{CT}{CU}\cdot\frac{1-(1-\frac{SQ}{SU})\cdot\frac{BC}{BC-CU}}{\frac{1}{BC-CU}}\cdot\frac{\frac{1}{AC-TC}}{1-(1-\frac{SP}{ST})\cdot\frac{AC}{AC-TC}}\\
&= \frac{CT}{CU}\cdot\frac{1-(1-\frac{SQ}{SU})\cdot\frac{BC}{BC-CU}}{AC-TC}\cdot\frac{BC-CU}{1-(1-\frac{SP}{ST})\cdot\frac{AC}{AC-TC}}\\
&= \frac{CT}{CU}\cdot\frac{BC-CU-(1-\frac{SQ}{SU})\cdot BC}{AC-TC-(1-\frac{SP}{ST})\cdot AC}\\
&= \frac{CT}{CU}\cdot\frac{\frac{SQ}{SU}\cdot BC-CU}{\frac{SP}{ST}\cdot AC-TC}\\
&= \frac{\frac{SQ}{SU}\cdot\frac{BC}{CU}-1}{\frac{SP}{ST}\cdot\frac{AC}{CT}-1}
\end{aligned}$$

To show that this ratio is indeed 1, it suffices to show that $\frac{SQ}{SU}\cdot\frac{BC}{CU}=\frac{SP}{ST}\cdot\frac{AC}{CT}$. With $CT=SU$ and $CU=ST$, and that $\frac{SP}{SQ}=\frac{BC}{AC}$ (i.e. $SP\cdot AC=SQ\cdot BC$), we have

$$\frac{SQ}{SU}\cdot\frac{BC}{CU}=\frac{SP}{SU}\cdot\frac{AC}{CU}=\frac{SP}{CT}\cdot\frac{AC}{ST}$$

as desired.

G4. Let P be a point inside triangle ABC. Let AP meet BC at A_1, let BP meet CA at B_1, and let CP meet AB at C_1. Let A_2 be the point such that A_1 is the midpoint of PA_2, let B_2 be the point such that B_1 is the midpoint of PB_2, and let C_2 be the point such that C_1 is the midpoint of PC_2. Prove that points A_2, B_2, and C_2 cannot all lie strictly inside the circumcircle of triangle ABC.

Solution. Let AP, BP, CP meet the circumcircle of ABC again at A_3, B_3, C_3, respectively. Suppose that A_2, B_2, C_2 all lie inside the circumcircle. Then $A_1P < A_3A_1$, $B_1P < B_3B_1$ and $C_1P < C_3C_1$. We also have:

$$\frac{A_1P}{A_3A_1}=\frac{|\triangle PBC|}{|\triangle BCA_1|}=\frac{\frac{1}{2}\cdot d(P,BC)\cdot BC}{\frac{1}{2}\cdot d(A_1,BC)\cdot BC}=\frac{d(P,BC)/BC}{d(A_1,BC)/BC}$$

where $d(P,BC)$ is the distance from point P to line BC or the height from P of the triangle PBC (and similar for other notations). Now we have

$$\frac{d(P,BC)}{BC}=\frac{d(P,BC)}{PB}\cdot\frac{PB}{BC}=\sin\angle PBC\cdot\frac{\sin\angle PCB}{\sin\angle CPB}=\frac{\sin\angle PBC\cdot\sin\angle PCB}{\sin\angle(PBC+PCB)}$$

$$=\frac{1}{\cot\angle PBC+\cot\angle PCB}$$

and therefore

$$\frac{A_1P}{A_3A_1} = \frac{1}{\cot\angle PBC + \cot\angle PCB} \div \frac{1}{\cot\angle A_3BC + \cot\angle A_3CB} = \frac{\cot\angle A_3AC + \cot\angle A_3AB}{\cot\angle PBC + \cot\angle PCB}$$

$$= \frac{\cot\angle PAC + \cot\angle PAB}{\cot\angle PBC + \cot\angle PCB}$$

where $\angle A_3BC = \angle A_3AC$ and $\angle A_3CB = \angle A_3AB$ follow from that A, B, C, A_3 concyclic. The fact that $A_3P < C_3A_1$ therefore means

$$\cot\angle PAC + \cot\angle PAB < \cot\angle PBC + \cot\angle PCB$$

and similarly $B_1P < B_3B_1$ and $C_1P < C_3C_1$ means

$$\cot\angle PBC + \cot\angle PBA < \cot\angle PCA + \cot\angle PAC$$

$$\cot\angle PCA + \cot\angle PCB < \cot\angle PBA + \cot\angle PAB$$

so summing the 3 quantities gives

$$\cot\angle PAC + \cot\angle PAB + \cot\angle PBC + \cot\angle PBA + \cot\angle PCA + \cot\angle PCB$$

$$< \cot\angle PAC + \cot\angle PAB + \cot\angle PBC + \cot\angle PBA + \cot\angle PCA + \cot\angle PCB$$

which is a contradiction.

G6. Let I be the incentre of acute-angled triangle ABC. Let the incircle meet BC, CA, and AB at D, E, and F, respectively. Let line EF intersect the circumcircle of the triangle at P and Q, such that F lies between E and P. Prove that $\angle DPA + \angle AQD = \angle QIP$.

Solution. We denote the second intersection of QD and PD as X and Y, respectively, and the intersection of PQ and BC as T (possibly point of infinity).

We first claim that XY also passes through T. Since AD, BE, CF are concurrent (well-known), the points $(T, B; D, C)$ are harmonic. By Brokard's theorem, if ℓ is the pole of D w.r.t. to circle ABC, then PQ and XY intersect on ℓ. If D' is the projection from OD to ℓ, then $OD \cdot OD' = r^2 = OB^2 = OC^2$. Thus O, B, D', C are in fact cyclic and so OD' is an internal angle bisector of angle $BD'C$. Now, ℓ and OD' are perpendicular, so if ℓ intersects BC at T' then based on the two angle bisectors we have $(DT', DB; DD', DC)$ are harmonic pencil, so $(T', B; D, C)$ are harmonic. This gives $T = T'$ and therefore PQ, BC, XY concur at T.

Now, by Brokard's theorem again, PX and YQ intersect at the polar of TD which is BC. This point, namely U, is the intersection of tangents to of B and C to the circumcircle. This means $BXCP$ and $BYCQ$ are both harmonic quadrilaterals. Also using the identities on the cevians of triangle (with some trigonometric flavour) we have

$$\frac{\sin\angle PIF}{\sin\angle PIE} \cdot \frac{IF}{IE} = \frac{PF}{PE} = \frac{\sin\angle PAF}{\sin\angle PAE} \cdot \frac{AF}{AE}$$

but since $AF = AE$ and $IF = IE$ we have

$$\frac{\sin\angle PIF}{\sin\angle PIE} = \frac{\sin\angle PAF}{\sin\angle PAE} = \frac{\sin\angle PAB}{\sin\angle PAC} = \frac{PB}{PC} = \frac{XB}{XC} = \frac{\sin\angle BAX}{\sin\angle XAC}$$

and given that $\angle PIE = \angle PIF + \angle FIE = \angle PIF + 180^\circ - \angle BAC$ we have $\sin\angle PIE = \sin(\angle BAC - \angle PIF$, and $\sin\angle XAC = \sin(\angle BAC - \angle BAX)$, we have

$$\frac{\sin\angle PIF}{\sin(\angle BAC - \angle PIF)} = \frac{\sin\angle BAX}{\sin(\angle BAC - \angle BAX)}$$

which then gives $\angle PIF = \angle BAX$ after expanding. Similarly, we have $\angle QIE = \angle CAY$. Therefore,

$$\angle DPA + \angle AQD = \angle YPA + \angle AQX$$
$$= \angle CPA + \angle CAY + \angle AQB + \angle BAX = \angle CBA + \angle QIE + \angle ACB + \angle PIF$$
$$= 180^\circ - \angle BAC + \angle QIE + \angle PIF = \angle FIE + \angle QIE + \angle PIF = \angle QIP$$

as desired.

G7. (IMO 6) Let I be the incentre of acute triangle ABC with $AB \neq AC$. The incircle ω of ABC is tangent to sides BC, CA, and AB at D, E, and F, respectively. The line through D perpendicular to EF meets ω at R. Line AR meets ω again at P. The circumcircles of triangle PCE and PBF meet again at Q.

Prove that lines DI and PQ meet on the line through A perpendicular to AI.

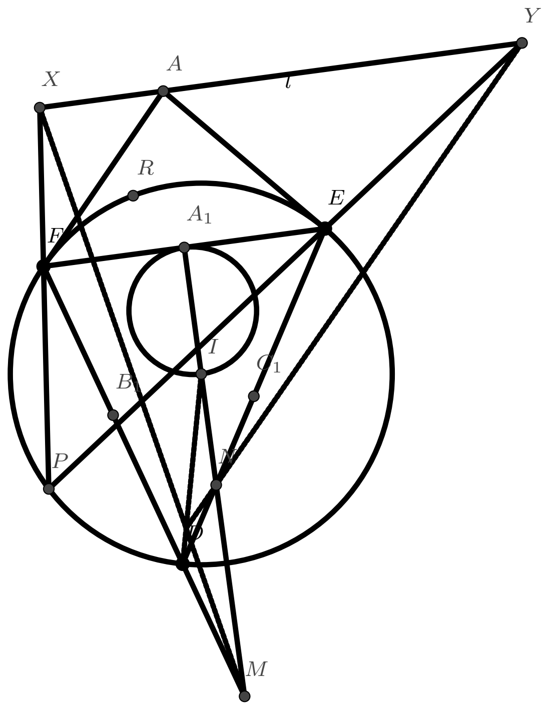
Solution. (Reproduced from AoPS) Here's a solution using inversion and trigonometric bashing: by inverting in ω we turn the problem into the follows: keep A, D, E, F, P as they are, A_1, B_1, C_1 as midpoints of EF, DF, DE. Let the circumcircles of triangles PC_1E and PB_1F meet again on Q_1. Let γ be the circle with diameter A_1I. Prove that the second intersection of circumcircle of PQ_1I and line DI (or I if tangent) meet on γ. The last statement (to prove) is the same as proving that the radical axis of γ, circumcircles of triangles PC_1E and PB_1F lie on line DI. Equivalently, the radical axis of PC_1E and γ, and the radical axis of PB_1F and γ concur on DI. This last statement is our focus. W.L.O.G. assume $AB < AC$, so $DF < DE$ too and we know DEF is acute. In this setting (details skipped), P will lie on minor arc DF, and in particular P, D, F are different points.

We need to identify those radical axes, and for each of them it's defined based on two points. We first focus on finding the radical axis of PB_1F and γ. We need the following lemma:

Lemma: A is on the radical axis of ω and γ.
Proof: The power of point of A to ω is $AE = AF$; to γ is $AA_1 \cdot AI$. But then $AA_1 \perp EF$ and $\angle AFI = 90^\circ$, the conclusion follows by similarity of triangles. This means the radical axis is actually line through A perpendicular to AA_1 (parallel to EF, in other words), let's name it ℓ.

Consider, now, the circles PB_1F, γ and ω. The radical axis of ω and γ has been established above; the one for ω and PB_1F is PF, so one such point must be $X = \ell \cap PF$. Consider, now, the circles PB_1F, γ and the circle B_1FA_1I (the four points are concylic bcz $\angle FB_1I = \angle FA_1I$). The radical axis of γ and B_1FA_1I is A_1I; the one for PB_1F and B_1FA_1I is B_1F which is DF. Thus $M = DF \cap A_1I$ is the radical centre of the three circles, hence on the radical axis of γ and PB_1F. The radical axis, therefore, is XM (we need to be careful in showing $X \neq M$; suppose $X = M$, then the fact that both pass through F via FP and FD and that F, D, P are not collinear because they are different point on ω means $X = M = F$. But the power of point of F to FPB_1 is 0 while to γ is $A_1F \neq 0$, assuming nondegenerate here). In a similar fashion, if $Y = \ell \cap PE$ and $N = DE \cap A_1I$ then YN is the radical axis of γ and PC_1E.

We are left with proving that YN, XM, DI are concurrent, which we will use bashing (trigonometric) here! Extend line MFD to meet ℓ at X_1 and NED to meet ℓ at Y_1. We now have $MF > DF$ and $NE < DE$ According to our assumption, X will be between P and X_1 and MX will be in the angle domain of $\angle AMX_1$. So goes to segment DI so MX will intersect segment DI (and not anything outside). Similarly, Y_1 will be between A and Y, which also means NY is outside angle domain $\angle ANY_1$. But then segment DI won't be on the angle domain either so YN intersects DI in its segment. These realization are here to free us from using signed convention later (well we could but I am lazy now).



Now a not-so-well-known trigonometric identity says that considering the triangle AFX_1 and the cevian FX we get $\frac{AX}{XX_1} = \frac{AF}{FX_1} \cdot \frac{\sin\angle AFX}{\sin\angle XFX_1}$. Considering the triangle AMX_1 and cevian MX we get $\frac{AX}{XX_1} = \frac{AM}{MX_1} \cdot \frac{\sin\angle AMX}{\sin\angle XMX_1}$. Finally, if MX intersects DI at Z_1, considering triangle MDI and the cevian MZ_1 we get, $\frac{IZ_1}{Z_1D} = \frac{MI}{MD} \cdot \frac{\sin\angle AMX}{\sin\angle XMX_1}$. Thus we have

$$\frac{IZ_1}{Z_1D} = \frac{MI}{MD} \cdot \frac{\sin\angle AMX}{\sin\angle XMX_1} = \frac{MI}{MD} \cdot \frac{AX}{XX_1} \div \frac{AM}{MX_1} = \frac{MI}{MD}\frac{AF}{FX_1} \cdot \frac{\sin\angle AFX}{\sin\angle XFX_1} \div \frac{AM}{MX_1}$$

and similarly if NY intersects DI at Z_2 we get

$$\frac{IZ_2}{Z_2D} = \frac{NI}{ND}\frac{AE}{EY_1} \cdot \frac{\sin\angle AEY}{\sin\angle YEY_1} \div \frac{AN}{NY_1}$$

so we need to prove the two ratios are equal (and this would be sufficient since we know that Z_1 and Z_2 are both on segment DI, i.e.

$$\frac{MI}{MD}\frac{AF}{FX_1} \cdot \frac{\sin\angle AFX}{\sin\angle XFX_1} \div \frac{AM}{MX_1} = \frac{NI}{ND}\frac{AE}{EY_1} \cdot \frac{\sin\angle AEY}{\sin\angle YEY_1} \div \frac{AN}{NY_1}$$

First, notice that $AF = AE$, so these can be cancelled out. Next, $\angle AFX = \angle FEP$ and $\angle AEY = \angle EFP$, but then by sine rule $\sin\angle FEP / \sin\angle EFP = FP/EP$. These two angles are on numerators of two different sides so they can be replaces with FP and EP, respectively. Then, $\angle XFX_1 = \angle FPD = \angle PED = \angle YEY_1$, again can be cancelled. Also, since $\ell \parallel EF$ we have $FX_1/EY_1 = DF/DE$. Thus we now need to check the following:

$$\frac{MI}{MD}\frac{1}{DF} \cdot \frac{FP}{1} \div \frac{AM}{MX_1} = (?)\frac{NI}{ND}\frac{1}{DE} \cdot \frac{EP}{1} \div \frac{AN}{NY_1}$$

Next, AM/MX_1 is actually $\cos\angle AMX_1 = \sin\angle DFA_1 = \sin\angle DFE$ and similarly $AN/NY_1 = \sin\angle DEA_1 = \sin\angle DEF$ (notice the implicit use of the fact that $A_1I \perp EF$ and $A_1I \perp XY = \ell$. But then by sine rule $\sin\angle DFE/\sin\angle DEF = DE/DF$. Thus we have the equation to prove above becomes the following:

$$\frac{MI}{MD}\frac{1}{DF}\cdot\frac{FP}{1} \div DE = (?)\frac{NI}{ND}\frac{1}{DE}\cdot\frac{EP}{1} \div DF$$

so now it suffices to show that

$$\frac{MI}{MD}\cdot FP = (?)\frac{NI}{ND}\cdot EP$$

Using sine rule again, $\frac{MI}{MD} = \frac{\sin\angle MDI}{\sin\angle MID}$ and $\frac{NI}{ND} = \frac{\sin\angle NDI}{\sin\angle NID}$. But then both $\angle MID$ and $\angle NID$ are angle between DI and IM so they must be either equal or supplementary, hence having equal since. We now reduce everything to the following: $\sin\angle MDI \cdot FP = (?)\sin\angle NDI \cdot EP$. But now, $\sin\angle MDI = \sin\angle IDF$ and notice that $\angle IDF = 90^\circ - \angle DEF = \angle EDR = \angle EFR$ so $\sin\angle MDI = \angle EFR$ and $\sin\angle NDI = \sin\angle FER$ for the similar reason. But then $\sin\angle EFR/\sin\angle FER = ER/FR$ so we are left with proving that $ER\cdot FP = FR/EP$. But this follows from the fact that $PEFR$ is a harmonic quadrilateral!

Number Theory

N1. (IMO 4) Find all pairs (k, n) of positive integers such that

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

Answer. $(1,1),(3,2)$.
Solution. It's obvious that the answers above satisfy the condition, and we will show that these are the only pairs. By considering the power of 2 dividing both sides (denote by $v_2(\cdot)$) we have the v_2 of right hand side as

$$= \sum_{i=0}^{n-1} v_2(2^n-2^i) = \sum_{i=0}^{n-1} i = \frac{n(n-1)}{2}$$

On the other hand we have

$$v_2(k!) = \sum_{i=1}^{\infty}\lfloor\frac{k}{2^i}\rfloor < \sum_{i=1}^{\infty}\frac{k}{2^i} = k$$

hence $k > \frac{n(n-1)}{2}$.

Now we compare the size of both sides: on right hand side we have

$$\prod_{i=1}^{n-1}(2^n-2^i) \le \prod_{i=1}^{n-2} 2^n\cdot(2^{n-1}) = 2^{n^2-1}$$

while for $k > 16$ (valid for $n \ge 7$ as of the first identity) we have

$$k! = 15!\cdot\prod_{i=16}^{k} \ge 15!\cdot 2^4 = 15!\cdot 2^{4(k-15)}$$

and we have $15! = 7! \cdot 8 \cdot \dots \cdot 15 > 5040 \cdot 8^8 > 2^12 \cdot 8^8 = 2^36$ so we in fact have $k! > 2^{4(k-15)+36} = 2^{4(k-6)}$. This means, $2^{4(k-6)} < k! \le 2^{n^2-1}$ and therefore $4(k-6) < n^2 - 1$. But then $k \ge \frac{n(n-1)}{2} + 1$ so

$$4(\frac{n(n-1)}{2} + 1 - 6) \le 4(k-6) < n^2 - 1$$

Or rather, $2n(n-1) - 20 < n^2 - 1$ or $n^2 - 2n = n(n-2) < 19$. Since $n(n-2)$ is an increasing functin and when $n = 7$ we have $7(5) = 35 > 19$, this inequality is false for $n \ge 7$.

We therefore only need to consider $n \le 6$. When $n \ge 5$, the factor 31 is present on the right hand side (via $2^n - 2^{n-5} = 2^{n-5}(2^5 - 1)$) and 31 is prime, so $k \ge 31$ here. This means $100 = 4(31-6) < n^2 - 1$, which is only true when $n \ge 11$ so this case is eliminated. To consider the rest manually:

- $n = 1$ gives $1 = 1!$ and $n = 2$ gives $2 \times 3 = 6 = 3!$.
- $n = 3$ gives $7 \times 6 \times 4 = 168$ which lies strictly between $5! = 120$ and $6! = 720$.
- $n = 4$ gives $15 \times 14 \times 12 \times 8 = 20160$ which lies strictly between $7! = 5040$ and $8! = 40320$.

Hence only $n = 1, 2$ work.

N2. Find all triples (a, b, c) of positive integers such that $a^3 + b^3 + c^3 = (abc)^2$.

Answer. $(a, b, c) = (3, 2, 1)$ and permutations. Both sides are equal to 36 in this case.

Solution. W.L.O.G. assume $a \ge b \ge c$, and consider the following representation:

$$a + \frac{b^3 + c^3}{a^2} = (bc)^2$$

Here, we have $a < (bc)^2$.

We now consider $a + \frac{b^3+c^3}{a^2}$ is increasing in a. Differentiating this w.r.t. a gives $1 - \frac{2(b^3+c^3)}{a^2}$, which is increasing in a, so $a + \frac{b^3+c^3}{a^2}$ is indeed convex. This means, since we're considering $a \in [b, (bc)^2 - 1]$, we have

$$a + \frac{b^3 + c^3}{a^2} \le \max\{b + \frac{b^3 + c^3}{b^2}, (bc)^2 - 1 + \frac{b^3 + c^3}{((bc)^2 - 1)^2}\}$$

i.e. considering the endpoints. This means one of $b + \frac{b^3+c^3}{b^2}$ and $(bc)^2 - 1 + \frac{b^3+c^3}{((bc)^2-1)^2}$ must be $\ge (bc)^2$.

The case $b + \frac{b^3+c^3}{b^2} \ge (bc)^2$ means $(bc)^2 \le 2b + \frac{c^3}{b^2} \le 3b$ since $c \le b$, or simply, $bc^2 \le 3$. Thus we have $bc \le 3$ here.

The case $(bc)^2 - 1 + \frac{b^3+c^3}{((bc)^2-1)^2} \ge (bc)^2$ simply means $b^3 + c^3 \ge ((bc)^2 - 1)^2$. Since $b^3c^3 + 1 \ge b^3 + c^3$, we have $b^3c^3 + 1 \ge ((bc)^2 - 1)^2$, or, substituting bc with x gives $x^3 + 1 \ge (x^2 - 1)^2$, or $x^2(x-2)(x+1) \le 0$. Since $x = bc > 0$, we have $x \le 2$, so $bc \le 2$ here.

Therefore combining both cases, we're limited to $bc \le 3$.

If $bc = 1$, then $a \le (bc)^2 - 1 = 0$ which is impossible. If $bc = 2$, we have $b = 2, c = 1$ which leaves us to solve $a^3 + 2^3 + 1^3 = a^2(2)^2$, or $a^3 + 9 = 4a^2$. This gives the factorization $(a-3)(a^2 - a - 1) = 0$, and since $a^2 - a - 1$ has discriminant 5 (not a square), it has no integer root, so $a = 3$ is the only solution. Finally, if $bc = 3$, then $b = 3, c = 1$ so $a^3 + 28 = 9a^2$. Here, $a = 2$ is a root and we have the factorization $(a-2)(a^2 - 7a - 14) = 0$ and $a^2 - 7a - 14$ has discriminant 105 (not square) so no integer root, so $a = 2$ is the only solution (which is simply another permutation of $(3, 2, 1)$).

N3. We say that a set S of integers is rootiful if, for any positive integer n and any $a_0, a_1, \cdots, a_n \in S$, all integer roots of the polynomial $a_0 + a_1 x + \cdots + a_n x^n$ are also in S. Find all rootiful sets of integers that contain all numbers of the form $2^a - 2^b$ for positive integers a and b.

Answer. The only rootiful set is the set of all integers, $\mathbb{Z}$.

Solution. First, taking $a = b$ in $2^a - 2^b$ means that $0 \in S$, and taking $a = 1, b = 0$ means $1 \in S$. Next, considering the polynomial $P(x) = a_0 + x$ means if $a_0 \in S$, so is $-a_0$. Considering $P(x) = a_0 + a_1 x$ means if $a_0, a_1 \in S$ and $a_1 \mid a_0$, then $\frac{-a_0}{a_1} \in S$, and therfore $\frac{a_0}{a_1} \in S$.

It then suffices to show that $n \in S$ for all $n \in \mathbb{N}$ since $-n \in S$ will follow, where we will proceed by induction. Let $n \geq 2$ and suppose that $0, 1, 2, \cdots, n-1 \in S$. Write $n = 2^m k$ where k odd, then by Euler's theorem there exists a such that $k \mid 2^a - 1$, and therefore $n = 2^m k \mid 2^{a+m} - 2^m$. Let $2^{a+m} - 2^m = n\ell$. There exists a unique way to write ℓ as base n:

$$\ell = \sum_{i=0}^{p} a_i n^i$$

where $0 \leq a_i \leq n-1$ (hence all $a_i \in S$). We therefore have

$$n\ell = \sum_{i=0}^{p} a_i n^{i+1}$$

and therefore n is a root of the polynomial

$$\sum_{i=0}^{p} a_i x^{i+1} - n\ell$$

Finally, since $n\ell, a_0, \cdots, a_{n-1} \in S$ (recall that $n\ell = 2^{a+m} - 2^m \in S$), $n \in S$, as desired.

N5. Let a be a positive integer. We say that a positive integer b is a-good if $\binom{an}{b} - 1$ is divisible by $an + 1$ for all positive integers n with $an \geq b$. Suppose b is a positive integer such that b is a-good, but $b+2$ is not a-good. Prove that $b+1$ is prime.

Solution. We show that b is a-good if and only if for all primes $p \leq b$, $p \mid a$. Suppose, now, $p \nmid a$ for some prime $p \leq b$. Then for any $m, k \geq 0$ we can always choose n such that $an + 1 \equiv m \pmod{p^k}$, and $an + 1 > b$. In addition, we need the condition $\gcd(\binom{an}{b}, an+1) = 1$ for all n, so if p is a prime such that $p \mid an+1$ then $p \nmid \binom{an}{b}$.

The next thing is to determine $\binom{an}{b} \pmod{p}$ assuming $p \mid an+1$. We'll use Lucas' theorem on $\binom{an}{b}$. If $p \leq b$, then b has at least two digits in base p. Let the $i > 0$-th digit position to be nonzero. Since $\gcd(a, p) = 1$, we can choose n such that an has i-th position $= 0$ and 0-th position $= p - 1$, so $p \mid an + 1$ but given the i-th position of an is less than that of b, we have $p \mid \binom{an}{b}$. This gives a contradiction.

Now that we have established $p \mid a$ for all primes $p \leq b$ (which tbh is the key to the solution), any prime dividing $an + 1$ are greater than b so $\gcd(b!, an+1) = 1$ for all n. Therefore, we have the following congruence:

$$\binom{an}{b} = \frac{an(an-1)\cdots(an-b+1)}{b!} \equiv \frac{(-1)(-2)\cdots(-b)}{b!} = (-1)^b \pmod{an+1}$$

so in this case, b is a-good iff b is even.

Now it's easy to complete the solution. Since b is a-good but $b+2$ is not, b and $b+2$ are both even, and $p \leq a$ for all $p \leq b$ but there exists a prime $q \leq b+2$ with $q \nmid a$. This q must be greater than b but at most $b+2$, and since $b+2$ is even (hence cannot be prime), $q = b+1$. Thus $b+1$ is prime.

N6. Let $H = \{\lfloor i\sqrt{2} \rfloor : i \in \mathbb{Z}_{>0}\} = \{1, 2, 4, 5, 7, \dots\}$ and let n be a positive integer. Prove that there exists a constant C such that, if $A \subseteq \{1, 2, \dots, n\}$ satisfies $|A| \geq C\sqrt{n}$, then there exist $a, b \in A$ such that $a - b \in H$. (Here $\mathbb{Z}_{>0}$ is the set of positive integers, and $\lfloor z \rfloor$ denotes the greatest integer less than or equal to z.)

Solution. We notice that the problem conclusion is similar to the following variants:

- There exists a constant C_0 such that if $A \subseteq \{1, \cdots, n\}$ satisfying $a - b \notin H$ for any $a, b \in A$, then $|A| \leq C_0\sqrt{n}$.
- There exists a constant C_1 such that if $a_1 < \cdots < a_k$ are such that $a_i - a_j \notin H$ for each $i \neq j$. Then $|a_k - a_1| \geq C_1 k^2$.

The idea of the solution is to consider everything "modulo $\sqrt{2}$". For each positive real number x, we define $f(x)$ as the quantity $x - \sqrt{2}\lfloor \frac{x}{\sqrt{2}} \rfloor$. Notice that $f(x) \in [0, \sqrt{2})$. We know that f is additive up to modulo $\sqrt{2}$.

We first notice that for positive integer k, $k \in H$ if and only if there exists a number ℓ such that $0 < \ell\sqrt{2} - k < 1$, which is to say $k - (\ell - 1)\sqrt{2} \in (\sqrt{2} - 1, \sqrt{2})$, and therefore $f(k) \in (\sqrt{2} - 1, \sqrt{2})$. Conversely, by considering $\sqrt{2}\lceil \frac{k}{\sqrt{2}} \rceil$ we notice that $f(k) \in (\sqrt{2} - 1, \sqrt{2})$, then we such $\sqrt{2}\lceil \frac{k}{\sqrt{2}} \rceil - k < 1$. Therefore $k \in H$ if and only if $f(k) \in (\sqrt{2} - 1, \sqrt{2})$.

Next, consider $a_1 < \cdots < a_k$ (with $k \geq 2$) with $a_j - a_i \notin H$ for any i, j. We'll show the following:

$$0 < \sum_{i=1}^{k-1} f(a_{i+1} - a_i) \leq \sqrt{2} - 1$$

We'll establish this via induction on k, focusing only on the right inequality since the left is trivial (f only takes nonnegative values with equality only when $a_{i+1} - a_i$ is an integer multiple of $\sqrt{2}$, which is not the case since $\sqrt{2}$ is irrational): for $k = 2$ it's rather immediate because $a_{i+1} - a_i \notin H$. Suppose that for some k we have the inequality hold true as above. Consider an additional term:

$$\sum_{i=1}^{k} f(a_{i+1} - a_i) = \sum_{i=1}^{k-1} f(a_{i+1} - a_i) + f(a_{k+1} - a_k)$$

We know that $f(a_{k+1} - a_k) \in (0, \sqrt{2} - 1)$. If $\sum_{i=1}^{k} f(a_{i+1} - a_i) > \sqrt{2} - 1$, then we have

$$\sqrt{2}-1 < \sum_{i=1}^{k} f(a_{i+1}-a_i) = \sum_{i=1}^{k-1} f(a_{i+1}-a_i)+f(a_{k+1}-a_k) \leq (\sqrt{2}-1)+(\sqrt{2}-1) = 2(\sqrt{2}-1)$$

However, since $2(\sqrt{2} - 1) < \sqrt{2}$ (given that $\sqrt{2} < 2$), given the additivity of f up to modulo $\sqrt{2}$, we have $f(a_{k+1} - a_1) \in (\sqrt{2} - 1, 2(\sqrt{2} - 1))$, which gives $a_{k+1} - a_1 \in H$. This is a contradiction, and therefore establishes the inequality.

Now, denote $d_i = a_{i+1} - a_i$ for $i = 1, 2, \cdots, k - 1$. We're interested in gauging the magnitude of d_i. Notice now that d_i is an integer, and there exists a constant c_i such that $f(d_i) = d_i - c_i\sqrt{2}$. Now, we have $(d_i - c_i\sqrt{2})(d_i + c_i\sqrt{2}) = d_i^2 - 2c_i^2$. Given that $0 < d_i - c_i\sqrt{2} < \sqrt{2} - 1$ and $d_i > 0$, we have $d_i^2 - 2c_i^2 \geq 1$. Therefore $d_i + c_i\sqrt{2} \geq \frac{1}{d_i - c_i\sqrt{2}} = \frac{1}{f(d_i)}$. We also have $c_i\sqrt{2} \leq d_i$, so $2d_i \geq d_i + c_i\sqrt{2} = \frac{1}{f(d_i)}$, making $d_i \geq \frac{1}{2f(d_i)}$.
Finally, by Cauchy-Schawz inequality, we have

$$(\sum_{i=1}^{k-1} f(d_i))(\sum_{i=1}^{k-1} \frac{1}{2f(d_i)}) \geq (\sum_{i=1}^{k-1} \sqrt{\frac{1}{2}})^2 = \frac{(k-1)^2}{2}$$

and given that $\sum_{i=1}^{k-1} f(d_i) \leq \sqrt{2}-1$, we have $\sum_{i=1}^{k-1} \frac{1}{2f(d_i)} \geq \frac{(k-1)^2}{2(\sqrt{2}-1)}$. Therefore,

$$a_k - a_1 = \sum_{i=1}^{k-1} d_i \geq \sum_{i=1}^{k-1} \frac{1}{2f(d_i)} = \frac{(k-1)^2}{2(\sqrt{2}-1)}$$

and therefore $a_k - a_1 \in \Omega(k^2)$, as desired.