

Follow the Arrows

Eric Shen

Friday, April 24, 2020

Contents

1	Review material	2
1.1	Functional equations	2
1.2	What is a graph?	3
2	Lecture notes	4
3	Practice problems	8
4	Solutions to walkthroughs	10
4.1	Solution 1.1 (IMO 2017/2)	10
4.2	Solution 1.2 (Friendship paradox)	10
4.3	Solution 2.1 (AIME II 2013/11)	11
4.4	Solution 2.2 (AIME II 2018/10)	11
4.5	Solution 2.3	12
4.6	Solution 2.4 (APMO 2013/4)	12
4.7	Solution 2.5 (IMO 1987/4')	13
4.8	Solution 2.6 (ISL 2013 A5)	13
5	Solutions to practice problems	15
5.1	Solution 3.1 (JMO 2019/2')	15
5.2	Solution 3.2 (ELMO SL 2018 A1)	15
5.3	Solution 3.3 (USAMO 2019/1)	15
5.4	Solution 3.4 (USAMO 2019/1')	16
5.5	Solution 3.5 (IMO 2017/1)	17
5.6	Solution 3.6 (ISL 2017 A3)	17
5.7	Solution 3.7	18
5.8	Solution 3.8 (ISL 2012 A6)	18
5.9	Solution 3.9 (IMO 2018/3)	18
5.10	Solution 3.10 (Vietnam TST 1990/3)	19
5.11	Solution 3.11 (ISL 2010 A6)	20

§1 Review material

§1.1 Functional equations

You should probably be familiar with the notion of a functional equation. To remind ourselves how they taste like, let's run through one of the most infamously difficult functional equations in history :D.

Example 1.1 (IMO 2017/2)

Find all functions $f : \mathbb{R} \rightarrow \mathbb{R}$ such that for all real numbers x and y ,

$$f(f(x)f(y)) + f(x + y) = f(xy).$$

Walkthrough. This problem has two main parts. The first is the conventional F.E. stuff, and the second is more conceptual.

- (a) There are three linear solutions; find them. Also check that there are no more polynomial solutions.
- (b) Observe what happens when $f(0) = 0$, and settle that case. In what follows, we will assume $f(0) \neq 0$.
- (c) Note that f obeys the functional equation if and only if $-f$ obeys the functional equation, so we may actually assume $f(0) > 0$. A single solution to the F.E. remains, and it is our goal to prove it is the only solution with $f(0) > 0$.
- (d) Prove that $f(z) = 0$ for some z . It may help to scream **“DURR WE WANT THINGS TO CANCEL.”**
- (e) Prove that if for some z , we have $f(z) = 0$, then $z = 1$. (Hence $f(1) = 0$ and it is the only root.) *Hint:* let $x = z$ and make the $f(x + y)$ and $f(xy)$ terms cancel once again.
- (f) Show that $1 + f(x + 1) = f(x)$, and compute $f(n)$ for all integers n .
- (g) Prove that $f(f(x)) = 1 - f(x)$.
- (h) To utilize (g), we present the **fff trick** (sometimes called the **triple involution trick**). Express $f(f(f(x)))$ in two ways to prove $f(x) = f(1 - f(x))$.

It suffices to prove f is injective (i.e. $f(a) = f(b)$ implies $a = b$). Then it will follow that $x = 1 - f(x)$, end proof. This is the more unconventional part of the problem.

- (i) Assume $f(a) = f(b)$, and prove that $f(a + C) = f(b + C)$ for all integers C .
- (j) Prove there are x, y, C with $x + y = a + C + 1$ and $xy = b + C$. Prove that $f(f(x)f(y)) = 1$ in this case.
- (k) Show that $f(x)f(y) = 0$, and deduce $1 \in \{x, y\}$.
- (l) Conclude that $a = b$.

§1.2 What is a graph?

Graphs aren't that scary. By $G = (V, E)$, we mean G is a graph with **vertex set** V and **edge set** E . A **vertex**, or **node**, can be thought of as a point, and an **edge** connects two nodes. The **degree** of a vertex v (denoted $\deg v$) is the number of edges incident to v .

We will use this example to get to know graphs better.

Example 1.2 (Friendship paradox: On average, your friends are more popular than you)

Let G be a simple graph and v a randomly selected vertex in G . We randomly select a neighbor u of v (meaning uv is an edge). Prove that $\mathbb{E}[\deg u - \deg v] \geq 0$, and determine when equality holds.

Walkthrough. First, we prove a common result:

Lemma — Let $|E|$ be the number of edges in the graph G , and let V be the vertex set. It follows that $2|E| = \sum_{v \in V} \deg v$.

- (a) This is a double-counting exercise. Check that as we sum over $\deg v$, each edge is counted twice, and use this to prove the lemma.

Now let $N(v)$ be the set of vertices adjacent to v , so $|N(v)| = \deg v$. We want to prove

$$\mathbb{E}_{v \in G} \left[\mathbb{E}_{u \in N(v)} [\deg u] \right] \geq \mathbb{E}_{v \in G} [\deg v],$$

- (b) Replace all expected values with summations instead: we want to prove

$$\sum_{v \in G} \frac{1}{\deg v} \sum_{u \in N(v)} \deg u \geq \sum_{v \in G} \deg v.$$

- (c) Instead of summing over v , and then u on the left-hand side, sum over all edges $uv \in E$.
 (d) Minimize each individual summand via AM-GM. It should look something like

$$\frac{\deg u}{\deg v} + \frac{\deg v}{\deg u} \geq 2.$$

- (e) Prove the left-hand sum is at least $2|E|$, and establish the bound.
 (f) Identify the equality case.

§2 Lecture notes

This lecture is less theory and more “watch me do it.”

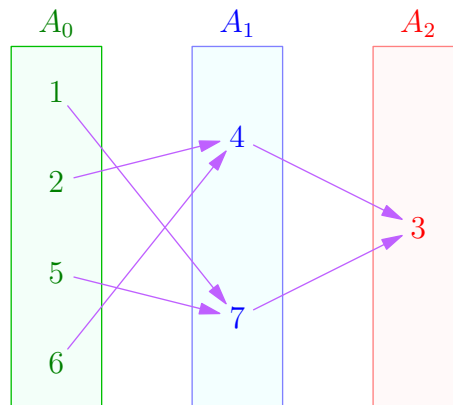
An *arrows problem* is more-or-less any problem in which drawing arrows may help solve the problem. The most common olympiad example would probably be an iterated F.E., such as Example 2.3; the flagship computational example might be one of the following two.

Example 2.1 (AIME II 2013/11)

Let $A = \{1, 2, 3, 4, 5, 6, 7\}$, and let N be the number of functions $f : A \rightarrow A$ such that $f(f(x))$ is the constant function. Find the remainder when N is divided by 1000.

Walkthrough. In most arrow function problems, we construct a graph whose vertices are the elements of A , and we draw an arrow from $a \rightarrow b$ whenever $f(a) = b$.

- (a) Draw some possible graphs, and try to partition the vertices into three different groups whose elements have similar properties.
- (b) Shown below is how I did (a). Figure out what I did.



- (c) Do casework on the number of elements in A_1 . For each $k = 1, \dots, 6$, how many functions have $|A_1| = k$?
- (d) Sum and determine the final answer.

Example 2.2 (AIME II 2018/10)

Find the number of functions $f(x) : \{1, 2, 3, 4, 5\} \rightarrow \{1, 2, 3, 4, 5\}$ that satisfy $f(f(x)) = f(f(f(x)))$ for all $x \in \{1, 2, 3, 4, 5\}$.

Walkthrough. Draw the graph as we did above. This problem can be done in a similar fashion to the 2013 AIME problem, but it’s much more nasty. Here’s an awesome way to do it with minimal casework.

- (a) In addition to drawing the arrows graph, delete all self loops, i.e. all edges of the form $f(n) = n$. What does the graph look like now?
- (b) Prove that the graph is a forest (collection of trees).
- (c) Connect all the trees together by introducing a sixth vertex r , such that for all n with $f(n) = n$, draw the edge $n \rightarrow r$. Check that the graph is now a tree with root r .

(d) Here is the statement of Cayley's formula:

For every positive integer n , the number of (undirected) trees on n labeled vertices is exactly n^{n-2} .

Use this to determine the number of directed trees with root r , such that from any vertex, following the arrows leads us to r .

(e) It is necessary and sufficient for each vertex to be a distance at most 3 from r . Subtract the overcount: in the conflicting cases, there is a "backbone" of the form $v_4 \rightarrow v_3 \rightarrow v_2 \rightarrow v_1 \rightarrow r$.

To count the number of such cases, let v_5 be the remaining vertex, and take cases on which vertex v_5 points to. (Remember that if v_5 points to v_3 , we can swap v_4, v_5 .)

(f) Finish the problem.

Example 2.3

Describe all functions $f : \mathbb{N} \rightarrow \mathbb{N}$ such that $f(f(n)) = f(n) + 1$ for all positive integers n .

Walkthrough. The answer is not as obvious as it may seem. But don't worry about that now. As we did above, draw a bunch of positive integers, and draw an arrow $a \rightarrow b$ whenever $f(a) = b$.

- (a) We use $f_*(\mathbb{N})$ to denote the range of f (formally, the image of $\mathbb{N} = \{1, 2, \dots\}$). Determine $f(m)$ for all $m \in f_*(\mathbb{N})$.
- (b) In addition to (a), give a description of what the range of f looks like. (Think: is it possible for $f_*(\mathbb{N}) = \{1, 3, 5, \dots\}$?)
- (c) Prove that for some constant c , we have $f(n) = n + 1$ for all $n \geq c$. Thus it remains to select $f(1), f(2), \dots, f(c-1)$.
- (d) Trying to prove $f(n) = n + 1$ for all n is a dead end. Check that $f(1) = f(2) = \dots = f(c-1) = c$ is actually a solution (!).
- (e) Take inspiration from (d) to come up with the complete solution set.

Example 2.4 (APMO 2013/4)

Let a and b be positive integers, and let A and B be finite disjoint sets of positive integers. Assume that for every $i \in A \cup B$, we have $i + a \in A$ or $i - b \in B$. Prove that $a|A| = b|B|$.

Walkthrough. (a) Draw arrows from i, j whenever $j = i + a \in A$ or $j = i - b \in B$.

- (b) Show that each vertex has indegree at most 1.
- (c) Each vertex has an outdegree at least 1. Combine this with (b). Consequently what does our graph look like?
- (d) The graph should consist of disjoint cycles. We know by the problem statement that the problem is true for each cycle. Furthermore if we prove the problem for each cycle, then summing over all cycles gives the general problem. Henceforth look at each individual cycle.

- (e) Start at some vertex of the cycle and follow the arrows, keeping track of how much you add and how much you subtract. Eventually you return to the starting vertex, so the amount added and the amount subtracted must be equal. Use this observation to solve the problem.

Example 2.5 (IMO 1987/4')

Determine all pairs of positive integers (a, b) for which there is a function $f : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$ satisfying

$$\underbrace{f(f(\dots f(n)\dots))}_{a \text{ times}} = n + b$$

for all nonnegative integers n .

Walkthrough. Draw the arrows.

- (a) Describe an infinite family of solutions (a, b) .
- (b) Prove, using the fff trick, that $f(n + b) = f(n) + b$ for all n .
- (c) Consider the sequence $n, f(n), f(f(n)), \dots$ for each n . Prove that the sequence is unbounded, and no two elements of the sequence are equal.
- (d) Prove that the sequence mentioned in (c) produces exactly a distinct remainders upon division by b .
- (e) Conclude that $a \mid b$.

Example 2.6 (ISL 2013 A5)

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 nonnegative integers n .

Walkthrough. This problem really consists of three parts: (i) standard FE tactics, (ii) arrow analysis, (iii) answer extraction.

We open with part (i).

- (a) Find a solution to the functional equation.
- (b) Apply the older sibling of the fff trick—the ffff trick—to prove $f^4(n) = f^4(n + 1) - 1$. Hence $f^4(n) = n + c$ for some constant c .
- (c) Prove that $f(n + c) = f(n) + c$.

Part (ii) is the pith of the problem. Per usual, we draw an arrow $a \rightarrow b$ whenever $f(a) = b$. The first step is to prove the following claim:

Claim — Each nonnegative integer is an element in some infinite chain of the form

$$w \rightarrow x \rightarrow y \rightarrow z \rightarrow w + c \rightarrow x + c \rightarrow y + c \rightarrow z + c \rightarrow w + 2c \rightarrow \dots,$$

where w, x, y, z are distinct nonnegative integers less than c .

- (d) For each n , write down the chain $n \rightarrow f(n) \rightarrow f^2(n) \rightarrow \dots$. Prove that the chain is of the form $w \rightarrow x \rightarrow y \rightarrow z \rightarrow w + c \rightarrow \dots$, and check that all its elements are distinct.
- (e) Formulate a process taking a chain \mathcal{C} of the above form with $z \geq c$ to another chain \mathcal{C}' that contains all the elements of \mathcal{C} , but whose fourth term is strictly smaller.
- (f) Conclude the claim.

For the remainder of (ii), we will prove $c = 4$.

- (g) Via Example 2.5, prove that $4 \mid c$.
- (h) For how many nonnegative integers $n < c$ is $f(n) < c$? For how many nonnegative integers $n < c$ is $f^3(n) < c$?
- (i) Prove the two following summations:

$$\sum_{n=0}^{c-1} f(n) = \frac{c(c-1)}{2} + \frac{c^2}{4} \quad \text{and} \quad \sum_{n=0}^{c-1} f^3(n) = \frac{c(c-1)}{2} + \frac{3c^2}{4}.$$

- (j) Sum the functional equation $f^3(n) = f(n+1) - 1$ over $n = 0, 1, \dots, c-1$, and use $f(c) = f(0) + c$. Conclude that $c = 4$. (Thus there is actually only one chain that traverses all of \mathbb{N} .)

Finally comes part (iii), the answer extraction. We will extensively use $f^4(n) = f(n) + 4$, $f(n+4) = f(n) + 4$, $f^3(n) = f(n+1) + 1$.

- (k) Verify $z = f(w+1) + 1$. We will use it in all four possible cases for w .
- (l) If $w = 0$, prove that $z = 3$ and $f(1) = 2$. Determine all possible chains.
- (m) If $w = 1$, exhaustively prove $z \neq 1$, $z \neq 2$, $z \neq 3$.
- (n) If $w = 2$, exhaustively prove $z \neq 2$, $z \neq 3$. Determine all possible chains in the remaining subcase $z = 1$.
- (o) If $w = 3$, prove that $z \geq 5$, which is clearly absurd.
- (p) There are two possible chains and two solutions to the functional equation. What are they?

Problem 3.10 (Vietnam TST 1990/3). Is there a function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $f(f(x)) = x^2 - 2$ for all real numbers x ?

Problem 3.11 (ISL 2010 A6). Let $f, g : \mathbb{Z}_{>0} \rightarrow \mathbb{Z}_{>0}$ be functions so that

$$f(g(n)) = f(n) + 1 \quad \text{and} \quad g(f(n)) = g(n) + 1$$

for all positive integers n . Prove that $f(n) = g(n)$ for all positive integers n .

§4 Solutions to walkthroughs

§4.1 Solution 1.1 (IMO 2017/2)

The answer is $f \equiv 0$, $f(x) \equiv x - 1$, $f(x) \equiv 1 - x$, which all clearly work. Let $P(x, y)$ denote the assertion. First note that $P(0, 0)$ gives $f(f(0)^2) = 0$. In what follows we assume $f \not\equiv 0$. We may assume that $f(0) > 0$ since:

- if $f(0) = 0$, then $P(x, 0)$ gives $f(x) = 0$ for all x ;
- if f obeys the functional equation, then so does $-f$.

The task is to show $f(x) \equiv 1 - x$.

Claim 1. $f(z) = 0 \iff z = 1$, and $f(0) = 1$, $f(1) = 0$.

Proof. For all z ,

$$P\left(z, \frac{z}{1-z}\right) \implies f\left(f(z)f\left(\frac{z}{z-1}\right)\right) = 0.$$

If $z \neq 1$ and $f(z) = 0$, then $f(0) = 0$, contradiction, thus $f(z) = 0 \implies z = 1$. But $f(f(0)^2) = 0$, so f attains 0 — thus we must have $f(1) = 0$.

Furthermore $f(0)^2 = 1$, but since $f(0) > 0$, it follows that $f(0) = 1$. \square

Claim 2. f is injective.

Proof. Note that $P(x, 1)$ gives $1 + f(x+1) = f(x)$, so if $f(a) = f(b)$, then for all nonnegative integers C , $f(a+C) = f(b+C)$. Choose C sufficiently large so that there exist x and y with

$$x + y = a + C + 1 \quad \text{and} \quad xy = b + C;$$

therefore $f(x+y) = f(xy) - 1$.

Then $P(x, y)$ gives $f(f(x)f(y)) = 1$, so $f(x)f(y) = 0$ and $1 \in \{x, y\}$. This implies $a = b$, as claimed. \square

Finally, $P(x, 0)$ gives $f(f(x)) + f(x) = 1$, so

$$f(x) = 1 - f(f(x)) = f(f(f(x))) = f(1 - f(x)).$$

From this $f(x) \equiv 1 - x$, and we are done.

§4.2 Solution 1.2 (Friendship paradox)

The proof is by taking the sum over all v :

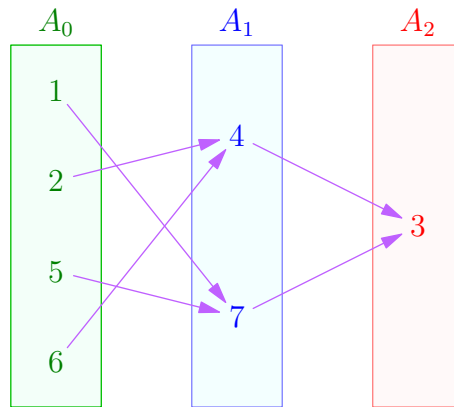
$$\sum_{v \in G} \frac{1}{\deg v} \sum_{u \in N(v)} \deg u = \sum_{uv \in E} \left(\frac{\deg u}{\deg v} + \frac{\deg v}{\deg u} \right) \geq 2|E| = \sum_{v \in G} \deg v.$$

Equality holds if and only if $\deg u = \deg v$ for all edges uv ; i.e. if G is comprised of disjoint regular graphs.

§4.3 Solution 2.1 (AIME II 2013/11)

The elements of A can be partitioned into three subsets A_0 , A_1 , A_2 with the following properties:

- $A_2 = \{c\}$, the constant value of $f \circ f$.
- A_1 is the set of x for which $f(x) = c$.
- A_0 is the set of x for which $f(f(x)) = c$ but $f(x) \neq c$.



Fix c ; there are 7 ways to do so. Let $k = |A_1|$, so $|A_0| = 6 - k$. There are $\binom{6}{k}$ ways to select A_1 , and each of its elements point to c . For the $6 - k$ elements of A_0 , each may point to any of the k elements of A_1 . Hence the answer is

$$7 \cdot \sum_{k=1}^6 \binom{6}{k} \cdot k^{6-k} = 7399.$$

§4.4 Solution 2.2 (AIME II 2018/10)

Sketch of intended solution (based on AIME II 2013/11) By considering the number of elements in $f(f(f(\{1, 2, 3, 4, 5\})))$ and casework on which map to the identity element, the answer ends up being

$$\sum_{j=1}^4 \sum_{k=1}^j \binom{j}{k} (5-j)^k k^{j-k} = 756.$$

Arrows approach Per usual, construct a graph on five vertices labeled 1, 2, 3, 4, 5. Draw an edge $a \rightarrow b$ whenever $f(a) = b$, and remove self-loops. Evidently no cycles are allowed, so this graph is a forest. Add a vertex r to the graph and draw an edge from each leaf (i.e. each n with $f(n) = n$) to r . Then we are counting the number of trees on $\{1, 2, 3, 4, 5, r\}$ with each edge a distance at most 3 from r .

Ignoring the distance condition, Cayley's formula gives $6^4 = 1296$ trees. A description of the overcounted trees is as follows: there must be a "backbone" $v_4 \rightarrow v_3 \rightarrow v_2 \rightarrow v_1 \rightarrow r$, and the sixth vertex v_5 points to any one of these. If v_5 points to r , v_1, v_2, v_4 , there are $5!$ trees each. If v_5 points to v_3 , there are $5!/2$ trees since v_4, v_5 may be swapped.

The answer is $1296 - 120 \cdot 4 - 60 = 756$.

§4.5 Solution 2.3

Let c be positive integer. All functions obeying the two following conditions clearly work:

- (i) $n < c$ implies $f(n) \geq c$;
- (ii) $n \geq c$ implies $f(n) = n + 1$.

These are the only solutions, since $n \in f_*(\mathbb{N})$ implies $f(n) = n + 1$ and $n + 1 \in f_*(\mathbb{N})$.

§4.6 Solution 2.4 (APMO 2013/4)

We present three solutions that are actually isomorphic. Nonetheless they present different ways to think of the solution.

First solution, by arrows Consider the directed graph between elements of $A \cup B$, where i points to j if $j = i + a \in A$ or $j = i - b \in B$. Hence each vertex has outdegree at least 1. I contend each vertex has indegree exactly 1. Indeed, if $i \in A$, any incoming edge has other endpoint $i - a$, and if $i \in B$, any incoming edge has other endpoint $i + b$.

It follows that each vertex has outdegree and indegree exactly 1, so the graph is comprised of disjoint cycle. Consider any one of these cycles, and say it contains p elements in A and q elements in B . I claim $ap = bq$, and summing will give the desired result.

Let x be any vertex of the cycle. There are p edges of the form $i \rightarrow i + a$ and q edges of the form $i \rightarrow i - b$. After traversing all the edges of the cycle in some order, we will end up where we started, at $x = x + ap - bq$. Hence $ap = bq$, and we are done.

Second solution, by Inclusion-Exclusion We begin with the following claim:

Claim. $A \cup B = (A - a) \cup (B + b)$, and $A - a, B + b$ are disjoint.

Proof. We know for each $i \in A$, we have $i - a \in A \cup B$, and for each $i \in B$, we have $i + b \in A \cup B$. Hence $A \cup B \subseteq (A - a) \cup (B + b)$. It then follows that

$$|A \cup B| \leq |(A - a) \cup (B + b)| \leq |A - a| + |B + b| = |A| + |B| = |A \cup B|,$$

so equality holds, and the claim follows. \square

Let $\sum S$ denote the sum of the elements of S . Summing over the elements of $A \cup B = (A - a) \cup (B + b)$, we have

$$\sum A + \sum B = \sum(A - a) + \sum(B + b) = \left(\sum A + \sum B\right) - a|A| + b|B|,$$

whence $a|A| = b|B|$, as needed.

Third solution, by generating functions Let $A(x) = \sum_{i \in A} x^i$ and $B(x) = \sum_{i \in B} x^i$ be the generating functions for A, B . Apply the claim in the above solution, so $A \cup B = (A - a) \cup (B + b)$. It follows that $A(x) + B(x) = x^{-a}A(x) + x^bB(x)$. This rewrites as

$$\frac{A(x)}{x^a} \cdot \frac{x^a - 1}{x - 1} = B(x) \cdot \frac{x^b - 1}{x - 1}.$$

Plugging in $x = 1$ yields the desired conclusion. (Alternatively differentiate at $x = 1$.)

§4.7 Solution 2.5 (IMO 1987/4')

The answer is $a \mid b$, achieved by $f(n) = n + b/a$. Clearly f is injective. Note that $f(n + b) = f(f(f(n))) = f(n) + b$, so each n determines a sequence $n, f(n), f^2(n), \dots$ that contains at most a residues modulo b . But if the two residues in the sequence are the same, then some two elements in the sequence are equal, so the sequence is eventually periodic. But the sequence is unbounded, contradiction.

Hence we can partition the residues $\{0, 1, \dots, b - 1\}$ into subsets of size a . In particular, a divides b .

§4.8 Solution 2.6 (ISL 2013 A5)

First note that

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

Hence $f^4(n) = n + c$ for some c ; as a corollary, f is injective.

By the same trick, we have $f(n + c) = f^5(n) = f(n) + c$. As a corollary, if $n \geq c$, then $f(n) \geq c$; otherwise $f(n - c) = f(n) - c < 0$, absurd. In what follows, we draw an arrow from a to b whenever $f(a) = b$.

Claim 1. Each nonnegative integer n is an element in some infinite chain of the form

$$w \rightarrow x \rightarrow y \rightarrow z \rightarrow w + c \rightarrow x + c \rightarrow y + c \rightarrow z + c \rightarrow w + 2c \rightarrow \dots,$$

where w, x, y, z are distinct nonnegative integers less than c .

Proof. Begin by writing down the chain $n \rightarrow f(n) \rightarrow f^2(n) \rightarrow \dots$. Since $f^4(n) = n + c$, the chain looks like the one presented in the claim; in particular, it is unbounded, so it is not periodic and all its elements are distinct. Now repeatedly apply the following process to force $w, x, y, z < c$:

Let the current chain be $w \rightarrow x \rightarrow y \rightarrow z \rightarrow \dots$. If $z < c$, we are done. Otherwise if $z \geq c$, instead start the chain at $z - c$, so it reads $z - c \rightarrow w \rightarrow x \rightarrow y \rightarrow \dots$.

Since the fourth term of the chain strictly decreases, the process must eventually terminate, and the claim is proven. \square

It follows that $4 \mid c$, since we can partition the elements of $\{0, 1, \dots, c - 1\}$ into subsets $\{w, x, y, z\}$ of size 4 such that $w \rightarrow x \rightarrow y \rightarrow z$.

Claim 2. $c = 4$.

Proof. By Claim 1, for $3c/4$ of the nonnegative integers $n < c$ (specifically those with $n \in \{w, x, y\}$), we have $f(n) < c$, and for the other $c/4$, we have $c \leq f(n) < 2c$. Similarly $c/4$ of the nonnegative integers $n < c$ (specifically those with $n = w$), we have $f^3(n) < c$, and for the other $3c/4$, we have $c \leq f^3(n) < 2c$.

Summing over the given functional equation,

$$\frac{c(c-1)}{2} + \frac{3c^2}{4} = \sum_{n=0}^{c-1} f^3(n) = c + \sum_{n=1}^c f(n) = 2c + \sum_{n=0}^{c-1} f(n) = \frac{c(c-1)}{2} + \frac{c^2}{4} + 2c.$$

This solves to $c = 4$. \square

What remains is a finite case check: there is a single chain $w \rightarrow x \rightarrow y \rightarrow z \rightarrow \dots$ (with $\{w, x, y, z\} = \{0, 1, 2, 3\}$) that covers all nonnegative integers. We take cases:

- If $w = 0$, then $z = f(1) + 1 \geq 3$, so $z = 3$ and $f(1) = 2$; the corresponding chain is $0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow \dots$.
- If $w = 1$, then $z = f(2) + 1 \geq 1$. Clearly $z \neq 1$; $z \neq 2$ since $f(2) \neq 1$; and $z \neq 3$ since $f(2) \neq 2$.
- If $w = 2$, then $z = f(3) + 1 \geq 1$. Clearly $z \neq 2$; $z \neq 3$ since $z \neq f(3) + 1$; and $z = 1$ implies $f(3) = 0$, and the corresponding chain is $2 \rightarrow 3 \rightarrow 0 \rightarrow 1 \rightarrow \dots$.
- If $w = 3$, then $z = f(4) + 1 \geq 5$.

In summary, the only valid chains are $0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow \dots$ and $2 \rightarrow 3 \rightarrow 0 \rightarrow 1 \rightarrow \dots$, so the answer is

$$\boxed{f(n) \equiv n + 1} \quad \text{or} \quad \boxed{f(n) \equiv \begin{cases} n + 1 & \text{if } n \text{ even} \\ n + 5 & \text{if } n \equiv 1 \pmod{4} \\ n - 3 & \text{if } n \equiv 3 \pmod{4} \end{cases}}.$$

Both work, so we are done.

§5 Solutions to practice problems

§5.1 Solution 3.1 (JMO 2019/2')

By JMO 2019/2, no such functions exist when $|a| \neq |b|$. If $a = b = 0$, any pairs (f, f^{-1}) , where f is a bijection, work.

In general, we have $f(x + b) = f(x) + a$ and $g(x + a) = g(x) + b$. When $|a| = |b| = n \neq 0$, the functions are uniquely determined by $f(0), f(1), \dots, f(n-1)$, given that they are distinct modulo n . It follows that f is invertible, and the unique function g given f is $g(x) = f^{-1}(x) + b$.

§5.2 Solution 3.2 (ELMO SL 2018 A1)

The answer is yes, infinitely many g always exist. Note that if f doesn't have finite order, then f commutes with f^0, f^1, f^2, \dots , which are all distinct. Henceforth $f^n \equiv \text{id}$ for some n .

If there are infinitely many fixed points, swap any two of them and fix everything else. In what follows, we will assume f has finite order and fixes finitely many points.

First, official solution Draw an arrow $a \rightarrow b$ whenever $f(a) = b$. This forms a graph comprised of disjoint cycles. Each cycle has length $d \mid n$, so for some d there are infinitely many cycles with length d .

Select a countably infinite number of the cycles and label them $x_{r,1} \rightarrow x_{r,2} \rightarrow \dots \rightarrow x_{r,d} \rightarrow x_{r,1}$. For each integer s , consider the function g sending $x_{r,a} \rightarrow x_{r+s,a}$ for each r, a , and fixing every other point not of the form $x_{k,\ell}$ for some k, ℓ .

There are infinitely many such g , and each commutes with f , the end.

Second solution Again draw arrows $a \rightarrow b$ whenever $f(a) = b$, so that we have a graph comprised of disjoint cycles. By assumption there are infinitely many cycles of length greater than 1.

For each cycle $x_1 \rightarrow x_2 \rightarrow \dots \rightarrow x_d \rightarrow x_1$ with $d > 1$, let $g(x_i) = f(x_i)$ for each i , and let g fix everything else. There are infinitely many ways to choose this cycle, and each such g commutes with f , end proof.

§5.3 Solution 3.3 (USAMO 2019/1)

The answer is all evens. It is not hard to check that any f that fixes the odds and is an involution on the evens works. To prove $f(1000)$ must be even, we present two solutions.

First solution, by induction In what follows, $f^k(n)$ means f iterated k times. Here we will prove all solutions to the functional equation are of the above form: f fixes odds and is an involution on the evens. Note that to prove only the original problem statement, Claims 1 and 2 suffice.

Claim 1. f is injective.

Proof. If $f(a) = f(b)$, then

$$a^2 = f^2(a)f^{f(a)}(a) = f^2(b)f^{f(b)}(b) = b^2,$$

so $a = b$ follows. □

Claim 2. If n is odd, then $f(n) = n$.

Proof. We use strong induction on n , with no base case. Assume the claim holds for all odd positive integers less than n .

Consider the equation

$$f^2(n)f^{f(n)}(n) = n^2.$$

I contend both terms on the left equal n . Let $m = f^2(n)$. If $m < n$, then $f^2(n) = f^2(m)$ by inductive hypothesis, so $n = m$ by injectivity, absurd. Thus $f^2(n) \geq n$. Analogously if $m = f^{f(n)}(n)$ and $m < n$, then $f^{f(n)}(n) = f^{f(n)}(m)$, so $n = m$ by injectivity, absurd. Thus we have $f^2(n) = f^{f(n)}(n) = n$.

Now the sequence $n, f(n), f^2(n), \dots$ repeats with period 2, so if $f(n)$ is odd, then $f(n) = f^{f(n)}(n) = n$. Otherwise suppose $f(n)$ is even, and let $m = f(n)$. Then $f(m) = n$, so $m^2 = f^2(m)f^{f(m)}(m) = n^2$, contradiction. \square

Claim 3. If n is even, then $f^2(n) = n$.

Proof. The proof is similar to that of Claim 2. We use strong induction on n , with no base case. Assume the claim holds for all even positive integers less than n ; also recall that f fixes odds and is injective, so $f(n)$ is even for all even n .

Again consider the equation

$$f^2(n)f^{f(n)}(n) = n^2.$$

It will suffice to show, once more, both terms on the left equal n . Let $m = f^2(n)$. If $m < n$, then $f^2(n) = f^2(m)$, so $n = m$ by injectivity, absurd, so $f^2(n) \geq n$. Similarly if $m = f^{f(n)}(n)$ and $m < n$, then $f^{f(n)}(n) = m = f^{f(n)}(m)$, so $n = m$ by injectivity, absurd.

Hence $f^2(n) = n$, as needed. \square

Second solution, by arrows (Espen Slettnes) Just as we did above, we first show f injective. Indeed, if $f(a) = f(b)$, then $a^2 = f^2(a)f^{f(a)}(a) = f^2(b) = f^{f(b)}(b) = b^2$, so f is injective.

Consider the sequence defined by $x_0 = 1000$ and $x_i = f(x_{i-1})$ for all $i \geq 1$. Letting $n = x_i$ in the functional equation gives

$$a_{i+2}a_{i+a_{i+1}} = a_i^2. \quad (*)$$

Take i so that a_i is minimal. Then $a_{i+2} \geq a_i$ and $a_{i+a_{i+1}} \geq a_i$, so by $(*)$ we must have $a_{i+2} = a_i$. By injectivity this reduces to $a_2 = a_0$, so $(a_i)_{i \geq 0}$ repeats with period 2.

Take $i = 0$ in $(*)$ to obtain $a_{a_1} = 1000$. Then $a_1 = 1000$ or a_1 is even. End proof.

§5.4 Solution 3.4 (USAMO 2019/1')

Just like USAMO 2019/1, the answer is f fixes odd and is an involution on the evens. It is not hard to check these work, so now we prove they are the only ones.

Claim 1. Let $k \in \mathbb{N}$ and let n be odd. Then $f(k) = n$ if and only if $k = n$.

Proof. Strong induct on n , with no base case.

Suppose the claim holds for odd integers less than n . Consider the equation $f^2(n)f^n(n) = n^2$. It is easy to verify by inductive hypothesis that $f^2(n) \geq n$ and $f^n(n) \geq n$, thus $f^2(n) = f^n(n) = n$. It follows readily that $f(n) = n$ by noting that the sequence $n, f(n), f^2(n), \dots$ repeats with period 2.

Finally assume $f(k) = n$. Then $k^2 = f^2(k)f^k(k) = n^2$, so $k = n$. \square

Claim 2. Let $k \in \mathbb{N}$ and let n be even. Then $f^2(k) = n$ if and only if $k = n$.

Proof. The proof is similar to that of Claim 1. Strong induct on n , with no base case.

Suppose the claim holds for even integers less than n . Consider the equation $f^2(n)f^n(n) = n^2$. It is easy to verify by inductive hypothesis that $f^2(n) \geq n$ and $f^n(n) \geq n$, so $f^2(n) = n$.

Finally if $f^2(k) = n$, then k is odd by Claim 1. We have $k^2 = f^2(k)f^k(k) = n^2$, so $k = n$. \square

§5.5 Solution 3.5 (IMO 2017/1)

The answer is $3 \mid a_0$. The proof is easy, but rigorizing it is substantially harder. Here is one among many ways of doing so.

Proof of sufficiency: Let m be the minimum value attained by the sequence. If $3 \mid a_0$, then all elements of the sequence are divisible by 3.

I claim $m = 3$. Indeed, if $m \geq 6$, then $(m - 3)^2 \geq m$, so the sequence reaches k^2 for some $k \leq m - 3$ and subsequently $k \leq m - 3 < m$. This contradicts the minimality of m , as required.

Then the sequence repeats $3 \rightarrow 6 \rightarrow 9 \rightarrow 3 \rightarrow \dots$, end proof.

Proof of necessity: Firstly if $a_k \equiv 2 \pmod{3}$ for some k , then $a_{k+i} = a_k + 3i$ for all i since 2 is not a quadratic residue modulo 3; thus the sequence is unbounded and the problem does not hold. Let m be the minimum value attained by the sequence.

I claim $m \equiv 2 \pmod{3}$, which proves the problem. Assume for contradiction $m \equiv 1 \pmod{3}$; then $m \geq 4$ implies $(m - 2)^2 \geq m$, so the sequence reaches k^2 for some $k \leq (m - 2)^2$ and subsequently $k \leq m - 2 < m$. This contradicts the minimality of m , as desired.

§5.6 Solution 3.6 (ISL 2017 A3)

Since S is a finite set, f is eventually periodic, say with period M . Take n large with $n \equiv -1 \pmod{M}$, so

$$f^{n+2} = f^{2n+1}.$$

I contend $f = f^n$. Assume not. Then set $g = f^n$ and note

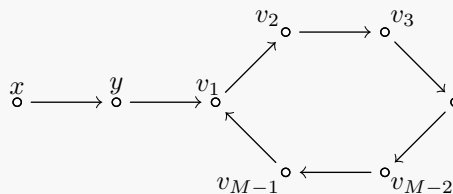
$$f \circ g \circ f = f^{n+2} = f^{2n+1} = g \circ f \circ g,$$

contradiction.

Hence f is a bijection on $f(S)$.

Remark (Motivation). The above solution is algebraic, but this problem is inherently an arrows problem. We consider the directed graph as usual, and task is to show that if $f(f(S)) \neq f(S)$, then there is a g with $f \circ g \circ f = g \circ f \circ g$.

Following the arrows eventually leads to a cycle. The idea is that $f(f(S)) \neq f(S)$ means $f(x)$ is not in a cycle for some $x \in S$; i.e. there is a tree of height ≥ 2 growing out of the cycle. For instance, see below:



Then we can pretend the vertices of the tree are actually part of the cycle as follows: for each x a distance k away from the cycle, let $g(x) = v_{M-k+1}$ (with indices modulo M).

§5.7 Solution 3.7

First we obtain subtraction and inversion. Take $a \diamond b = (a - b)^{-1}$. Set $b = 0$ to obtain inversion, and invert $a \diamond b$ to obtain subtraction.

Remark. Taking $a \diamond b = \frac{1}{a} - b$ also works.

Now we can negate by taking $0 - x$, and thus addition is possible. Halving is also obvious by $\frac{1}{x/2} = \frac{1}{x} + \frac{1}{x}$. It suffices to show that multiplication is possible, as consequently division comes from multiplying the inverse. The key is to notice that

$$\frac{1}{x-1} + \frac{1}{x+1} = \frac{2}{x^2-1},$$

whence we can square. Finally, $ab = \frac{(a+b)^2 - a^2 - b^2}{2}$, and we are done.

§5.8 Solution 3.8 (ISL 2012 A6)

Assume the contrary — that $k_n \leq M$ for each n . It is given the sequence $1, f(1), f^2(1), \dots$ is unbounded, so it contains no repeat elements; that is, $f^n(1)$ is injective in n .

Consider the two sequences $(x_n), (y_n)$ defined by $x_n = f^{2n-1}(1) - n$ and $y_n = f^{2n}(1) - n$. The given condition is equivalent to this:

For each n , there is a $k \leq M$ such that $x_n = x_{n+k}$. Similarly for each n , there is a $k \leq M$ such that $y_n = y_{n+k}$.

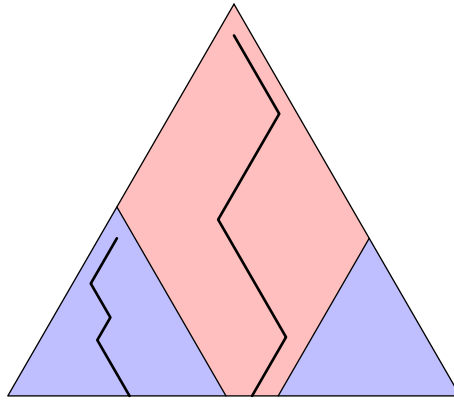
It follows that $(x_n), (y_n)$ each contain finitely many distinct values, so they are bounded above by some positive integers X, Y respectively.

Finally, consider $S = \{x_n + n : n \leq m\} \cup \{y_n + n : n \leq m\}$ for each m . Since $f^n(1)$ is injective, S contains $2m$ distinct elements. But for any $m > \max\{X, Y\}$, we have $|S| \leq \max\{X, Y\} + m < 2m$, contradiction.

§5.9 Solution 3.9 (IMO 2018/3)

The answer is no. Let $N = 2018$, and assume for the sake of contradiction such a triangle exists. For each number x not at the bottom, let its children be y and z . Draw an arrow from x to $\max(y, z)$, so that if $y > z$, we draw an arrow from x to y , and also $y = x + z$.

Let the chain starting from the top element be a_1, a_2, \dots, a_N (so that a_N is in the bottom row). Since at each step we increment by a different positive integer, it can be shown by induction that $a_i \geq 1 + 2 + \dots + i$. That is, $a_N \geq 1 + 2 + \dots + N$. Since every number in the triangle does not exceed $1 + 2 + \dots + N$, equality holds, and the numbers 1 through 2018 are all adjacent to some number in the chain. Consider the two subtriangles shown below.



We do not include $1 + 2 + \dots + N$ nor the two adjacent numbers, so neither triangle contains any integer 1 through 2018. By the Pigeonhole Principle one triangle has at least 1008 elements in its bottom row, so the number X at the bottom of the chain from the triangle's topmost element is greater than

$$\begin{aligned} X &\geq (N + 1) + (N + 2) + \dots + (N + 1008) \\ &= 1008N + 504 \cdot 1009 \\ &= 1009(N + 504) - N \\ &> 1009(N + 1) \\ &= 1 + 2 + \dots + N, \end{aligned}$$

a contradiction.

§5.10 Solution 3.10 (Vietnam TST 1990/3)

No such function exists. Say a function $g : \mathbb{R} \rightarrow \mathbb{R}$ is *good* if g has exactly two fixed points $\{a, b\}$ and $g \circ g$ has exactly four fixed points $\{a, b, c, d\}$.

Lemma (Generalization of problem)

If g is good, then there is no function f such that $f \circ f = g$.

Proof. Assume the contrary. As we did above, let the fixed points of g be $\{a, b\}$ and the fixed points of $g \circ g$ be $\{a, b, c, d\}$.

- Let $x = g(c)$; then $g(x) = g(g(c)) = c$ and $g(g(x)) = g(c) = x$, so $x \in \{c, d\}$. But if $x = c$, then $g(c) = c$, so $c \in \{a, b\}$, absurd. Hence $g(c) = d$ and $g(d) = c$.
- Next note that $f(g(x)) = f(f(f(x))) = g(f(x))$. Hence $f(a) = f(g(a)) = g(f(a))$, so $f(a) \in \{a, b\}$. By the same argument, we have $\{f(a), f(b)\} = \{a, b\}$.
- Check that $f(c) = f(g(d)) = g(f(d)) = g(f(g(c))) = g(g(f(c)))$, so $f(c) \in \{a, b, c, d\}$. Similarly $f(d) \in \{a, b, c, d\}$.

Finally we exhaust the three possible cases for $f(c)$:

- If $f(c) = a$, then $d = g(c) = f(f(c)) = f(a) \in \{a, b\}$, absurd. Similarly $f(c) \neq a$.
- If $f(c) = c$, then $f(c) = f(f(c)) = g(c) = d$, contradiction.
- If $f(c) = d$, then $c = g(d) = g(f(c)) = f(g(c)) = f(d) = f(f(c)) = g(c) = d$, absurd.

Thus the lemma is proven. \square

Finally I claim $g(x) = x^2 - 2$ is good. Indeed, $g(g(x)) = x^4 - 4x^2 + 2$, so $\{a, b\} = \{2, -1\}$ and $\{c, d\} = \{\frac{-1+\sqrt{5}}{2}, \frac{-1-\sqrt{5}}{2}\}$. We are done by the lemma.

§5.11 Solution 3.11 (ISL 2010 A6)

First note the following identities:

- For each n , we have $f(n) \neq n$; else $g(n) = g(f(n)) = g(n) + 1$. (Similarly $g(n) \neq n$.)
- $f(a) = f(b) \implies g(f(a)) = g(f(b)) \implies g(a) = g(b)$, i.e. $f(a) = f(b) \iff g(a) = g(b)$.
- $f(f(a)) = f(f(b)) \iff g(f(a)) = g(f(b)) \iff f(a) = f(b)$.

In particular, if $f(a) = f(b)$ and $a, b \in f_*(\mathbb{N})$, then $a = b$. Henceforth we will refer to this property as *pseudo-injectivity*.

The key is to describe the ranges of f, g , which we will do in the following two claims.

Claim 1. For some j, k , we have $f_*(\mathbb{N}) = \{j, j+1, \dots\}$ and $g_*(\mathbb{N}) = \{k, k+1, \dots\}$.

Proof. Note that if $f(n) \in f_*(\mathbb{N})$, then $f(n) + 1 = f(g(n)) \in f_*(\mathbb{N})$, and similarly $g(n) \in g_*(\mathbb{N})$ implies $g(n) + 1 \in g_*(\mathbb{N})$. \square

Claim 2. $f(j) = g(k)$.

Proof. Assume without loss of generality $f(j) \leq g(k)$.

Since $f(f(j)) \neq f(j)$, for some ℓ we have $f(f(j)) = f(\ell) + 1 = f(g(\ell))$. But $g(\ell) \geq g(k) \geq f(j)$, so $g(\ell) \in f_*(\mathbb{N})$. By pseudo-injectivity $f(j) = g(\ell) \geq g(k)$, thus $f(j) = g(k)$. \square

Henceforth we let $f(j) = g(k) = c$, so that $f_*(\mathbb{N}) = g_*(\mathbb{N}) = \{c, c+1, \dots\}$. In the next two claims we prove $f(n) = g(n)$ for cofinitely many n :

Claim 3. $f(c) = c + 1$.

Proof. Assume for contradiction $f(c) - 2 \geq c$. Then for some ℓ , we have

$$f(c) - 2 = f(\ell) = f(g(\ell)) - 1 = f(g(g(\ell))) - 2.$$

Note $g(\ell) = f(\ell')$ for some ℓ' . By pseudo-injectivity

$$c = g(g(\ell)) = g(f(\ell')) = g(\ell') + 1,$$

so $c - 1 \in g_*(\mathbb{N})$, which is absurd. \square

Claim 4. $f(n) = g(n) = n + 1$ for $n \geq c$.

Proof. The proof is by induction on n , where the base case $n = c$ is the above claim. If $f(n-1) = g(n-1) = n$, then

$$f(n) = f(g(n-1)) = f(n-1) + 1 = n + 1,$$

and similarly $g(n) = n + 1$, as claimed. \square

To finish, note that for each n , we have $f(f(n)) = f(n) + 1 = f(g(n))$ by the above claim. By pseudo-injectivity, $f(n) = g(n)$, and we are done.