

# Monsters

EVAN CHEN 《陳誼廷》

2 October 2016

Whoever fights monsters should see to it that in the process he does not become a monster. And if you gaze long enough into an abyss, the abyss will gaze back into you.

— Friedrich Nietzsche

## Acknowledgments

THANKS to Ashwin Sah for detailed suggestions on an early draft, and to Kaan Dokmeci, Mihir Singhal, David Stoner, Colin Tang, Franklyn Wang, Jacob Zhang, and Kairat Zhubayev for several corrections.

## §1 Warm-Up

**Problem 1.1** (Tournament of Towns 1983). Suppose  $f: \mathbb{Z}^+ \rightarrow \mathbb{Z}^+$  is strictly increasing and satisfies  $f(f(n)) = 3n$  for every positive integer  $n$ . Evaluate  $f(1983)$ .

Solution at <http://www.aops.com/community/c6h390679p2170777>.

## §2 Introduction

Your typical garden-variety functional equation will ask “find all functions  $f$ ”, and there will be an obvious function like  $f(x) = x$  which works. Most of the time your job will be to prove these are the only solutions.

Sometimes, though, the functional equation will have a nasty surprise: the obvious solutions aren’t the only ones! The classic example is the relatively innocent-looking

$$f: \mathbb{R} \rightarrow \mathbb{R} \quad f(x + y) = f(x) + f(y)$$

which is called **Cauchy’s Functional Equation**. It is easy enough to get that  $f(x) = x \cdot f(1)$  for  $x \in \mathbb{Q}$ , yet there exist plenty more pathological solutions: we will discuss this example in depth later.

Monsters are most dangerous when you don’t know they are there. If you stubbornly try to prove that  $f(x) = x$  is the only solution when it isn’t, you are destined to fail. On the flip side, if you correctly guess the existence of a pathological solution, this gives you a huge upper hand!

## §3 Clues

Here are some clues that you might be dealing with a functional equation with some bizarre solutions.

- **Some stubborn case appears that can't be resolved.** For example, suppose you obtain that  $f(0) \in \{0, 1\}$ , and you try without success to dispel the  $f(0) = 1$  case. Might it be possible there is actually a solution? Check to see if  $f(x) = 1 - x$  might be a solution too. What if you have  $f(x)^2 = x^2$  for all  $x$ , but you can't get the sign? Might the function change sign at some values of  $x$ ?
- **Values of the function seem “too discrete”.** For example, you have  $f: \mathbb{Z} \rightarrow \mathbb{Z}$  and kind of find a way to relate  $f(n+1)$  to  $f(n)$ , but there is still some degree of freedom left. An extreme example of this is [Example 6.1](#).
- **You have some values of  $f$  down, but others seem out of reach.** This usually happens when  $f: \mathbb{R} \rightarrow \mathbb{R}$ . Cauchy's Functional Equation is the classical example of this: you can get  $f(x)$  at rational values, but how on Earth are you going to get  $f(\sqrt{2})$ ?
- **All values are “wrapped by  $f$ 's” in the equation.** In such cases you should immediately check for constant solutions. But it's also possible for the function to have a small range in a bizarre way. The legendary example of this is Shortlist 2004 A6, which is [Problem 7.13](#).
- **The problem only uses one operation.** The real numbers have *two* operations,  $+$  and  $\times$ . So it loses some of its structure if, say, there is no multiplication. This is the real reason that Cauchy's Functional Equation has bad solutions: it *ignores* the multiplication structure of  $\mathbb{R}$  and only looks at the additive structure.

Similarly, you can generally classify monsters into a few types.

- The mildest type of extra solution is when you get an extra function like  $f(x) = 1 - 2x$  when you originally were only expecting  $f(x) = x$ .  
It may be worth it, at the start of any problem, to **begin by plugging in polynomials, especially  $f(x) = ax + b$** , and seeing if there are any non-obvious solutions. This will help you notice solutions like  $x + c$  or  $cx$  or  $0$ , and it will give you a huge advantage should there be an unexpected solution.
- Some monsters take the form of functions  $f: \mathbb{Z} \rightarrow \mathbb{Z}$  that behave in cases based on the inputs mod  $n$ . You often pick up such functions when you try to compute  $f$  inductively, but find that the proof just won't go through due to some degrees of freedom.
- Still other monsters might take values or signs at certain inputs and other values at different inputs. For example, you might have a case  $f(x)^2 = 1$  for all  $x$ , and be unable to progress past that. This is more likely to happen in the “limited range” case specified earlier.
- Finally, you might be dealing with a functional equation  $f: \mathbb{R} \rightarrow \mathbb{R}$  which requires a Hamel basis (explained below). This probably won't happen in a real olympiad any time soon. . .but it is good to at least be aware of it.

## §4 Linear Algebra Terminology

Before I proceed, I want to introduce some linear algebra terms, so that I can explain things the “morally correct” way, rather than having to use clumsy terminology. If you know linear algebra well, you should skip this section.

Let  $K$  be a field (for our purposes, either  $\mathbb{Q}$  or  $\mathbb{R}$ ).

**Definition 4.1.** Informally, a  **$K$ -vector space** is a set  $V$  such that

- One can add any two elements of  $V$ , and
- One can scale elements of  $V$  by *scalars* in  $K$ .

There are some additional axioms<sup>1</sup> (addition should be commutative,  $0 \in V$ , and  $-v \in V$  for all  $v \in V$ ) but we won't concern ourselves with these.

The two best examples I have:

- The set of real polynomial of degree  $\leq 2$ , that is,

$$W = \{ax^2 + bx + c \mid a, b, c \in \mathbb{R}\},$$

is a real vector space. You can *add* any two such polynomials, and you can multiply them by real numbers. (Here, possibly  $a = 0$ ; what goes wrong if I try to force  $a \neq 0$ ?).

- The set of real polynomials is a real vector space, full stop. The sum of two polynomials is a polynomial, and if  $P$  is a polynomial then so is  $c \cdot P$ .

Stranger example:  $\mathbb{R}$  is a  $\mathbb{Q}$ -vector space. This will be important later.

Now, let's return to the example  $W = \{ax^2 + bx + c \mid a, b, c \in \mathbb{R}\}$ . You'll instantly recognize that the set  $\{1, x, x^2\}$  plays some special role: these elements generate all of  $W$  in some clean fashion.

To make this formal:

**Definition 4.2.** A set  $\mathcal{B}$  of vectors is a **basis** for a vector space  $V$  if every vector  $v \in V$  can be written *uniquely* as a finite sum of the form

$$v = t_1e_1 + t_2e_2 + \cdots + t_me_m \tag{1}$$

where  $t_i \in K$ ,  $e_i \in \mathcal{B}$ .

So,  $\{1, x, x^2\}$  is a basis of  $W$ . It's not the only one:  $\{2, x, x^2\}$  and  $\{x+4, x-2, x^2+x\}$  are other examples of bases, though not as natural. However, the set  $S = \{3+x^2, x+1, 5+2x+x^2\}$  is not a basis: it fails for the following two reasons.

- Note that

$$0 = (3+x^2) + 2(x+1) - (5+2x+x^2).$$

This violates our uniqueness condition, since  $0 = 0$ . In this way, we say the elements of  $S$  are not **linearly independent**.

- It's not possible to write  $x^2$  as a sum of elements of  $S$ . (Try it and see why not.) So  $S$  fails to be **spanning**.

With these new terms, we can just say a basis is a linearly independent, spanning set. As you might guess, you always need exactly three elements for  $W$ . More generally:

<sup>1</sup>Reminder for experts:  $(K, +)$  is an abelian group with compatible multiplication by scalars in  $K$ .

**Theorem 4.3** (Dimension Theorem)

Let  $V$  be a vector space which has a basis of size  $n$ .

- (a) Any other basis of  $V$  has size  $n$ , so we say  $V$  is  **$n$ -dimensional**.
- (b) Given  $n$  linearly independent elements, they form a basis.
- (c) Given  $n$  spanning elements, they form a basis.

It's also possible to have an infinite basis. For example, consider the set of polynomials. It has a basis  $\{1, x, x^2, \dots\}$  in the sense that any polynomial is just a *finite* sum  $\sum c_k x^k$ . (Note that (1) only permits finite sums!)

**Remark 4.4.** Possible spoiler: the Axiom of Choice is actually equivalent to the fact that every vector space has a (possibly infinite) basis.

## §5 Cauchy's Functional Equation

### §5.1 The Hamel Basis

**Example 5.1** (Cauchy's Functional Equation)

Find all functions  $f: \mathbb{R} \rightarrow \mathbb{R}$  satisfying  $f(x + y) = f(x) + f(y)$ .

Let's do this example in closer detail. Of course, our first naïve guess is that the solution set is  $f(x) = cx$  for some real number  $c$ . So, we let  $f(1) = c$  (as we may, you can think of this as "scaling"). Then

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

Next

$$f(3) = f(2) + f(1) = 3c$$

and readily we discover  $f(n) = nc$ , which is right on track. We can extend this to get all rational numbers, as for any integers  $p, q$  we see that

$$f\left(\frac{p}{q}\right) = f\left(\frac{p}{q} + \dots + \frac{p}{q}\right) = qf\left(\frac{p}{q}\right)$$

so  $f(p/q) = f(p)/q = c \cdot (p/q)$ , which is still spot on.

However, if you try solving the rest of the problem from here you will quickly get stuck. We've pinned down the value of  $f$  for all rational numbers, but how would we get  $f(\sqrt{2})$ , for example? Try all you want, but it won't work.

Here's why. What if, rather than talking about  $f: \mathbb{R} \rightarrow \mathbb{R}$ , I asked for solutions of  $f: \mathbb{Q}[\sqrt{2}] \rightarrow \mathbb{R}$ ? You might at this point be able to guess a solution:

$$f(a + b\sqrt{2}) = a + 2015b.$$

Convince yourself that this is well-defined and works – the point is that  $\sqrt{2}$  and 1 don't talk to each other. Analogously, the function

$$f(a + b\sqrt{2} + c\sqrt{3}) = a + 2015b + \sqrt{11}c$$

is a perfectly good solution for the inputs on which it's defined.

At this point you might guess how to construct a monster: keep throwing in “foreign” (linearly independent) elements in this way until you get all of  $\mathbb{R}$ . The bad news is that you will need not only infinitely many elements, but in fact *uncountably* many elements, so this can’t be done with normal induction. The good news is that this has all been worked out for you, and this is where the Axiom of Choice gets used in a so-called *transfinite induction*.

If you’re interested in the details, check out <https://blog.evanchen.cc/2015/04/10/cauchys-functional-equation-and-zorns-lemma/>. For this handout I will just state the result.

### Proposition 5.2 (Hamel Basis)

$\mathbb{R}$  has a basis as a  $\mathbb{Q}$ -vector space. Thus, there exists an infinite collection of real numbers  $\{e_\alpha\}$  such that for any real number  $x \in \mathbb{R}$ , there is a **unique** way to write  $x$  as a **finite** linear combination

$$x = a_1 e_{\alpha_1} + a_2 e_{\alpha_2} + \cdots + a_n e_{\alpha_n}.$$

The numbers  $\{e_\alpha\}$  are called a **Hamel basis**.

Concretely, you can almost think of this as saying something like

$$\mathbb{R} = \left\{ a_1 + \sqrt{2}a_2 + \sqrt{3}a_3 + a_4\pi + a_5e + \cdots \mid a_1, \dots \in \mathbb{Q} \right\}.$$

One literally just keeps throwing in elements until we get all of  $\mathbb{R}$ , and the Axiom of Choice is used to make this rigorous.

In any case, this resolves the original Cauchy’s Functional Equation. We simply take a Hamel basis, and assign  $f(e_\alpha)$  arbitrarily for each  $\alpha$ . Then, declare

$$f\left(\sum a_\alpha e_\alpha\right) = \sum a_\alpha f(e_\alpha).$$

Those of you familiar with linear algebra may recognize this as the following assertion: to specify a linear map, it suffices to specify it on the basis elements.

**Exercise 5.3 (Combinatorics Practice).** Show that any Hamel basis has *uncountably infinitely* many elements. (This is why I insist on calling it  $\{e_\alpha\}$  rather than  $e_1, e_2, \dots$ .)

If you know linear algebra well, then you can summarize the entire section as follows: view  $\mathbb{R}$  as a  $\mathbb{Q}$ -vector space. The Axiom of Choice lets you take a basis, which trivializes Cauchy’s Functional Equation.

## §5.2 Repairing Cauchy’s Functional Equation

Nonetheless, the “pathological” solutions to Cauchy’s Functional Equation exhibit some non-standard behaviors. Roughly, they ignore the “order” structure of  $\mathbb{R}$ .

### Theorem 5.4 (Properties of the Nonlinear Solutions)

Suppose  $f: \mathbb{R} \rightarrow \mathbb{R}$  is additive but not linear. Consider the *graph* of  $f$ , i.e. the set of points

$$G = \{(x, y) \in \mathbb{R}^2 \mid y = f(x)\}.$$

Then  $G$  is dense in the plane: any disk in  $\mathbb{R}^2$  contains points of  $G$ .

Hence, if we tried to print a plot of the function, the entire page would be covered in ink!

*Sketch of Proof.* Suppose  $f$  is an additive nonlinear function; WLOG  $f(1) = 0$  (why?). I'll show one can find points of  $f$  close to any given point  $(X, Y)$ .

Since  $f$  isn't linear, there's some  $\alpha$  such that  $f(\alpha) \neq 0$ , say  $f(\alpha) = z \neq 0$ . Thus, for any other rational numbers  $q$  and  $r$ , we have

$$f(q\alpha + r) = qz.$$

So, pick  $q$  such that  $qz$  is close to  $Y$ ; then pick  $r$  such that  $q\alpha + r$  is close to  $X$ . Since  $\mathbb{Q}$  is "dense in  $\mathbb{R}$ " we can make this as accurate as we want.  $\square$

Looking at the non-empty half of the glass of water, we can mold the above pathology into the following corollary.

#### Corollary 5.5 (Continuous/Bounded Cauchy)

Suppose  $f: \mathbb{R} \rightarrow \mathbb{R}$  satisfies Cauchy's Functional Equation, and **any** of the following conditions holds:

- $f$  is continuous
- $f$  is monotone
- $f \geq 0$  for  $x \geq 0$  (or any nontrivial interval).

Then  $f$  is a linear function.

In general, if an olympiad problem tells you a function is continuous (or something even stronger like "smooth"), this is a strong hint that you are trying to prove the function is additive. Another possible hint is if the domain is  $\mathbb{Q}$  (but more often this means induction-like solutions that don't carry over to  $\mathbb{R}$ ).

### §5.3 Involutions for Cauchy's Functional Equation

Here is an example of a condition that's *not* enough to get rid of pathology.

#### Example 5.6 (Involutions of Cauchy's FE)

Find all functions  $f: \mathbb{R} \rightarrow \mathbb{R}$  satisfying

$$f(x + y) = f(x) + f(y) \quad \text{and} \quad f(f(x)) = x.$$

This might be surprising, since  $f(f(x)) = x$  (such functions are called **involutions**) is a fairly strong condition, for example:

**Exercise 5.7.** Check that an involution must be both injective and surjective.

Nonetheless, this does have pathological solutions. To understand why, we have to think about vector spaces more deeply.

Our hope would be that the only solutions are  $f(x) = \pm x$ . Unfortunately, as we saw in the last problem, you can decompose  $\mathbb{R}$  in such a way that "elements in different parts don't talk to each other". We can do the same thing here. Imagine we have a Hamel basis  $\{e_\alpha\}$ . What we can do is, for every  $\alpha$ , pick a sign  $\varepsilon_\alpha \in \{\pm 1\}$ . Then, we have a solution given by sending each  $e_\alpha$  to  $\varepsilon_\alpha e_\alpha$ , and then extending linearly.

There's another way to think of this in terms of decompositions.

**Definition 5.8.** Let  $V$  be a  $K$ -vector space, and suppose  $A, B \subset V$  are subspaces (meaning they are themselves  $K$ -vector spaces). We write  $V = A \oplus B$  for  $A, B \subset V$  if the following holds: every element of  $V$  is *uniquely* a sum of the form  $a + b$ , where  $a \in A$ ,  $b \in B$ . Equivalently,  $A \cap B = \{0\}$  and every element of  $V$  is the sum of two elements in  $A$  and  $B$ . We say  $V$  is a **direct sum** of  $A$  and  $B$ .

For example, if  $V$  is the set of real polynomials of degree  $\leq 2$ , then if we let

$$A = \{kx^2 \mid k \in \mathbb{R}\} \quad \text{and} \quad B = \{kx + \ell \mid k, \ell \in \mathbb{R}\}$$

then we have  $V = A \oplus B$ .

In that case, we can generate solutions to [Example 5.6](#) in the following form: given any decomposition  $\mathbb{R} = A \oplus B$ , we take any given  $x = a + b$ , and send it to  $f(x) = a - b$ . Thus,  $f(f(x)) = f(a + (-b)) = a + b$  back. In fact, all solutions look like this!

*Solution to Example 5.6.* To be clear, we claim the solutions are exactly those of the following form: for any decomposition  $\mathbb{R} = A \oplus B$ , we let  $f(a + b) = a - b$ . Since such solutions clearly work, we need to prove all solutions can be written in this way.

Given such a function  $f$ , let  $A$  be the set of  $x$  such that  $f(x) = +x$  and  $B$  be the set of  $x$  such that  $f(x) = -x$ .

**Exercise 5.9.** Show that  $A$  and  $B$  are vector spaces with  $A \cap B = \{0\}$ .

We prove that every number  $x \in \mathbb{R}$  can be written as the sum of an element of  $A$  and the sum of an element from  $B$  now. Here's why:

$$\begin{aligned} f\left(\frac{x + f(x)}{2}\right) &= \frac{f(x) + f(f(x))}{2} = \frac{f(x) + x}{2} \\ f\left(\frac{x - f(x)}{2}\right) &= \frac{f(x) - f(f(x))}{2} = \frac{f(x) - x}{2}. \end{aligned}$$

So  $\frac{1}{2}(x + f(x)) \in A$  and  $\frac{1}{2}(x - f(x)) \in B$ , but these two elements sum to  $x$ .

So since  $A + B$  generates  $\mathbb{R}$  but  $A \cap B = \{0\}$ , we deduce that  $A$  and  $B$  are  $\mathbb{Q}$ -supplementary subspaces of  $\mathbb{R}$  — in other words, for every  $x \in \mathbb{R}$ , we can *uniquely* write  $x = a + b$ . This completes the proof.  $\square$

If this was confusing, it may help to redo this problem over  $\mathbb{Q}[\sqrt{2}] \rightarrow \mathbb{Q}[\sqrt{2}]$ . An example of a solution in this simpler situation is

$$f(a + b\sqrt{2}) = a - b\sqrt{2}.$$

or even

$$f(a + b(\sqrt{2} + 1)) = a - b(\sqrt{2} + 1).$$

It's harder to visualize for  $\mathbb{R}$  because we have no idea what the Hamel basis looks like.

**Remark 5.10** (For Linear Algebra Experts). This boils down to the following: we're trying to find all  $\mathbb{Q}$ -linear endomorphisms  $f: \mathbb{R} \rightarrow \mathbb{R}$  such that  $f^2 = \text{id}$ . So we see that  $A$  is the set of  $+1$  eigenvectors of  $f$  and  $B$  is the set of  $-1$  eigenvectors of  $f$ . These are the only possible eigenvalues: if  $f(v) = \lambda v$  we have  $\lambda^2 v = v$ , thus  $\lambda = \pm 1$ .

## §6 Example: An Integer Functional Equation

We next give an example of a functional equation which has behavior modulo  $n$ . It stems from the fact that one can determine its values from a few inputs, but that there is still some freedom left over.

### Example 6.1 (Gabriel Dospinescu, 2005)

Find all  $f: \mathbb{Z} \rightarrow \mathbb{R}$  satisfying

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

This solution is involved, and I split it into many pages. In my opinion it is not a realistic example of a problem that will show up on a contest. But it illustrates the kind of technical difficulties you run into when you have a function that is “kind of” determined by induction, and how they can lead to the existence of pathological solutions.

**Exercise 6.2.** Before reading the solution to follow, try and find as many functions satisfying the condition as you can.

### §6.1 Shifting

First, we plug in  $f(x) = ax + b$  to see which linear functions work. Surprisingly, all of them do! While this might be discouraging, it gives us a starting point, by shifting  $f$  appropriately, we may assume that  $f(0) = 0$ , and then we may assume that  $f(1) = 1$ . To be more verbose, if  $\tilde{f}$  is a solution to the functional equation then so is

$$f(x) = \tilde{f}(x) - c_1x - c_2$$

and we can select  $c_1, c_2$  such that  $f(0) = 0, f(1) = 1$ . So there is no harm in making these assumptions for now.

It’s worth noting at this point that this functional equation is *highly linear*. Specifically, if  $f$  is a solution, so is  $c \cdot f$ ; similarly, if  $g$  is a solution, then so is  $f + g$ . In linear algebra terms, the set of solutions is a *real vector space*. This is worth noting, because if we find even more solutions they can all add together to generate further solutions.

### §6.2 Reduction to a Finite System

Plugging in  $x = y = 0$  is the natural first move, but here it yields nothing. In fact, even plugging in  $y = 0$  yields nothing (and so there is no point in trying  $x = 0$  either, since the equation is symmetric). So, let’s instead try plugging in  $y = 1$  to obtain

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

Incredibly, still nothing.

How about  $y = 2$ ? We now have that

$$f(2x+1) + f(2) + f(x) = f(2x) + 1 + f(x+2) \tag{2}$$

In a similar line of thought, if we use 4 instead of 2 we have

$$f(4x+1) + f(4) + f(x) = f(4x) + 1 + f(x+4) \tag{3}$$

We now want to get terms of (2) and (3) to cancel. So we replace  $x$  with  $2x$  in (2) to get

$$f(4x+1) + f(2) + f(2x) = f(4x) + 1 + f(2x+2) \tag{4}$$



Thus, if we subtract (3) and (4), we get massive cancellation; when the dust settles we are left with

$$f(x) + f(4) + f(2x + 2) = f(x + 4) + f(2x) + f(2).$$

In short, we've found that

$$\begin{aligned} f(2x + 1) &= f(2x) + f(x + 2) + 1 - f(x) - f(2) \\ f(2x + 2) &= f(2x) + f(x + 4) + f(2) - f(4) - f(x). \end{aligned}$$

This is good, because it lets us construct  $f$  inductively given enough base cases, at least for positive inputs. For example, if we know  $f(0), f(1), \dots, f(99)$  then we can compute  $f(100) = f(98) + f(53) + 1 - f(4) - f(49)$ , and so on. In fact:

**Exercise 6.3.** Check that if we know  $f(n)$  for  $n = 0, 1, 2, 3, 4, 6$  then we know  $f(n)$  for all  $n \geq 0$ .

Let's now see if we can pin down some information about the negative values of  $x$ . The natural thing to do is put  $y = -1$ , which tells us

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

Thus, we see that if we know  $f(-1)$ , we can also extract the values of all negative integers by induction. Happily enough, we already know (or rather, have assumed) that  $f(0) = 0$  and  $f(1) = 1$ , and we have narrowed our sights down to seven values of  $n$  we'd like to compute.

In fact, we can do even better with the negative integers. If we plug in  $(-x, -y)$  into the given, we obtain that

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

This looks quite like the given, and motivates the following idea: let  $h(x) = f(x) - f(-x)$ , then

$$h(x + y) = h(x) + h(y)$$

by subtracting the given from the previously derived equation. Thus,  $h(x) = x \cdot h(1)$  by induction, and now we obtain

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

### §6.3 Pinning Down

The next part is somewhat bloody, so I'll spare the details and just say what things to plug in.

**Exercise 6.4.** By plugging in  $(x, y) = (-2, 2)$  and  $(-1, 4)$ , show that

$$f(-1) = f(0) - 2f(2) + 2f(1) - f(3) + f(4).$$

**Exercise 6.5.** By considering  $f(-x)$  for  $x \in \{1, 2, 5, 6\}$  and also  $(x, y) = (2, 2)$  and  $(x, y) = (-2, 3)$ , show that

$$f(6) = f(0) - 3f(2) + 3f(4).$$

These are not *too* unrealistic to come up with yourself. The idea is that we've pinned the problem down to a small, finite set of values. So we keep plugging in "local" values: small values of  $x$  and  $y$  and see what we can make out of it. Given enough time, you'll manage to get rid of  $f(-1)$  and  $f(6)$ , but not so with the other values of  $x$ .

(Remark: I think if you change the problem to  $\mathbb{Z}_{\geq 0} \rightarrow \mathbb{R}$ , then you can't get rid of  $f(6)$  either.)

## §6.4 The Solution Set

So, it remains to compute  $f(0)$ ,  $f(1)$ ,  $f(2)$ ,  $f(3)$ ,  $f(4)$ , and as we've seen already we may safely assume that  $f(0) = 0$ ,  $f(1) = 1$ . Unfortunately, try as you might, you'll find that it's impossible to determine any of  $f(2)$ ,  $f(3)$ ,  $f(4)$  from the  $f(0) = 0$  or  $f(1) = 1$  alone.

Eventually, you get tired of this, and might ask the flip direction: “what *does* go wrong if  $f(4) \neq 4$ , say?”. So, let's put  $f(2) = 2$ ,  $f(3) = 3$ ,  $f(4) = 5$  and see where it takes us. By pumping the recursion and the formula for  $f(6)$ , we get the first few values of  $f$  “should” be:

$f$	0	1	2	3	4	5	6	7	8	9	10	...
	0	1	2	3	5	7	9	12	15	18	22	...

and so on. Now, we look for the pair  $(x, y)$  which causes the given to be false.

And then the surprise comes: there is none. In fact, the function defined by the terrible recursion is an honest-to-god solution to the original functional equation. And in fact, it turns out you can repeat the experiment with *any* values of  $f(2)$ ,  $f(3)$ ,  $f(4)$ , and it will yield a valid solution.

What's going on? Well, depending on how many solutions you found in [Exercise 6.2](#), you might be more or less surprised than others. In particular, if you found the following five solutions then this is no surprise at all:

$$\begin{aligned} f(x) &= 1, x, x^2 \\ f(x) &= x \pmod{2} \in \{0, 1\} \\ f(x) &= x^2 \pmod{3} \in \{0, 1\}. \end{aligned}$$

Letting  $\pi_2(x) = x \pmod{2}$  and  $\pi_3(x) = x^2 \pmod{3}$ , we see that every function of the form

$$f(x) = a + bx + cx^2 + d\pi_2(x) + e\pi_3(x)$$

is a solution to the functional equation, for arbitrary reals  $a, b, c, d, e$ . Reason: the five functions exhibited above are linearly independent, and since the dimension of the solution space is  $\leq 5$  they must form a basis for the entire solution space. (In elementary terms, you can just solve for  $a, b, c, d, e$  given  $f(0), f(1), f(2), f(3), f(4)$ .)

So, the set of solutions is the real vector space of dimension 5 given by the basis elements  $1, x, x^2, x \pmod{2}$  and  $x^2 \pmod{3}$ .

## §7 Problems

Just to keep you on your toes: two of the problems here are completely benign, with only the “obvious” solutions. Note also that the last few problems are quite difficult (possibly too hard to use on a contest), and so even completely guessing the solution set is a fine achievement.

Thanks to Art of Problem Solving user **pco** for pointing me to many of these. Personal recommendations: [7.4](#), [7.5](#), [7.10](#), [7.12](#). By “solve over  $K$ ” I mean “find all functions  $f: K \rightarrow K$  such that the equation holds for all inputs in  $K$ ”.

**Problem 7.1** (Albania TST 2008). Solve  $f(x + f(y)) = y + f(x + 1)$  over  $\mathbb{R}$ .

**Problem 7.2** (IMO 2012/4). Find all functions  $f: \mathbb{Z} \rightarrow \mathbb{Z}$  such that, for all integers  $a, b, c$  that satisfy  $a + b + c = 0$ , the following equality holds:

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

**Problem 7.3** (ELMO 2014). Find all triples  $(f, g, h)$  of injective functions from the set of real numbers to itself satisfying

$$\begin{aligned}f(x + f(y)) &= g(x) + h(y) \\g(x + g(y)) &= h(x) + f(y) \\h(x + h(y)) &= f(x) + g(y)\end{aligned}$$

for all real numbers  $x$  and  $y$ .

**Problem 7.4** (Shortlist 2001 A4). Find all functions  $f: \mathbb{R} \rightarrow \mathbb{R}$  satisfying

$$f(xy)(f(x) - f(y)) = (x - y)f(x)f(y)$$

for all real numbers  $x$  and  $y$ .

**Problem 7.5** (ELMO Shortlist 2013). Let  $f^k(n)$  denote the function  $f$  applied  $k$  times. Solve over positive integers the functional equation

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

**Problem 7.6** (EGMO 2012/6). Determine all functions  $f: \mathbb{R} \rightarrow \mathbb{R}$  satisfying the condition

$$f(y^2 + 2xf(y) + f(x)^2) = (y + f(x))(x + f(y))$$

for all real numbers  $x$  and  $y$ .

**Problem 7.7** (January TST for IMO 2015, “Cauchy’s FE modulo 1”). Let  $f: \mathbb{Q} \rightarrow \mathbb{Q}$  be a function such that for any  $x, y \in \mathbb{Q}$ , the number  $f(x + y) - f(x) - f(y)$  is an integer. Decide whether there must exist a constant  $c$  such that  $f(x) - cx$  is an integer for every rational number  $x$ .

**Problem 7.8** (TSTST 2011/1). Find all  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  with the property that for any  $a, b, c$ , the median of  $a, b, c$  is the median of  $f(a, b), f(b, c), f(c, a)$ .

**Problem 7.9** (IMO 2015/5). Solve the functional equation  $f(x + f(x + y)) + f(xy) = x + f(x + y) + yf(x)$  over the real numbers.

**Problem 7.10** (Kazakhstan 2005). Solve over  $\mathbb{R}$ :

$$f(f(x) + x + y) = 2x + f(y).$$

**Problem 7.11** (Shortlist 2013 A5). Solve  $f(f(f(n))) = f(n+1) + 1$  over the nonnegative integers.

**Problem 7.12** (ELMO Shortlist 2012). Solve over  $\mathbb{Q}$ :

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

**Problem 7.13** (Shortlist 2004 A6). Find all functions  $f: \mathbb{R} \rightarrow \mathbb{R}$  which obey

$$f(x^2 + y^2 + 2f(xy)) = (f(x + y))^2$$

for all real numbers  $x$  and  $y$ .

Hints are on the next page. Most of them merely tell you the set of solutions.

## §8 Hints

**7.1** Substituting  $f(x) = g(x) + 1$ , we get  $g$  satisfies [Example 5.6](#).

**7.2** For arbitrary  $k \in \mathbb{Z}$ , we have

(i)  $f(x) = kx^2$ ,

(ii)  $f(x) = 0$  for even  $x$ , and  $f(x) = k$  for odd  $x$ , and

(iii)  $f(x) = 0$  for  $x \equiv 0 \pmod{4}$ ,  $f(x) = k$  for odd  $x$ , and  $f(x) = 4k$  for  $x \equiv 2 \pmod{4}$ .

**7.3** Benign.  $x + c$ . Let  $a = f(0)$ ,  $b = g(0)$ ,  $c = h(0)$ , and eliminate  $g, h$ .

**7.4** For any  $c \in \mathbb{R}$ ,  $f(x) \in \{0, cx\}$  for every  $x$  and  $\{x : f(x) \neq 0\}$  is a multiplicatively closed group.

**7.5**  $f$  is a bijection, so consider its finite cycles, and look for divisibility (“order”) conditions.

**7.6**  $x, -x$  and  $\frac{1}{2} - x$ .

**7.7** Many constructions exist. Try powers of two, or defining  $f(p/q)$  in terms of  $p, q$ .

**7.8** Benign, only two trivial solutions.

**7.9**  $x$  and  $2 - x$ .

**7.10**  $f(a + b) = a - 2b$ , where  $\mathbb{R} = A \oplus B$  again. You could have seen this coming if you realized both  $x$  and  $-2x$  were solutions.

**7.11**  $f(n) = n + 1$  and

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

**7.12** There are seven solutions. To describe them, first for  $\gcd(p, q) = 1$  we let

$$g(p/q) = \begin{cases} 0 & pq \equiv 0 \pmod{3} \\ -2 & pq \not\equiv p - q \equiv 0 \pmod{3} \\ 2 & pq \not\equiv p + q \equiv 0 \pmod{3}. \end{cases}$$

Then the solutions are  $f(x) = 0$ ,  $f(x) = 2$ ,  $f(x) = x$ ,  $f(x) = g(x)$ , or the restrictions of these solutions to  $\mathbb{R}^+$  with  $f(x) = 0$  for  $x \leq 0$ .

**7.13**  $f(x) = x$ ,  $f(x) = 0$ , and for any set  $S \subset (-\infty, -\frac{2}{3})$ , the solution

$$f(x) = \begin{cases} 1 & x \notin S \\ -1 & x \in S. \end{cases}$$