

## Algebra

**A1** Let  $n$  be a positive integer and let  $a_1, \dots, a_{n-1}$  be arbitrary real numbers. Define the sequences  $u_0, \dots, u_n$  and  $v_0, \dots, v_n$  inductively by  $u_0 = u_1 = v_0 = v_1 = 1$ , and  $u_{k+1} = u_k + a_k u_{k-1}$ ,  $v_{k+1} = v_k + a_{n-k} v_{k-1}$  for  $k = 1, \dots, n-1$ .

Prove that  $u_n = v_n$ .

**Solution.** For each  $n$ , denote the set  $S_n$  as the set of subsets of  $\{1, 2, \dots, n\}$  such that a set  $W \in S_n$  iff  $W$  does not contain a pair of neighbouring numbers (that is, if  $i \in W$

then  $i+1 \notin W$ ). We claim that  $u_n = \sum_{\{x_1, \dots, x_k\} \in S_{n-1}} \prod_{i=1}^k a_{x_i}$  (if  $W = \emptyset$  then the term is the

constant term 1). We will go by induction. For  $n = 2$  we have  $u_2 = u_1 + a_1 u_0 = 1 + a_1$ , and  $S_1$  has elements  $\{\}$  and  $\{1\}$ . For  $n = 3$  we have  $(1 + a_1) + a_2$  and it turned out that  $S_2$  has elements  $\{\}$ ,  $\{1\}$  and  $\{2\}$  ( $\{1, 2\}$  is invalid). These settle the base cases. For inductive step, suppose for some  $n$ , we have  $u_n$  and  $u_{n-1}$  following the recurrence relations. This means we have:

$$u_{n+1} = u_n + a_n u_{n-1} = \sum_{\{x_1, \dots, x_k\} \in S_{n-1}} \prod_{i=1}^k a_{x_i} + a_n \cdot \sum_{\{y_1, \dots, y_m\} \in S_{n-2}} \prod_{i=1}^m a_{y_i}$$

We first notice that  $S_{n-1}$  is a subset of  $S_n$  that contains all subsets in  $S_n$  not containing  $n$ , thus  $\{x_1, \dots, x_k\} \in S_{n-1}$  handles this. We also recognize that when  $\{y_1, \dots, y_m\} \in S_{n-2}$ ,  $\{y_1, \dots, y_m\} \cup \{n\} \in S_n$  since  $\max\{y_i\} \leq n-2$ . Conversely, this is also a necessary condition for  $\{y_1, \dots, y_m\} \cup \{n\} \in S_n$  since  $n-1$  cannot be in it. Combining these together, we see that  $S_{n-1}$  handles every element in  $S_n$  not containing  $n$  and  $S_{n-2}$  with  $n$  appended to each element handles every element in  $S_n$  containing  $n$ . This shows that  $u_{n+1}$  fulfills this property too.

Now that  $u_n = \sum_{\{x_1, \dots, x_k\} \in S_{n-1}} \prod_{i=1}^k a_{x_i}$ , we can deduce similarly that

$$v_n = \sum_{\{x_1, \dots, x_k\} \in S_{n-1}} \prod_{i=1}^k a_{n-x_i} = \sum_{\{n-x_1, \dots, n-x_k\} \in S_{n-1}} \prod_{i=1}^k a_{x_i}$$

But then  $\{x_1, \dots, x_k\} \in S_{n-1}$  iff  $\{n-x_1, \dots, n-x_k\} \in S_{n-1}$  by symmetry, so the coefficient of  $a_{x_1} \cdots a_{x_k}$  of  $u_n$  and  $v_n$  are both equal regardless of the combinations of  $x_i$ 's. This proves that  $u_n$  and  $v_n$ .

**A3** (IMO 5) Let  $\mathbb{Q}_{>0}$  be the set of all positive rational numbers. Let  $f : \mathbb{Q}_{>0} \rightarrow \mathbb{R}$  be a function satisfying the following three conditions:

- (i) for all  $x, y \in \mathbb{Q}_{>0}$ , we have  $f(x)f(y) \geq f(xy)$ ;
- (ii) for all  $x, y \in \mathbb{Q}_{>0}$ , we have  $f(x+y) \geq f(x) + f(y)$ ;
- (iii) there exists a rational number  $a > 1$  such that  $f(a) = a$ .

Prove that  $f(x) = x$  for all  $x \in \mathbb{Q}_{>0}$ .

**Solution.** Notice that (i) has the more general form:

$$f(a_1 a_2 \cdots a_n) \leq f(a_1) f(a_2 \cdots a_n) \leq f(a_1) f(a_2) f(a_3 \cdots a_n) \leq \cdots \leq f(a_1) f(a_2) \cdots f(a_n)$$

and same goes to (ii):

$$f(a_1 + \cdots + a_n) \geq f(a_1) + f(a_2 + \cdots + a_n) \geq \cdots \geq f(a_1) + \cdots + f(a_n)$$

We first show that  $f$  produces only positive values. We now have the following:

- Plugging  $x := a, y := 1$  into (i) we get  $f(a)f(1) \geq f(a)$ , or  $af(1) \geq a$ . This means  $f(1) \geq 1$  since  $a > 0$ .
- Next, by considering the general form of (ii), we let  $a_1 = a_2 = \cdots a_1 = 1$  and get  $f(n) \geq nf(1) \geq n$ , so for each positive integer  $n$ ,  $f(n)$  is also positive.
- For any rational number  $p/q$  (with  $p$  and  $q$  integers) we have, from (i),  $f(p/q)f(q) \geq f(p) \geq p$ , so  $f(p/q)f(q) > 0$ . Since  $f(q) > 0$ , we have  $f(p/q) > 0$ .

Having this in mind, we notice from above that  $f(n) \geq n$  so for any positive non-integer rational number  $x$  with  $\lfloor x \rfloor = n$ , we have  $n - \lfloor x \rfloor > 0$ , so  $f(x) \geq f(n) + f(n - \lfloor x \rfloor) > f(n) \geq n$ , so  $f(x) \geq n$  if  $x > n$ . In other words,  $f(x) \geq f(\lfloor x \rfloor)$  with equality holding only when  $x$  is an integer. More generally, for all  $z > x$ , there's  $y > 0$  with  $x + y = z$  (and if  $z, x \in \mathbb{Q}$  then so is  $y$ ) so  $f(z) = f(x + y) \geq f(x) + f(y) > f(x)$ , so  $f$  is also increasing. We therefore have  $f(x) \geq f(\lfloor x \rfloor) \geq \lfloor x \rfloor > x - 1$ , in particular.

With this, we consider the sequences  $f(a^n), n = 1, 2, \dots$ . From the generalized (i), we can plug  $a_1 = \cdots a_n = a$  to get  $a^n = f(a)^n \geq f(a^n)$ . To claim that equality must hold, suppose otherwise, and  $f(a^n) < a^n$  for some  $n > 0$ . Then by (i) we have the following:

$$a^m f(a^n) \geq f(a^n) f(a^m) \geq f(a^{n+m})$$

Let  $\epsilon = a^n - f(a^n) > 0$ , then  $f(a^{n+m}) \leq a^m(a^n - \epsilon)$ . Given that  $a > 1$ , for  $m$  sufficiently large we have  $a^m \epsilon > 1$ . This means that  $f(a^{n+m}) \leq a^m(a^n - \epsilon) < a^{m+n} - 1$ . But this contradicts what we had before:  $f(x) > x - 1$ . Therefore,  $f(a^m) = a^m$  for all positive integers  $m$ .

Next, we show that  $f(x) \leq x$  for all  $x$ . Suppose otherwise, then for each integer  $n$  the generalized form of (ii) gives  $f(nx) = nf(x)$  by plugging  $a_i = x, \forall i = 1, \dots, n$ . If  $\delta > 0$  is such that  $f(x) = x + \delta$ , then  $f(nx) \geq n(x + \delta)$  and for sufficiently large  $n$  we have  $f(nx) \geq nx + 1$ . Choose one such  $n$ , for now. Now by (ii) again, for each positive integer  $m$  we have  $f(nx + m) \geq f(nx) + m \geq nx + m + 1$ . Since  $a > 1$  again, there exists  $p$  such that  $a^p > nx + 1$ , and let  $m$  be the maximum integer with  $nx + m \leq a^p$ , i.e.  $nx + m + 1 > a^p$ . We have  $nx + m \leq a^p$ , but  $f(nx + m) \geq nx + m + 1 > a^p = f(a^p)$ . This contradicts our previous claim that  $f$  is increasing.

Therefore  $f(n) \leq n$  for each positive integer  $n$ , and coupled with  $f(n) \geq n$  from before we get  $f(n) = n$  for each positive integer  $n$ . Finally for each rational number  $p/q$  we have  $f(q)f(p/q) \geq f(p)$ , so  $qf(p/q) \geq p$ , or  $f(p/q) \geq p/q$ . But we have shown that  $f(p/q) \leq p/q$  must hold, too hence  $f(p/q) = p/q$  for each rational number  $p/q$ .

**A4** Let  $n$  be a positive integer, and consider a sequence  $a_1, a_2, \dots, a_n$  of positive integers. Extend it periodically to an infinite sequence  $a_1, a_2, \dots$  by defining  $a_{n+i} = a_i$  for all  $i \geq 1$ . If

$$a_1 \leq a_2 \leq \cdots \leq a_n \leq a_1 + n$$

and

$$a_{a_i} \leq n + i - 1 \quad \text{for } i = 1, 2, \dots, n,$$

prove that

$$a_1 + \cdots + a_n \leq n^2.$$

**Solution.** The fact that  $a_{a_1} \leq n$  means that there exists  $i \in [1, n]$  with  $a_i \leq n$  and since  $a_1 \leq a_2 \leq \cdots a_n, a_1 \leq n$ . Next, since  $a_1, a_2, \dots, a_{a_1} \leq n$ , we can consider the numbers  $a_{a_1}, a_{a_2}, \dots, a_{a_{a_1}}$ , which are all smaller than  $a_n$ . Also, each number is bounded by all the terms following itself, so if  $x \leq a_k$  we have  $a_x \leq a_{a_k} \leq n + k - 1$ . We can now split the

numbers up into the following:

$$\begin{aligned}
\sum_{i=1}^n a_i &= \sum_{i=1}^{a_1} a_i + \sum_{i=a_1+1}^{a_2} a_i + \cdots + \sum_{i=a_{a_1}+1}^n a_i \\
&\leq \sum_{i=1}^{a_1} a_i + \sum_{i=a_1+1}^{a_2} (n+1) + \cdots + \sum_{i=a_{a_1}+1}^n (a_1+n) \\
&= \sum_{i=1}^{a_1} a_i + (n+1)(a_2 - a_1) + (n+2)(a_3 - a_2) + \cdots + (n+(a_1-1))(a_{a_1} - a_{a_1-1}) \\
&\quad + (n+a_1)(n - a_{a_1}) \\
&= \sum_{i=1}^{a_1} a_i + (n+1)(a_2 - a_1) + n \left( \sum_{i=2}^{a_1} a_i - a_{i-1} \right) + n(n - a_{a_1}) - a_1 + \sum_{i=2}^{a_1} a_i((i-1) - i) + na_1 \\
&= \sum_{i=1}^{a_1} a_i + n(n - a_1) - \sum_{i=1}^{a_1} a_i + na_1 \\
&= n^2
\end{aligned}$$

as desired.

- A5** Let  $\mathbb{Z}_{\geq 0}$  be the set of all nonnegative integers. Find all the functions  $f : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$  satisfying the relation

$$f(f(f(n))) = f(n+1) + 1$$

for all  $n \in \mathbb{Z}_{\geq 0}$ .

**Answer.** There are two family of functions. One is  $f(n) = n + 1$ . The other is given by:  $f(2) = 3, f(3) = 0, f(0) = 1, f(1) = 6$ , and  $f(n+4) = f(n) + 4$  for all  $n \geq 0$ .

**Solution.** The key idea is to analyze  $f^k(n)$ , i.e.  $f$  applied to  $n$  for  $k$  times. The condition gives  $f^3(n) = f(n+1) + 1$ .

We first notice the following:

$$f^4(n+1) = f^3(f(n+1)) = f(f(n+1) + 1) + 1 = f(f^3(n)) + 1 = f^4(n) + 1$$

so starting from  $n = 0$  we can inductively deduce that  $f^4(n) = f^4(0) + n$ . This would also mean that for all integers  $m \neq n$  we have  $f^4(m) \neq f^4(n)$ , which means  $f^4$  is injective. It then follows that  $f$  must also be injective. If  $f$  were to be surjective, then so is  $f^3$ . In particular, there's  $n_1$  with  $f(n_1) = 0$ ,  $n_2$  with  $f(n_2) = n_1$  and  $n_3$  with  $f(n_3) = n_2$ . This gives  $f^3(n_3) = 0$  and therefore  $f(n_3 + 1) = -1$ , contradicting that  $f$  only takes nonnegative values.

For convenience we denote  $f^k(\mathbb{N}_{\geq 0})$  as  $\{f^k(n) : n \in \mathbb{N}_{\geq 0}\}$ . Since  $f^k(\mathbb{N}_{\geq 0}) \subseteq f^\ell(\mathbb{N}_{\geq 0})$  for all  $k \leq \ell$  and  $f^4(\mathbb{N}_{\geq 0}) = \mathbb{N}_{\geq 0} \setminus \{0, 1, \dots, f^4(0) - 1\}$ , we know that  $\mathbb{N}_{\geq 0} \setminus f(\mathbb{N}_{\geq 0})$  is finite. Let  $\mathbb{N}_{\geq 0} \setminus f(\mathbb{N}_{\geq 0}) = \{x_1, \dots, x_m\}$ . We show by induction that  $|\mathbb{N}_{\geq 0} \setminus f^k(\mathbb{N}_{\geq 0})| = mk$ , with the base case given by  $k = 1$ . In fact, we will verify that for each  $m \geq 1$ ,  $\mathbb{N}_{\geq 0} \setminus f^{k+1}(\mathbb{N}_{\geq 0}) = (\mathbb{N}_{\geq 0} \setminus f^k(\mathbb{N}_{\geq 0})) \cup \{f^k(x_1), \dots, f^k(x_m)\}$ . There are a few things to take care of:

- $f^k(x_1), \dots, f^k(x_m)$  are all distinct since  $f$  is injective (and so is  $f^k$ ). They are also disjoint from  $\mathbb{N}_{\geq 0} \setminus f^k(\mathbb{N}_{\geq 0})$  since these are values in  $f^k$ .
- Suppose for some  $x_i$ , there's some  $y$  with  $f^{k+1}(y) = f^k(x_i)$ . As mentioned,  $f^k$  is injective, so  $f^k(f(y)) = f^k(x_i)$  means  $f(y) = x_i$ . This contradicts that  $x_i$  is not a value of  $f$  and so  $\{f^k(x_1), \dots, f^k(x_m)\} \subseteq \mathbb{N}_{\geq 0} \setminus f^{k+1}(\mathbb{N}_{\geq 0})$ .
- Finally if  $y \neq x_i$  then  $y = f(z)$  for some  $z$  and therefore  $f^k(y) = f^{k+1}(z)$ .

This justifies the claim of the set size.

Another realization is that, by the given identity,  $f^3(\mathbb{N}_{\geq 0}) = f(\mathbb{N}_{\geq 0} + 1) + 1$ . To determine  $f(\mathbb{N}_{\geq 0} + 1) + 1$ , we notice that:

- $f(\mathbb{N}_{\geq 0}) = \mathbb{N}_{\geq 0} \setminus \{x_1, \dots, x_k\}$ .
- $f(\mathbb{N}_{\geq 0} + 1) = \mathbb{N}_{\geq 0} \setminus \{x_1, \dots, x_k, f(0)\}$ .
- $(\mathbb{N}_{\geq 0} + 1) + 1 = \mathbb{N}_{\geq 0} \setminus \{0, x_1 + 1, \dots, x_k + 1, f(0) + 1\}$

and notice that 0 is different from  $x_1 + 1, \dots, x_k + 1, f(0) + 1$  (since these numbers are all  $\geq 1$ ), and  $f(0)$  is also different from  $x_1, \dots, x_k$  ( $x_i$ 's are not values of  $f$  but  $f(0)$  is). Thus  $|\mathbb{N}_{\geq 0} \setminus f(\mathbb{N}_{\geq 0} + 1) + 1| = k + 2$ . Combined with  $|\mathbb{N}_{\geq 0} \setminus f^3(\mathbb{N}_{\geq 0})| = 3k$  we have  $3k = k + 2$ , so  $k = 1$ .

Now that we have  $k = 1$ , it's rather easy so see how to proceed. Since  $f^4(\mathbb{N}_{\geq 0}) = \{n \in \mathbb{N}_{\geq 0} : n \geq f^4(0)\}$ , we have  $f^4(0) = 4k = 4$ , so  $f^4(n) = n + 4$ . This also means  $f$  cannot have any fixed point. That is, there's no  $n$  with  $f(n) = n$ , as we will see later. In an easier way, we have one value  $x_1$  not a value of  $f$ , which also means  $f(x_1), f^2(x_1), f^3(x_1)$  are not a value of  $f^4$ . By the injectivity of  $f$  we also have  $\{x_1, f(x_1), f^2(x_1), f^3(x_1)\} = \{0, 1, 2, 3\}$ . Now we have a few cases:

- If  $x_1 = 0$ , then  $f(1) + 1 = f^3(0) \leq 3$  so  $f(1) \leq 2$ .  $f(1) = 0$  is impossible since here 0 is not a value of  $f$ .  $f(1) = 1$  implies  $f^k(1) = 1$  for all  $k$ , contradicting  $f^4(1) = 1 + 4 = 5$ . Thus we can only have  $f(1) = 2$ , i.e.  $f^3(0) = 3$ . which means the sequence  $(x_1, f(x_1), f^2(x_1), f^3(x_1))$  must be in the form  $(0, 1, 2, 3)$ . We will deal with this later.
- If  $x_1 = 1$ , then by  $f(2) + 1 = f^3(x_1) \leq 3$  we have  $f(2) \leq 2$ . Again by the similar logic above,  $f(2) \neq 2$  so  $f(2) = 0$  or  $f(2) = 1$ . Since 1 is not a value of  $f$  here, we can only have  $f(2) = 0$ . But this forces  $f^3(1) = f(2) + 1 = 0 + 1 = 1$ , contradicting that 1 is not a value of  $f$  (and hence not a value of  $f^3$ ).
- If  $x_1 = 2$ , by the similar logic above we have by the same logic above we have  $f(3) \leq 2$ . Again 2 is not a value of  $f$  here so we have  $f(3) = 0$  or  $f(3) = 1$ . In the first case,  $f(3) = 0$  so  $f(2) = 1$ , which means the sequence  $(x_1, f(x_1), f^2(x_1), f^3(x_1))$  has the form  $(2, 3, 0, 1)$ . We will come back and deal with this later. In the second case,  $f(3) = 1$  so  $f(2) = 2$ , again contradicting  $f$  cannot have fixed point.
- If  $x_1 = 3$  then we have  $f(4) \leq 2$ . Nevertheless,  $f(4) = f^5(0) \in f^4(\mathbb{N}_{\geq 0})$  so  $f(4) \geq 4$ , which is a contradiction.

So our sequence of  $(x_1, f(x_1), f^2(x_1), f^3(x_1))$  must go in the sequence  $(0, 1, 2, 3)$  or  $(2, 3, 0, 1)$ . With  $f^4(n) = f^4(0) + n = n + 4$ , for each  $n$  and  $k$  we have  $f^{4k}(n) = n + 4k$  (by repeatedly apply  $f^4$ , which is  $k$  times addition of 4). This means that  $f(n + 4k) = f(f^{4k}(n)) = f(n) + 4k$ , too, so  $f$  is uniquely determined by  $f(0), f(1), f(2), f(3)$ . In the first case  $(0, 1, 2, 3)$  we have  $f(n) = n + 1$ , which works easily. In the trickier second case, the function is extrapolated in the form  $(2, 3, 0, 1, 6, 7, 4, 5, \dots)$ . We now have  $f^3(2) = 1 = 0 + 1 = f(3) + 1$ ,  $f^3(3) = f^4(2) = 6 = 5 + 1 = f(4) + 1$ ,  $f^3(0) = 7 = 6 + 1 = f(1) + 1$  and  $f^3(1) = 4 = 3 + 1 = f(0) + 1$ . Also if  $f^3(n) = f(n + 1) + 1$  then  $f^3(n + 4) = f^3(f^4(n)) = f^4(f(n) + 1) = f(n) + 1 + 4 = f(n) + 5 = f(n + 4) + 1$  (since  $f(n) + 4 = f^4(f(n)) = f(f^4(n)) = f(n + 4)$ ) so the fact that our identity holds for  $n = 0, 1, 2, 3$  means it holds for all  $n$  too. This also gives the following construction for  $f$ :  $f(n) = n + 1$  for  $n$  even,  $n - 3$  for  $n \equiv 3 \pmod{4}$  and  $n + 5$  for  $n \equiv 1 \pmod{4}$ , as claimed.

## Combinatorics

- C1** Let  $n$  be an positive integer. Find the smallest integer  $k$  with the following property;  
Given any real numbers  $a_1, \dots, a_d$  such that  $a_1 + a_2 + \dots + a_d = n$  and  $0 \leq a_i \leq 1$  for

$i = 1, 2, \dots, d$ , it is possible to partition these numbers into  $k$  groups (some of which may be empty) such that the sum of the numbers in each group is at most 1.

**Answer.**  $2n - 1$ .

**Solution.** To see that  $k = 2n - 1$  is necessary, consider  $d = 2n - 1$  and  $a_i = \frac{n}{2n-1}$  for all  $i$ . Since each  $a_i = \frac{n}{2n-1} > \frac{n}{2n} = \frac{1}{2}$ , each  $a_i$  must be in its own isolated group, hence  $2n - 1$  groups are necessary.

To see that  $2n - 1$  is sufficient, we sort the numbers into  $a_1 \geq a_2 \geq \dots \geq a_d$  and place them into the  $2n - 1$  groups in the following algorithmic manner: we let  $a_1, a_2, \dots, a_{2n-1}$  go into their individual groups (assuming  $d \geq 2n - 1$  otherwise the problem is trivial). Then for each  $j \geq 2n$ , we place  $a_j$  sequentially into the group that has sum of numbers at most  $1 - a_j$ .

To show that we can always find those groups, we first note that  $\sum_{i=1}^j a_i \leq \sum_{i=1}^d a_i = n$  and since  $a_i \geq a_j$  for all  $i \leq j$ ,  $n \geq \sum_{i=1}^j a_i \geq \sum_{i=1}^j a_j = ja_j \geq 2na_j$  (since  $j \geq 2n$ ), and therefore  $a_j \leq \frac{1}{2}$ . The above identity also implies  $\sum_{i=1}^{j-1} a_i \leq n - a_j$ , and the first summation is the total size (i.e. sum of numbers) of groups by considering only  $a_1, \dots, a_{j-1}$ . The average group size under this consideration is currently  $\frac{\sum_{i=1}^{j-1} a_i}{2n-1} \leq \frac{n - a_j}{2n-1}$ . But also notice the following:

$$\frac{\sum_{i=1}^{j-1} a_i}{2n-1} + a_j \leq \frac{n - a_j}{2n-1} + a_j = \frac{n + (2n-2)a_j}{2n-1} \leq \frac{n + (2n-2) \cdot \frac{1}{2}}{2n-1} = 1$$

so the average of the size of groups are currently at most  $1 - a_j$ . This means, there must be a group of size at most  $1 - a_j$ , so we are done.

**C2** (IMO 2) In the plane, 2013 red points and 2014 blue points are marked so that no three of the marked points are collinear. One needs to draw  $k$  lines not passing through the marked points and dividing the plane into several regions. The goal is to do it in such a way that no region contains points of both colors.

Find the minimal value of  $k$  such that the goal is attainable for every possible configuration of 4027 points.

**Answer.** 2013.

**Solution.** To show that 2013 is a lower bound, consider the regular polygon of 4027 vertices  $A_i, i = 1, 2, \dots, 4027$ . Moreover, let  $A_i$  blue if  $i$  odd, and  $A_i$  red if  $i$  even. Since each  $A_i, A_{i+1}$  are of different colour for  $i = 1, \dots, 4026$ , there must be at least a line passing through the segment  $A_i A_{i+1}$ . On the other hand, each line can only pass through two sides of a regular polygon, so at least  $\frac{4026}{2}$  lines needed.

Now to show that 2013 is sufficient, we generalize the problem into any configuration of  $2k + 1$  points (with no three collinear) with binary colours, and show that  $k$  lines suffice. To deal with  $k = 1$  as base case, if all three points are the same colour we are easily done. Otherwise, there must be two points (say  $A_1$  and  $A_2$ ) of the same colour, and a point (say  $B$ ) of the other colour. Now the segments  $BA_1$  and  $BA_2$  cannot be collinear because no three lines are collinear, so we can draw a line cutting through the internal (open) segments  $BA_1$  and  $BA_2$ .

For inductive step, consider an arbitrary configuration of  $2k + 1$  points, and the convex hull  $\mathcal{P}$  of the configuration. Consider any line  $\ell$  determined by any side of  $\mathcal{P}$ ; all other lines lie on the same side of  $\ell$ . If any of the sides have both points of the same colour,

then we can draw a line to isolate the points from the others, and the rest follows from induction hypothesis: draw  $k - 1$  lines to handle the case for the remaining  $2k - 1$  points.

Otherwise, the points on  $\mathcal{P}$  must be alternating in colours, so we choose a side with a red vertex  $R$  and a blue vertex  $B$  randomly and consider the  $k - 1$  lines that separate the remaining  $2k - 1$  points into regions with no region containing points of both colours. If  $R$  and  $B$  are in the same region, then this region must contain either only red points and blue points other than  $R$  and  $B$  (we call this region “red region” or “blue region” in these respective cases). Since  $R$  and  $B$  are on the convex hull, we can draw a line to isolate  $R$  if this region is blue, or a line to isolate  $B$  if this region is red. Otherwise,  $R$  and  $B$  are in different regions, and we can now draw a line to isolate  $R$  and  $B$  from the rest of the  $2k - 1$  points. This way,  $R$  and  $B$  are each isolated, and we are done.

- C4** Let  $n$  be a positive integer, and let  $A$  be a subset of  $\{1, \dots, n\}$ . An  $A$ -partition of  $n$  into  $k$  parts is a representation of  $n$  as a sum  $n = a_1 + \dots + a_k$ , where the parts  $a_1, \dots, a_k$  belong to  $A$  and are not necessarily distinct. The number of different parts in such a partition is the number of (distinct) elements in the set  $\{a_1, a_2, \dots, a_k\}$ . We say that an  $A$ -partition of  $n$  into  $k$  parts is optimal if there is no  $A$ -partition of  $n$  into  $r$  parts with  $r < k$ . Prove that any optimal  $A$ -partition of  $n$  contains at most  $\sqrt[3]{6n}$  different parts.

**Solution.** Let  $a_1 < a_2 < \dots < a_k$  be the different parts in an optimal partition of  $n$ , and let  $f_1, \dots, f_k$  be the frequency of the numbers in the representation (that is,  $n = \sum_{i=1}^n a_i f_i$ ).

We first see that there cannot be  $0 \leq i < j < k$  with  $a_{j+1} - a_j = a_{i+1} - a_i$ . Otherwise, we have  $a_i + a_{j+1} = a_{i+1} + a_j$ . Now suppose that  $m = \min\{f_i, f_{i+1}, f_j, f_{j+1}\}$ . Consider the following:

$$f_i a_i + f_{i+1} a_{i+1} + f_j a_j + f_{j+1} a_{j+1} = (f_i + m) a_i + (f_{i+1} - m) a_{i+1} + (f_j - m) a_j + (f_{j+1} + m) a_{j+1} \dots (1)$$

$$f_i a_i + f_{i+1} a_{i+1} + f_j a_j + f_{j+1} a_{j+1} = (f_i - m) a_i + (f_{i+1} + m) a_{i+1} + (f_j + m) a_j + (f_{j+1} - m) a_{j+1} \dots (2)$$

If  $m = f_{i+1}$  or  $m = f_j$ , then in (1) one of these coefficients are zero, while the rest remain positive. This means that there is a representation of  $n$  with the original set of parts with either  $f_{i+1}$  or  $f_j$  removed. Similarly, if  $m = f_{j+1}$  or  $m = f_i$  then there is a representation of  $n$  with the original set of parts with either  $f_i$  or  $f_{j+1}$  removed. Thus the pairwise distance of the neighbouring elements must be different, which also implies the following:

$$a_j = a_1 + \sum_{i=1}^{j-1} a_{i+1} - a_i \geq a_1 + \sum_{i=1}^{j-1} i = a_1 + \frac{j(j-1)}{2}$$

and since  $a_1 > 1$  (valid assumption, because otherwise we can have  $f_1 = 1$  and this gives the representaiton of 1 part),  $a_j \geq 2 + \frac{j(j-1)}{2}$ . Since each  $f_i \geq 1$ , we have

$$n \geq \sum_{i=1}^k a_i \geq \sum_{i=1}^k \left( 2 + \frac{j(j-1)}{2} \right) = 2k + \frac{k(k+1)(2k+1)}{12} - \frac{k(k-1)}{4} > \frac{k^3}{6}$$

and therefore  $k < \sqrt[3]{6n}$ .

- C6** In some country several pairs of cities are connected by direct two-way flights. It is possible to go from any city to any other by a sequence of flights. The distance between two cities is defined to be the least possible numbers of flights required to go from one of them to the other. It is known that for any city there are at most 100 cities at distance exactly three from it. Prove that there is no city such that more than 2550 other cities have distance exactly four from it.

**Solution.** Let's ditch the city and flight notion and instead use the graph theory ones: vertices and edge. Let  $V$  and  $E$  be the set of vertices, in the said graph, respectively.

Consider any node  $v_0$ , and consider the set  $S_i : \{v \in V : d(v, v_0) = i\}$  for  $i = 1, 2, 3, 4$  (where  $d(u, v)$  denoting distance of  $u$  and  $v$ ). We note the following:

- The non-strict triangle inequality holds: if we can go from  $u$  to  $v$  in  $d(u, v)$  steps and  $v$  to  $w$  in  $d(v, w)$  steps, using the shortest route from  $u$  to  $v$  and then from  $v$  to  $w$  we can go from  $u$  to  $w$  in  $d(u, w)$  steps. This means,  $d(u, w) \leq d(u, v) + d(v, w)$ .
- Consider any  $v \in S_n$ , and let  $v_0, v_1, \dots, v_n = v$  be such a path of length  $n$ . For each  $0 \leq i \leq n$ , this path from  $v_0$  to  $v_i$  has length  $i$  so for each  $i$ ,  $d(v_0, v_i) \leq i$ . Conversely,  $n = d(v_0, v_n) \leq d(v_0, v_i) + d(v_i, v_n) \leq d(v_0, v_i) + (n - i)$  since the path  $v_i, \dots, v_n$  takes  $n - i$  steps. This means  $d(v_0, v_i) \geq i$  and therefore  $d(v_0, v_i) = i$  for all  $i \leq n$ .

Now let  $S_4 = v_1, v_2, \dots, v_m$  and for each  $v_i$  consider its shortest path from  $v_0$ :  $v_0 = v_{0,i}, v_{1,i}, v_{2,i}, v_{3,i}, v_{4,i} = v_i$ . Consider all such  $v_{0,i}, v_{1,i}, v_{2,i}, v_{3,i}$ . If, say,  $v_{3,i} = v_{3,j}$  then  $v_{0,i}, v_{1,i}, v_{2,i}, v_{3,j}, v_{4,j}$  is also a shortest path. Similarly  $v_{2,i} = v_{2,j}$  implies that  $v_{0,i}, v_{1,i}, v_{2,j}, v_{3,j}, v_{4,j}$  is a shortest path. This means we can assume that if  $v_{1,i} \neq v_{1,j}$  then  $v_{2,i} \neq v_{2,j}$  and  $v_{3,i} \neq v_{3,j}$ . Consider the set  $T : \{v_{1,1}, \dots, v_{1,m}\}$ , and suppose that it has  $n$  distinct elements. Consider, now, an arbitrary vertex  $v_{1,i}$  and let  $T_i = \{v_{4,j} : v_{1,i} = v_{1,j}\} \subseteq S_4$ . From the definition of shortest path, for each  $v_{4,j} \in T_i$  we have  $d(v_{1,j}, v_{4,j}) = 4 - 1 = 3$ . Consider, also, any  $k \neq i$  with  $v_{1,k} \neq v_{1,i}$ , and the vertices  $v_{2,k}, v_{3,k}$  and  $v_{4,k}$ . We claim that  $v_{1,j}$  has a distance exactly three with at least one of those vertices. Now,  $d(v_{1,j}, v_{4,k}) \geq d(v_0, v_{4,k}) - d(v_0, v_{1,j}) = 4 - 1 = 3$ . If it's greater than 3 then we have  $4 \leq d(v_{1,j}, v_{4,k}) \leq d(v_{1,j}, v_{3,k}) + d(v_{3,k}, v_{4,k}) = d(v_{1,j}, v_{3,k}) + 1$ , so  $d(v_{1,j}, v_{3,k}) \geq 3$ . Again if  $d(v_{1,j}, v_{3,k}) > 3$  we have  $d(v_{1,j}, v_{2,k}) \geq 3$  by a similar logic, but then by triangle inequality again  $d(v_{1,j}, v_{2,k}) \leq d(v_{1,j}, v_0) + d(v_0, v_{2,k}) = 1 + 2 = 3$  so the equality must hold.

To summarize, let  $v_{1,a_1}, \dots, v_{1,a_n}$  be distinct elements in  $T$ . For each  $a_i$ , it has distance exactly 3 with each vertex in  $|T_{a_i}|$ , and with at least one member among  $v_{2,a_j}, v_{3,a_j}$  and  $v_{4,a_j}$  for any  $j \neq i$ . By our assumption, for each  $j \neq i$  we have  $v_{x,a_j} \neq v_{y,a_i}$  for  $x, y \in \{1, 2, 3, 4\}$ . This means that it has distance 3 with at least  $|T_{a_i}| + (n - 1)$  elements, which is at most 100. This means  $|T_{a_i}| \leq 100 - n + 1$ . In addition,  $T_{a_1}, \dots, T_{a_n}$  together represent the partition of  $S_4$ , we have

$$|S_4| = \sum_{i=1}^n |T_{a_i}| \leq \sum_{i=1}^n (100 - n + 1) = n(100 - n + 1)$$

The last expression obtains its maximum when  $n = 101/2$ ; for  $n$  integer the maximum is when  $n = 50$  or  $51$ , in which case  $n(100 - n + 1) = 50 \times 51 = 2550$ . Thus  $|S_4| \leq 2550$  as desired.

## Geometry

- G1** (IMO 4) Let  $ABC$  be an acute triangle with orthocenter  $H$ , and let  $W$  be a point on the side  $BC$ , lying strictly between  $B$  and  $C$ . The points  $M$  and  $N$  are the feet of the altitudes from  $B$  and  $C$ , respectively. Denote by  $\omega_1$  is the circumcircle of  $BWN$ , and let  $X$  be the point on  $\omega_1$  such that  $WX$  is a diameter of  $\omega_1$ . Analogously, denote by  $\omega_2$  the circumcircle of triangle  $CWM$ , and let  $Y$  be the point such that  $WY$  is a diameter of  $\omega_2$ . Prove that  $X, Y$  and  $H$  are collinear.

**Solution.** Consider, now, the point  $Z$  (possibly  $Z = H$ ) which is the intersection of the circles  $BWN$  and  $CWM$  that is not  $W$ . As  $WX$  is a diameter of  $\omega_1$ ,  $WZ$  and  $ZX$  are perpendicular to each other. Similarly,  $WZ$  and  $ZY$  are perpendicular to each other. This means that both  $X$  and  $Y$  are on the perpendicular to  $WZ$  passing through  $Z$ , say  $\ell$ . In addition, from  $\angle BYC = \angle BXC = 90^\circ$ ,  $BYXC$  is cyclic, so  $A$  lies on the radical axis of the circles  $BWN$  and  $CWM$ , which is  $WZ$ . Therefore,  $A, W, Z$  are collinear.

By Miquel's theorem,  $Z$  also lies on the circumcircle of  $AMN$ , and since  $\angle AMH = \angle ANH = 90^\circ$ , we have  $AH$  a diameter of this circle, and therefore  $AZH = 90^\circ$ . We have shown that  $A, W, Z$  are collinear and  $X, Y$  on  $\ell$ . Moreover,  $AZ \perp ZH$  so  $H$  is also on  $\ell$ . Thus  $X, Y, H$  are collinear.

- G2** Let  $\omega$  be the circumcircle of a triangle  $ABC$ . Denote by  $M$  and  $N$  the midpoints of the sides  $AB$  and  $AC$ , respectively, and denote by  $T$  the midpoint of the arc  $BC$  of  $\omega$  not containing  $A$ . The circumcircles of the triangles  $AMT$  and  $ANT$  intersect the perpendicular bisectors of  $AC$  and  $AB$  at points  $X$  and  $Y$ , respectively; assume that  $X$  and  $Y$  lie inside the triangle  $ABC$ . The lines  $MN$  and  $XY$  intersect at  $K$ . Prove that  $KA = KT$ .

**Solution.** Denote  $\ell$  as the perpendicular bisector of segment  $AT$ . We first show that the perpendicular bisectors of  $AB$  and  $AC$  (namely  $\ell_C, \ell_B$ ) are symmetric about  $\ell$ . To see this, if  $O$  is the center of  $\omega$ , we have  $OA = OT$  so  $O \in \ell$ . Moreover,  $AT$  is the internal angle bisector of the angle  $\angle BAC$ , and thus:

$$\angle(\ell_C, AT) = \angle(\ell_C, AB) + \angle(AB, AT) = 90^\circ + \angle(AT, AC) = \angle(AC, \ell_B) + \angle(AT, AC) = \angle(AT, \ell_B)$$

and with the fact that the lines  $\ell_C$  and  $\ell_B$  intersect at  $O$  (which is on  $\ell$ ),  $\ell_C$  and  $\ell_B$  are indeed symmetric in  $\ell$ .

Next, we show that  $M$  and  $X$  are also symmetric about  $\ell$ . Let  $M'$  to be the reflection of  $M$  in  $\ell$ . Then since  $M$  is on  $\ell_C$ ,  $M'$  is on  $\ell_B$  and therefore  $M'A = M'C$ . Moreover,  $AMM'T$  is an isosceles trapezoid, hence cyclic. It follows that  $M'$  is one of the intersections of the line  $\ell_B$  and the circle  $AMT$ . Since  $\angle AMX > \angle AMO = 90^\circ$ ,  $\angle AMX$  and similarly  $\angle ANX$  are obtuse. It follows that the other intersection of  $\ell$  and circle  $AMT$  must be on the different side with  $X$  with respect to  $AC$ , so it reduces to consider the half-line of  $\ell_B$  passing through  $O$ . We see that the reflection of this half line intersects  $M$  (for  $X$  to stay inside  $\triangle ABC$ ), so  $M'$  is on the half-line of  $\ell_B$  passing through  $O$ , and therefore  $M = X'$ . Similarly  $N$  and  $Y$  are symmetric in  $\ell$ .

The above argument ready implies that  $MXYN$  is an isosceles trapezoid with parallel sides  $MX$  and  $NY$ . It then follows that the intersection of  $MN$  and  $XY$  (i.e.  $K$ ) lies on  $\ell$ . Therefore  $KA = KT$ .

- G3** In a triangle  $ABC$ , let  $D$  and  $E$  be the feet of the angle bisectors of angles  $A$  and  $B$ , respectively. A rhombus is inscribed into the quadrilateral  $AEDB$  (all vertices of the rhombus lie on different sides of  $AEDB$ ). Let  $\varphi$  be the non-obtuse angle of the rhombus. Prove that  $\varphi \leq \max\{\angle BAC, \angle ABC\}$ .

**Solution.** Denote  $PQRS$  as the rhombus, with  $P, Q, R, S$  lying on  $AB, BD, DE, EA$  respectively. Now let  $U$  and  $V$  be the intersections the diagonal  $QS$  and the lines  $AD$  and  $BE$ , respectively. Now suppose for a contradiction, that  $\varphi$  is greater than both  $\angle BAC$  and  $\angle ABC$ . Since  $QS$  bisects angle  $\angle RSP$ , we also have  $\angle QSP = \frac{\varphi}{2} > \frac{\angle BAC}{2} = \angle DAB$ . Looking at the quadrilateral  $USAP$ , we have  $\angle USP > \angle UAP$ , which implies that  $S$  lies strictly inside the circumcircle of  $UAP$ , too. This has the following consequences:

- By a similar logic we immediately have  $\angle ASP > \angle AUP$ .
- The opposite angles,  $\angle ASU$  and  $\angle APU$ , have sum greater than  $180^\circ$ .
- Point  $A$  is now outside the circumcircle of  $SUP$ , and therefore  $\angle UPS > \angle UAS$ . Similarly we have  $\angle SAP > \angle SUA$ .

(TL; DR we are tweaking properties of a cyclic quadrilateral on a non-cyclic quadrilateral). Similarly we have  $\angle QPB > \angle QVB$ . If we denote  $I$  and the incenter of  $ABC$  then we



now have the following:

$$\begin{aligned}
\angle SRV &= \angle SPV \\
&= \angle SPQ - \angle VPQ \\
&= 180^\circ - \varphi - \angle VPQ \\
&< 180^\circ - \varphi - \angle VQB \\
&< 180^\circ - \varphi - \frac{\angle ABC}{2}
\end{aligned}$$

and bearing in mind that  $\angle SEV = \angle AEB = 180^\circ - \angle BAC - \frac{\angle ABC}{2}$ . Since  $\varphi > \angle BAC$ , this inequality also implies

$$\angle SRV < 180^\circ - \varphi - \frac{\angle ABC}{2} = 180^\circ - \angle BAC - \frac{\angle ABC}{2} = \angle SEV$$

and therefore by the similar logic  $\angle RES > \angle RSV = \frac{\varphi}{2}$ . Similarly  $\angle RDU > \angle RQU = \frac{\varphi}{2}$  and therefore

$$\angle AED + \angle EDB = \angle AEB + \angle ADB + \angle RES + \angle RDU > \varphi + 360^\circ - \frac{3}{2}\angle CAB - \frac{3}{2}\angle CBA$$

which would mean that  $\angle CER + \angle CDR < \frac{3}{2}\angle CAB + \frac{3}{2}\angle CBA - \varphi$ . But then  $E, R, D$  are collinear so this angle  $\angle CER + \angle CDR$  is supposed to be the same as  $\angle CAB + \angle CBA$  so we have

$$\angle CAB + \angle CBA < \frac{3}{2}\angle CAB + \frac{3}{2}\angle CBA - \varphi$$

which means  $\varphi < \frac{1}{2}\angle CAB + \frac{1}{2}\angle CBA$ . This clearly contradicts that  $\varphi > \angle CAB$  and  $\varphi > \angle CBA$ .

**G4** Let  $ABC$  be a triangle with  $\angle B > \angle C$ . Let  $P$  and  $Q$  be two different points on line  $AC$  such that  $\angle PBA = \angle QBA = \angle ACB$  and  $A$  is located between  $P$  and  $C$ . Suppose that there exists an interior point  $D$  of segment  $BQ$  for which  $PD = PB$ . Let the ray  $AD$  intersect the circle  $ABC$  at  $R \neq A$ . Prove that  $QB = QR$ .

**Solution.** From  $\angle ACB = \angle PBA$  and  $PD = PB$  we have triangles  $PBA$  and  $PCB$  similar, so  $PD^2 = PB^2 = PA \cdot PC$ . In a similar way we can deduce  $AB^2 = AQ \cdot AC = \angle AD \cdot AR$  from the fact that  $\angle QBA = \angle ACB = \angle ARB$  (since  $R$  is on circle  $ABC$ ). Therefore, the quadrilateral  $DQCR$  is cyclic. We can now compute the following:

$$\angle QBR = \angle QBC + \angle CBR = \angle ABC - \angle ABQ + \angle RAC = \angle ABC - \angle ACB + (\angle PDA + \angle DPQ)$$

and the fact  $PD^2 = PA \cdot PC$  also implies  $\angle PDA = \angle PDC$  and therefore

$$\angle QRB = \angle QRD + \angle DRB = \angle QCD + \angle ACB = \angle PDA + \angle ACB$$

but since

$$2\angle ACB = \angle PBA + \angle ABQ = \angle PBQ = \angle PDB = \angle PDQ + \angle PQB = \angle DPQ + \angle ABC$$

and therefore

$$\begin{aligned}
\angle QBR &= \angle ABC - \angle ACB + (\angle PDA + \angle DPQ) \\
&= \angle ABC - \angle ACB + \angle PDA + (2\angle ACB - \angle ABC) \\
&= \angle ACB + \angle PDA \\
&= \angle QRB
\end{aligned}$$

and therefore  $QR = QB$ .

- G5** Let  $ABCDEF$  be a convex hexagon with  $AB = DE$ ,  $BC = EF$ ,  $CD = FA$ , and  $\angle A - \angle D = \angle C - \angle F = \angle E - \angle B$ . Prove that the diagonals  $AD$ ,  $BE$ , and  $CF$  are concurrent.

**Solution.** The fact that their opposite sides have equal length motivates us to try the spiral similarity between them. To be precise, let  $X$  be the intersection of  $AD$  and  $BE$  and we consider the second intersection  $O_1$  of circles  $ABX$  and  $DEX$ , with this intersection being equal to  $X$  only when the two circles are tangent at  $X$ . We first notice the following realizations by directed angles:

$$\angle(AO_1, BO_1) = \angle(AX, BX) = \angle(AD, BD) = \angle(DX, EX) = \angle(DO_1, EO_1)$$

$$\angle(AB, BO_1) = \angle(AX, XO_1) = \angle(DX, XO_1) = \angle(DE, EO_1)$$

$$\angle(AB, AO_1) = \angle(BX, XO_1) = \angle(EX, XO_1) = \angle(DE, DO_1)$$

and therefore triangles  $ABO_1$  and  $DEO_1$  are similar (it's not quite immediate from directed angles since the corresponding angles could either be equal or isupplementary, but then the claim follows that the angles in a triangle are positive and add up to  $180^\circ$ ). Since  $AB = DE$ , these triangles are indeed congruent, and therefore  $AO_1 = DO_1$  and  $BO_1 = EO_1$ . Since the congruent triangles can be obtained by each other via rotation (technically reflection might be involved but that's just rotation by  $180^\circ$ ) we have  $\angle(AB, DE) = \angle(AO_1, DO_1) = \angle(BO_1, EO_1)$ , which in turn motivates the following:

$$\angle(AB, DE) = \angle(AO_1, DO_1) = \angle(AO_1, AD) + \angle(AD, DO_1) = 2\angle(AO_1, AD)$$

and similarly  $\angle(AB, DE) = 2\angle(BO_1, BE)$ .

Now to properly investigate  $\angle(AB, DE)$ , we need the angle condition given above. In essence  $\angle(AB, DE) = \angle(AB, BC) + \angle(BC, CD) + \angle(CD, DE) = \angle B + \angle C + \angle D$  (we may assume that our hexagon  $ABCDEF$  is indeed oriented in this manner). Denoting the angle  $\alpha := \angle A - \angle D = \angle C - \angle F = \angle E - \angle B$ , and bearing in mind that the sum of the six angles  $\angle A$  through  $\angle F$  is  $720^\circ$ , we get  $\angle B + \angle C + \angle D = 360^\circ - \frac{\alpha}{2}$ . Therefore, denoting  $\beta = \frac{\alpha}{2}$  (we want to avoid fractions in directed angles...if you get what I mean) we have  $\angle(AB, DE) = 360^\circ - \beta = -\beta$ , i.e.  $2\angle(BO_1, BE) = -\beta$ .

Consider, now, the intersection  $Y$  of the lines  $BE$  and  $CF$ , and  $O_2$  as the second intersection of circles  $BCY$  and  $EFY$ . Then by the similar logic as above we get  $2\angle(BO_2, BE) = \angle(BC, EF)$  and we also have  $\angle(BC, EF) = \angle C + \angle D + \angle E$  which can be calculated as  $360^\circ + \beta$ , so we now have  $2\angle(BO_2, BE) = \angle(BC, EF) = \beta = -2\angle(BO_1, BE)$ . Thus considering the sides  $O_1$  and  $O_2$  belong to w.r.t.  $BE$  we have  $BE$  as the internal angle bisector of  $\angle O_1BO_2$  and similarly  $O_1E$  is the internal angle bisector of  $\angle O_1EO_2$ . Therefore  $BE$  is the perpendicular bisector of  $O_1O_2$ .

Finally, denote the intersection  $Z$  of the lines  $CF$  and  $AD$ , and  $O_3$  as the second intersection of  $CDZ$  and  $FAZ$ . By the similar logic as above, again,  $AD$  is the perpendicular bisector of  $O_1O_3$  and  $CF$  is the perpendicular bisector of  $O_2O_3$ . Therefore  $AD, BE, CF$  concur at the circumcenter of  $O_1O_2O_3$ , i.e. the point  $X = Y = Z$ .

## Number Theory

- N1** Let  $\mathbb{Z}_{>0}$  be the set of positive integers. Find all functions  $f : \mathbb{Z}_{>0} \rightarrow \mathbb{Z}_{>0}$  such that

$$m^2 + f(n) \mid mf(m) + n$$

for all positive integers  $m$  and  $n$ .

**Answer.** The only such function is the identity function  $f(x) = x$ , where both sides are equal to  $m^2 + n$  so this function works.

**Solution.** By plugging  $m = n = 2$  we have  $4 + f(2) \mid 2f(2) + 2 = 2(4 + f(2)) - 6$ , so  $4 + f(2) \mid 6$ . Since  $f(2) > 0$ , the only possibility is  $f(2) = 2$ . Then, plugging  $m = 2$  gives  $4 + f(n) \mid 2f(2) + n = 4 + n$ . Given that  $4 + f(n)$  and  $4 + n$  are both positive, we have  $4 + f(n) \leq 4 + n$ , so  $f(n) \leq n$ . Finally, plugging  $n = 2$  gives  $m^2 + 2 = m^2 + f(n) \mid mf(m) + 2$ , so  $m^2 \leq mf(m)$ , or  $m \leq f(m)$ . This means,  $f(m) \geq m$  and  $f(m) \leq m$  both hold for all  $m$ , hence  $f(m) = m$  for all  $m$ .

**N2** (IMO 1) Assume that  $k$  and  $n$  are two positive integers. Prove that there exist positive integers  $m_1, \dots, m_k$  such that

$$1 + \frac{2^k - 1}{n} = \left(1 + \frac{1}{m_1}\right) \cdots \left(1 + \frac{1}{m_k}\right).$$

**Solution.** A solution that fixes  $n$  and does induction on  $k$  is possible, as per my own solution to the IMO. But let me give a solution that's based on the following intuition: we want to find a sequence of  $k + 1$  numbers  $n = a_0 < a_1 < \dots < a_k = n + 2^k - 1$  such that for each  $0 \leq i < k$ ,  $a_{i+1} - a_i \mid a_i$ . This way, we have:

$$1 + \frac{2^k - 1}{n} = \frac{n + 2^k - 1}{n} = \prod_{i=1}^k \frac{a_i}{a_{i-1}} = \prod_{i=1}^k \left(1 + \frac{1}{\frac{a_i - a_{i-1}}{a_{i-1}}}\right)$$

and notice that  $\frac{a_i - a_{i-1}}{a_{i-1}}$  is an integer based on our construction. Hence this sequence of  $a_i$ 's gives rise of a valid construction.

To construct such sequence, we first notice that among  $n, n + 1, \dots, n + 2^k - 1$  which is a consecutive sequence of  $2^k$  numbers, exactly one of them is divisible by  $2^k$ . Name this number  $n_0$ . Next, we consider the binary representation of the two numbers:

$$n_0 - n = 2^{c_1} + 2^{c_2} + \dots + 2^{c_m} \quad (n + 2^k - 1) - n_0 = 2^{b_1} + 2^{b_2} + \dots + 2^{b_\ell}$$

with  $c_1 < c_2 < \dots < c_m < k$  and  $b_1 < b_2 < \dots < b_\ell < k$  ( $k$  and  $\ell$  could be 0 in case  $n$  or  $n + 2^k - 1$  coincides with  $n_0$ ). We now construct the sequence based on the following: for  $i \leq m$ , we let  $a_i = a_{i-1} + 2^{c_i}$  and for  $i > m$ ,  $a_i = a_{i-1} + 2^{b_{\ell-i+1}}$ . To see it works, we need the two following observations:

- First, we need  $a_i - a_{i-1} \mid a_i$ . We notice that for each  $j \leq m$ ,  $a_j = n + \sum_{i=1}^j 2^{c_i} = n_0 - \sum_{i=j+1}^m 2^i$ . Since  $n_0$  is divisible by  $2^k$  and  $k > a_i$  for all  $i$ , we have  $2^{\min\{c_i : i \geq j+1\}} = 2^{c_{j+1}}$

so in fact  $a_j - a_{j-1} = 2^{c_j} \mid a_j$ , as desired. As for  $j > m$ ,  $a_j = a_n + \sum_{i=0}^{j-m-1} 2^{b_{\ell-i}}$  so by similar logic,  $2^{\ell-(j-m-1)} \mid a_j$  and thus  $2^{\ell-(j-m)} \mid a_j$ . But  $2^{\ell-(j-m)} = a_j - a_{j-1}$ , so the conclusion follows.

- We can see that  $n_{\ell+m}$  is indeed equal to  $n_0 + 2^{b_1} + 2^{b_2} + \dots + 2^{b_\ell} = n + 2^k - 1$ , so it remains to show that  $\ell + m = k$ . This is the same as showing that  $n_0 - n$  and  $(n + 2^k - 1) - n_0$  have  $k$  ones in total in their binary representation. We however notice that both numbers are at most  $2^k - 1$ , so they can be written with  $k$  digits. Moreover, since they add up to  $2^k - 1$ , all digits within the first  $k$  digits are complementary of each other. This means, half of the  $2k$  digits are 1 and half are zero, and we have a total of  $k$  ones, as desired.

The conclusion follows from the combination of the two claims.

**N3** Prove that there exist infinitely many positive integers  $n$  such that the largest prime divisor of  $n^4 + n^2 + 1$  is equal to the largest prime divisor of  $(n+1)^4 + (n+1)^2 + 1$ .

**Solution.** Denote by  $f(n) = n^2 + n + 1$ . Then  $n^4 + n^2 + 1 = (n^2 + n + 1)(n^2 - n + 1) = f(n)f(n-1)$ . Denote also the largest prime divisor of  $f(n)$  as  $p(n)$ . We first show that  $\gcd(f(n), f(n-1)) = 1$ . If  $k$  is a common divisor of  $f(n) = n^2 + n + 1$  and  $f(n-1) = n^2 - n + 1$ , then  $k \mid 2n$  too. We have  $\gcd(f(n), n) = 1$  (since ) and both  $f(n)$  and  $f(n-1)$  are odd, and therefore  $k = \pm 1$  must hold. Next, if  $p(n) \geq p(n-1)$  and  $p(n) \geq p(n+1)$ , considering that the largest prime divisor of  $n^4 + n^2 + 1$  is  $\max\{p(n), p(n-1)\} = p(n)$ , we have the largest prime divisor of  $n^4 + n^2 + 1$  and  $(n+1)^4 + (n+1)^2 + 1$  being equal, and therefore this  $n$  would work.

It remains to show that there are infinitely many  $n$  such that  $p(n) > p(n-1)$  and  $p(n) > p(n+1)$ . Since the  $\gcd\{f(n), f(n-1)\}$ ,  $p(n)$  cannot be equal to  $p(n-1)$ . Suppose our desired condition doesn't hold, then either  $p(n) < p(n-1)$  for all sufficiently large  $n$ , or  $p(n) > p(n-1)$  for sufficiently large  $n$ . The first case is impossible since all prime numbers are greater than 0, and thus for each  $n$  there's only finitely many prime numbers smaller than  $p(n)$ . For the second case to hold, let  $n_0$  be that  $p(n+1) > p(n)$  for all  $n \geq n_0$ . By taking a tail of the sequence  $\{n : n \geq n_0\}$ , we may also assume that  $n_0 \geq 2$ . Now  $(n_0+1)^2 > n_0+1$  for  $n_0 \geq 2$ , so it also follows that  $p((n_0+1)^2) > p(n_0+1) > p(n_0)$ . However, we have  $f((n_0+1)^2) = f(n_0)f(n_0+1)$  so  $p((n_0+1)^2) = \max\{p(n_0+1), p(n_0)\}$ , contradicting our initial assumption that  $p((n_0+1)^2) > p(n_0+1) > p(n_0)$ . The conclusion therefore follows.

**N5** Fix an integer  $k > 2$ . Two players, called Ana and Banana, play the following game of numbers. Initially, some integer  $n \geq k$  gets written on the blackboard. Then they take moves in turn, with Ana beginning. A player making a move erases the number  $m$  just written on the blackboard and replaces it by some number  $m'$  with  $k \leq m' < m$  that is coprime to  $m$ . The first player who cannot move anymore loses.

An integer  $n \geq k$  is called good if Banana has a winning strategy when the initial number is  $n$ , and bad otherwise.

Consider two integers  $n, n' \geq k$  with the property that each prime number  $p \leq k$  divides  $n$  if and only if it divides  $n'$ . Prove that either both  $n$  and  $n'$  are good or both are bad.

**Solution.** Alternatively, a number is good if the second player has the winning strategy and bad if otherwise. We have the following observation:

- The number  $k$  is good.
- For each number  $n > k$ , if there's a good number  $n'$  with  $k \leq n' < n$  and  $\gcd(n', n) = 1$ , then  $n$  is bad because the player handling  $n'$  is guaranteed to lose, and therefore player handling  $n$  can replace  $n$  with  $n'$ .
- Otherwise,  $n$  is not coprime to any good number  $n'$ , so whoever handling  $n$  will either lose a move or choose the number  $n_1$  that is bad. This means their opponent who handles  $n_1$  has a winning strategy, so the player handling  $n$  themselves lose. This means  $n$  is a good number.

This also means that  $k$  is good, any two good numbers cannot be coprime to each other, and any bad number  $n$  is coprime to at least one good number  $m$  with  $k \leq m < n$ . In particular, if  $n$  is coprime to  $k$  (and so is  $n'$  by the problem condition), then  $n$  is bad (and so is  $n'$ ). So we now assume that  $\gcd(n, k)$  and  $\gcd(n, k')$  are both greater than 1.

To solve the problem, we first consider any two numbers  $n \geq k$  and  $n_0$  such that they share the same set of prime divisors that are at most  $k$ , and  $n_0$  has no divisors greater than  $k$ . We show that such  $n_0$  with  $k \leq n_0 \leq n$  exists. If  $n$  itself has no divisor greater than  $k$  we are done. Otherwise, let  $n = n_1 n_2$ , with  $n_1$  having only prime divisors at most

$p$  and  $n_2$  having only prime divisors greater than  $p$ . Since  $n_2$  has at least one prime divisor (by assumption) and greater than  $k$ , we have  $n = n_1 n_2 > k n_1$ . If  $n_1 \geq k$ , then  $n_1$  has the same set of prime divisors at most  $k$  as  $n$ , and  $k \leq n_1 \leq n_1 n_2 = n$ , so  $n_0 = n_1$  would work. Otherwise, consider the sequence  $\{n_1^m : m \geq 1\}$ , where each member in the sequence also has prime divisors satisfying our desired property. Since  $\gcd(n, k) > 1$ , we have  $n_1 > 1$  and so we can pick  $m$  that is the smallest integer such that  $n_1^m \geq k$ . Since  $n_1 < k$ , we have  $m \geq 2$ . This also means that  $n_1^{m-1} < k$  and therefore  $n_1^m < k n_1 < n$ . But we have just assumed that  $n_1^m \geq k$ , so this  $n_0 := n_1^m$  would work.

Now if  $n_0$  is good, then for each prime  $p$  we have  $p \mid n_0 \rightarrow p \mid n$  and so for any good number  $m$ , we have  $p \mid m$  and  $p \mid n_0$  simultaneously for some prime  $p$ , and therefore for this prime we have  $p \mid \gcd(m, n)$ . Considering this for all good numbers  $m$  we know that  $n$  is also good. Conversely, suppose that for some  $n$ ,  $n$  is good but  $n_0$  with the said property as above is bad. By our argument above, we may assume that  $n$  is the smallest such example. This means there's a good number  $m < n_0$  such that  $\gcd(m, n_0) = 1$  but since  $n$  is good,  $m$  and  $n_0$  have a common prime divisor, say  $p_0$ . Since the set of prime divisors of  $n_0$  and  $n$  only differs on those that are greater than  $k$ , we have  $p_0 > k$ . Now, let  $m_0$  have the same set of prime divisors  $\leq k$  as  $m$ , but has no prime divisor  $> k$ . By our previous paragraph, we may assume that  $k \leq m_0 < m$ , and by the minimality of  $n$ , since  $m$  is good,  $m_0$  must also be good. Since  $m_0$  and  $n$  are both good, there must be a prime divisor  $p_1$  dividing both numbers, and by the property of  $m_0$ ,  $p_1 \leq k$ . This means  $p_1 \mid n_0$  and  $p_1 \mid m$  too, contradicting  $\gcd(m, n_0) = 1$ .

Finally, consider any  $n$  and  $n'$  that share the same set of prime divisors  $\leq k$ . Consider  $n_0$  that has this same property without any prime divisor  $> k$  (and smaller than both  $n$  and  $n'$ ). If  $n_0$  is good then both  $n$  and  $n'$  are good; otherwise,  $n_0$  is bad so  $n$  and  $n'$  are both bad.