Problem 0.1 (ELMO 2022 P6). *Find all functions f:$\mathbb{Z} \to \mathbb{Z}$ such that,for all integers m and n,*

$$f(f(m)-n)+f(f(n)-m)=f(m+n).$$

I want to write a solution which can show my thinking pace. So there are some results not necessarily be useful to the final result. But the most important thing is how we can get the answer, not the answer itself. This is more like a diary trying to keep the track of tackling this problem.

1 Part I: Warming up

We can easily show that

Observation 1.1.

$$f(n) \equiv 0, f(n) = 2n,$$

are two possible solutions.

So we may first show that $f(0) = 0$.
We construct a lemma first.

Lemma 1.2 (double lemma).

$$f(2n) = 2f(f(n)-n).$$

The proof is trivial.
Then we can show $f(0) = 0$ by the following lemma.

Lemma 1.3. *If n is divisible by 2^k,then $f(n)$ is divisible by 2^k.*

Proof. It follows by the previous lemma. □

Corollary 1.4. $f(0) = 0$.

2 Part II: Basic Calculation

When I first see this problem,the first thing I want to do is to determine how $f(n)$ is similar to $2n$. So I find the following lemma.

Lemma 2.1 (1,3 generate everything). *If $f(n) = 2n$ holds for $n, 3n$,then it holds for all $n\mathbb{Z}$, where n is an integer.*

Proof. First we notice if $f(k) = 2k$ is true for n, m, then if it is true for two of three $2n-m, 2m-n, m+n$, then it is true for the third.

So we can see $-n$ satisfies that condition if and only if n does, and if n does then $2n$ does.

For 3 we can not directly get since if we let $(m, n) = (1, 2)$, then $(2n-m, 2m-n, m+n) = (3, 0, 3)$,which involves a cycle reasoning.

But we can do for all odds larger than 3 by induction. If we let $(m, n) = (2k-1, 2)$, then $(2n-m, 2m-n, m+n) = (-(2k-5), 4(k-1), 2k+1)$, so the induction works. □

The lemma is interesting but not related to the main proof.

So I think I need to calculate $f(1)$ or try to figure out what could happen if $f(1) \neq 2$ and $f(1) \neq 0$ both hold, which is the most interesting case.

I first want to show $f(1) \neq 3$. Luckily it's really easy to show that.

Proposition 2.2. *$f(n) \neq 3n$ for $n \neq 0$.*

We see in fact we only need to show that the case $n = 1$, what we will do a lot of times later.

Proof. If $f(1) = 3$, then by double lemma, $f(2) = 2f(2)$, so $f(2) = 0$.So $f(4) = 0$.Let $(m, n) = (1, 2)$ we know

$$3 + f(-1) = f(3).$$

Let $(m, n) = (1, 0)$ we know

$$f(3) + f(-1) = 3.$$

So $f(-1) = 0$. But we can easily show that

Lemma 2.3. *If $f(n) = 0$,then for all $k \in n\mathbb{Z}$,$f(k) = 0$.*

So $f(1) = 0$,a contradiction. □

Next goal I set for myself is to show $f(1) \neq -2$, but I find it is really hard to find a direct proof.We will see why it's hard later.

3 Part III: First Difficulty – f(Whatever it is) $= 0$

I was totally defeated by the problem I set for myself to find why $f(1)$ can not be -2. For the next few hours I was trying to find the keypoint but I found nothing.

After a long time I find the keypoint is to show f(Whatever it is) $= 0$(that is what I called it).

Why can we show $f(1) \neq 3$? There must be some reason.The reason is that we show $f(2) = 0$ if $f(1) = 3$! And the method we use is the double lemma. But using double lemma we could only deal with the case $f(1) = 3$. So it introduces the idea.

If we want to use the condition,we need to find (m, n) such that $f(m) - n = m + n$, then $f(f(n) - m) = 0$.

But for $f(m) - m = 2n$,the most obvious choice is that $m = 2$(although $f(1)$ is also divisible by 2 in the end but until now we only know $f(2)$ is divisible by 2), and by double lemma $f(2) - 2 = 2(f(f(1) - 1) - 1) - 2$,so we have $n = f(f(1) - 1) - 1$, and $f(f(n) - m) = 0$.

Then we know

Lemma 3.1 (weak version)**.**

$$f(f(f(f(1) - 1) - 1) - 2) = 0.$$

or equivalent to say

Lemma 3.2 (general version)**.**

$$f(f(f(f(n) - n) - n) - 2n) = 0.$$

This is the most difficult step.

Before I show this lemma I have known some other results. I will list them now.Remember, all 1 can be substituted by n,and then k need to be substituted to kn. The proof is easy.

Lemma 3.3. *1.* $f(f(f(1)) - 1) = f(f(1) + 1)$,

2. $f(f(1) + 1) + f(f(-1) - 1) = 0$,

3. $f(f(1)) + f(-1) = f(1)$.

Obviously we want to simplfy the equation $f(f(f(f(1) - 1) - 1) - 2) = 0$. There are just too much $f's$. But all the formulas I proved can not directly fit in.However we can follow the idea of first idea.That is,to transfer f with the condition.You see $f(f(m) - n) + f(f(n) - m) = f(m + n)$, so we can move a pair of brackets from $f(m)$, and transfer them on n!Let$(m, n) = (f(1) - 1, 1)$,then

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

Now we can use the third formula so $f(f(f(1) - 1) - 1) = f(f(1)) - f(1) = -f(-1)$.So $f(-f(-1) - 2) = 0$.Or let's say,$f(-f(-n) - 2n) = 0$. So

Lemma 3.4.

$$f(2n - f(n)) = 0.$$

We can not do better now. We have not establish the tool to solve what $f(1)$ is.

4 Part IV:More Calculation

There are two cases $f(n) \equiv 0$ or $f(n) = 2n$ known now. So we want to distinguish them. The simplest choice is to calculate $f(1)$, which can not be directly solved by me.So I try to expolre more things about $f(n)$.

At first,when I have not complete the third part,I was thinking about to prove $f(n) = -f(-n)$ or $f(2n) = 2f(n)$. Because if these are true I can do a lot of work. However my roommate told me that is impossible to show them directly. But I still show that

Proposition 4.1. *1. If $f(d) = -f(-d)$ for d, same for $-d$ and $f(d)$.*

2. We can show $2f(-f(-d)) = f(2f(d))$, so if d satisfies the first condition,then $f(d)$ satisfies $f(2f(d)) = 2f(f(d))$.

But now we know $f(2n - f(n)) = 0$,so we can do further.

After that I turn to the zeros of $f(n)$. Using Lemma 2.3 we can show that

Lemma 4.2. *If $f(k) = f(l) = 0$,then $f(k + l) = 0$.*

So the zeros of $f(n)$ must have the form $n\mathbb{Z}$. Can n be a nonzero even number? No. Because using double lemma we know $f(2n) = 2f(f(n) - n)$. If $f(2n) = 0$, the$f(f(n) - n) = 0$,but $f(2n - f(n)) = 0$,so $f(n) = 0$. So we know

Lemma 4.3. *The zeros of $f(n)$ must have the form $d\mathbb{Z}$, where d is a positive odd number or* 0.

$d = 0$ and $d = 1$ correspond to the possible solutions. So let's say d an ordinary odd number. So $2n - f(n)$ need to be divisible by d for all n. Then we can easily prove $f(1) \neq -2$.

Proposition 4.4.

$$f(1) \neq -2.$$

Proof. If it is the case then because $2 - f(1) = 4$,so $d = 1$ must holds. But $f(1) \neq 0$. □

We can also see that if we let $(m, n) = (f(n) - 2n, n)$, then we get

$$\begin{aligned} f(f(n) - n) &= f(f(n) - 2n + n) \\ &= f(f(f(n) - 2n) - n) + f(f(n) - f(n) + 2n) \\ &= f(-n) + f(2n) \\ &= f(-n) + 2f(f(n) - n). \end{aligned}$$

So we can get

Lemma 4.5. *1.* $f(f(n) - n) = -f(-n)$,

2. $f(2n) = -2f(-n)$,

3. $f(n) \equiv 2n \pmod{d}$.

The second formula can be seen as a really nice substitution of linearity, which is a little different from the solution by others'. The third formula gives a useful congruence,which we will use later. Now let's show how useful they can be.

The next challenge I set for me is to show $d = 3$ is impossible. So I try and get the following:

Proposition 4.6. *Now let $d = 3$.*

1. $f(1) \neq -1$,

2. $f(1) \neq 2$,

3. $f(1) \neq 5$,

4. $f(1) \neq 8$,

5. $f(1) \neq 11$,

6. $f(1) \neq 14$.

Proof. 1. If $f(1) = -1$,we can easily show $f(-1) = -f(f(1)-1) = -f(-2) = 2f(1) = -2$ now. Then $f(2) = 4$.So $f(4) = 2f(f(2) - 2) = 2f(2) = 4$. But $f(4) = -2f(-2) = 4f(1) = -4$, a contradiction.

2. If $f(1) = 2$,then we know $f(-1) = -2$ (we early showed that). Then $f(2) = 4$. So $f(4) = 4$ also holds, but we also know $f(4) = 4f(1) = 8$, a contradiction.

3. If $f(1) = 5$, then $f(4) = 20$. But$f(-1) = -f(f(1) - 1) = -f(4) = 20$. So $f(2) = 40$. Now we will use a new formula. If we let $(m, n) = (x, rd)$ in the condition given,we then get

   Lemma 4.7 (magic formula)**.**

   $$f(x + rd) = f(f(x) - rd) + f(-x).$$

So we know in this case if we use $d = 3, x = 1$ we can get

$$f(-1) + f(5 - 3r) = f(3r + 1).$$

Let $r = 1$ we have $60 = f(-1) + f(2) = f(4) = 20$,a contradiction.

4. If $f(1) = 8$,we use $d = 3, x = 1, r = 0$ then get $f(1) = f(8) + f(-1) = -7f(-1)$,so $f(-1) = -\frac{8}{7}$, a contradiction.
5. If $f(1) = 11$,we use $d = 3, x = 1, r = 1$ then get $44 = 4f(1) = f(4) = f(8) + f(-1) = -7f(-1)$,so $f(-1) = -\frac{44}{7}$, a contradiction.
6. If $f(1) = 14$, the calculation is a little long and I don't want to bother myself to do that. Let's leave it as an exercise.

$\square$

However,all these calculation above seems to be some kind of small number luck.It works just because numbers of the form $\pm 2^k$ are quite common among small numbers, which is not for large numbers.But it gives some basic ideas.

First,$f(-2n) = -2f(n)$ gives us a really useful tool to simplify the $f(n)$ when n is even.Second,the Lemma 4.7 seems to have some magic. You see,in the condition given,there are two f's in the bracket of another f, so we can do little things. Now since$f(rd) = 0$, we only have one.

I want to use Lemma 4.7 to calculate large $f(n)$ just as before after that,but as we know if I use it one time,there will be a new f. So it's not that useful to calculate large $f(n)$ directly.I even can not prove $d \neq 3$!

As I was going to give up, I looked back to my previous idea, to identify $f(n)$ is injective or $f(n) \equiv 0$.

5 Part V:Last Step – Semi-Injectivity

I tried to prove it directly in the first few hours but my roommate said it would not work.But now it works. Why? If $f(x) = f(y)$, we can see $x \equiv y \pmod{d}$. Then if we put y in the position $x + rd$ like calculation we did before,then in the other side we get $f(y) = f(f(x) - y + x) + f(-x)$, we still can not get some useful thing. However if we do it symmetrically, then magic happens!

$$f(x + rd) = f(f(x) - rd) + f(-x)$$
$$f(y + rd) = f(f(y) - rd) + f(-y)$$

If $f(x) = f(y)$,then this annoying term $f(f(x) - rd)$ can be cancelled by subtracting the two lines.

then we get

$$f(x + rd) - f(y + rd) = f(-x) - f(-y)$$

must holds for all integers r. If we let $r = 0$,we get the right hand side is 0. So

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

If $x \neq y$,we know that $f(x + rd + c(x - y)) = f(x + rd)$ must hold for all integers c, r, so $f(x + rd)$ is bound. But it's a contradiction now.Because $f(-2n) = -2f(n)$. We can choose $(-2)^k \equiv 1 \pmod{d}$,then $f((-2)^k x) = (-2)^k f(x)$ and$(-2)^k x \equiv x \pmod{d}$.If $f(x) \neq 0$ then $x \not\equiv 0 \pmod{d}$.So $f(x + rd)$ is not bound, a contradiction.So we have

Lemma 5.1 (semi-injectivity)**.** *If $f(x) = f(y)$, then we must have either $x = y$ or $f(x) = 0$.*

Now we can throw f away.

If $d \neq 3, 1, 0$,we know if we let $(m, n) = (n, 2n)$ we get

$$f(3n) = f(f(n) - 2n) + f(f(2n) - n) = f(f(2n) - n).$$

So $f(2n) - n = 3n$, $f(2n) = 4n$.So d can not be any odd number since $f(2d) = 4d \neq 0$, a contradiction.

If $d = 3$, it is a little bit tricky.As we have seen a lot of cases,we are not afraid now.

6 Part VI:Exception – $d = 3$

3,always that 3. In Proposition 2.2 we see it, in Proposition 2.3 we love it, in Proposition 4.6 we hate it,now we will shut it.

Proposition 6.1. *The only possible solutions are $f(n) \equiv 0$ or $f(n) = 2n$.*

Proof. We just need to show $d = 3$ is impossible. By magic formula we know

$$f(1 + 3k) = f(f(1) - 3k) + f(-1).$$

But it still don't work. We know the magic formula can show its magic if we use it symetrically, so let's try another

$$f(-1 + 3l) = f(f(-1) - 3l) + f(1).$$

Take $3l = f(1) + 1 - 3k$ and add the two lines we get

$$f(1 + 3k) = f(1) + f(-1) + f(f(-1) - f(1) - 1 + 3k).$$

If $f(1) + f(-1) = 0$ then it is bound again and there is nothing to talk about.What about $f(1) + f(-1) \neq 0$? Then it is some sort of linear thing.If we take $r = f(1) + f(-1)$, $c = 2 - f(-1) + f(1)$(obviously we can not have $c = 0$), then

$$f(1 + 3d + nc) = f(1 + 3d) + nr.$$

By Lemma 4.5 we know it must have a reasonable linear speed, that is

$$c = 2 - f(-1) + f(1) = f(1) + f(-1) = r.$$

So $f(-1) = 1$,but it's impossible. Before I know $f(1) \neq 3$ I get

Proposition 6.2. *$f(n) \neq \pm n$ for $n \neq 0$.(Of course we only need to prove for $n = 1$ as before)*

I thought it was easy and has no relation to other propositions so I didn't write it into this article first. Suprisingly it occurs here!

Now we can prove it in a quite primitive way but let's just leave it as an exercise. Now we just use Proposition 4.6. So we complete the proof.

Congratulations! □