The OTIS Excerpts

A collection of 192 problems and solutions

Evan Chen

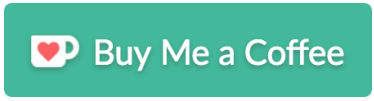
January 16, 2023

`web.evanchen.cc/excerpts.html`

The noblest art is that of making others happy.

P. T. Barnum

If you like this book and want to support me,
please consider buying me a coffee!



`http://ko-fi.com/evanchen/`

Preface

This book is a selection of notes from twelve of the lectures that I use in my year-round math olympiad classes. The program is affectionately named OTIS. The abbreviation officially stands for "Olympiad Training for Individual Study"; but in truth, I rigged the acronym so that it would match the name of an elevator in Lincoln, Nebraska of which I had fond childhood memories.

When I started teaching OTIS back in the fall of 2015, it was just a small group of students at Phillips Academy that I would meet with every couple of weeks. Despite my inexperience at the time, I have fond memories of this first group, and I maybe learned as much from them as they did from me.

Every year since then, the number of OTIS students has doubled, and the number of applications has increased even faster than that. At the same time, I became increasingly involved with volunteering with the organization of high-school contests. By the time I started graduate school, I was spending so little time on my own studies that I began to fear (rightly or wrongly) that I might fail out of the math PhD program.

Thus, despite my best efforts, I eventually had the heartbreaking task of having to tell eager and enthusiastic students that I simply did not have the time or space to take them all under my wing. I am the kind of person who finds it hard to say no, and so this was painful for me to do. OTIS taught me the reality that I am just one person.

In this way, these notes are an apology to everyone I turned down, and to everyone that I will have to turn down. I would have loved to be able to help everyone who came to my doorstep. I am sorry that I could not do more, but I wrote you a short book, as it was the least I could do.

As with all my works, there are bound to be numerous errors, mistakes, or points of unclarity. Comments, suggestions, corrections are very welcome. You can reach me at `evan@evanchen.cc`.

Evan Chen
December 30, 2018
Fremont, California, USA

Introduction

The book is divided into algebra, combinatorics, and number theory. We do not cover geometry, for which *Euclidean Geometry in Mathematical Olympiads* [Che16] already serves the role of "comprehensive book".

The twelve main chapters in this book are structured in to four sections.

- A **theoretical portion**, of varying length, in which relevant theorems or ideas are developed. Some of this material is new, but the majority of it is not. Most of it has been adapted, edited, and abridged from existing handouts that you can still find at

  `http://web.evanchen.cc/olympiad.html`.

  In general, the theoretical material here tries to stick to the basics, rather than being comprehensive.

- A couple **walkthroughs**. These are olympiad problems which are chosen to illustrate ideas, accompanied by an outline of the solution.

  When designing my lecture notes for OTIS, I wrote these walkthroughs with the idea of emulating a person. In a real classroom the student does not simply passively listen to solutions. The process is more interactive: the instructor walks a student through the example, but with a back-and-forth series of prepared questions. My hope with the walkthroughs is to simulate this as best I can with static text.

- A series of **problems**. These problems cover a range of difficulties. But in general, the first half of the problems in each chapter are intended to be fairly accessible, perhaps at the level of IMO 1/4. The difficulty increases quickly after that, with the closing problem usually being quite challenging.

- Full **solutions** to both the walkthroughs and problems. (Great for inflating page count!) Readers are encouraged to read solutions even to problems that they solved; comments, remarks, or alternate solutions frequently appear.

In addition, at the end of each part, a handful of problems chosen from USA selection tests are given, mostly for fun.

In general, I assume the reader has some minimal experience with reading and writing proofs. However, I nonetheless dedicated the first chapter to some mathematical and stylistic comments which may be helpful to beginners in proofs. Readers with significant proof experience should feel no shame in skipping this first chapter.

Contest abbreviations

Many problems have a source quoted, but there are a large number of abbreviations as a result. We tabulate some of the abbreviations here.

AIME American Invitational Math Exam, the qualifying exam for the USA national olympiad.

EGMO European Girl's Math Olympiad (not to be confused with [Che16])

ELMO The ELMO is a contest held at the USA olympiad training camp every year, written by returning students for newcomers.

The meaning of the acronym changes each year. It originally meant "Experimental Lincoln Math Olympiad" but future names have included "$e^{\log}$ Math Olympiad", "End Letter Missing", "Ex-Lincoln Math Olympiad", "English Language Master's Open". "Ego Loss May Occur", "vEry badLy naMed cOntest", "Eyyy LMaO".

ELMO Shortlist A list of problems from which each year's ELMO is chosen.

HMMT Harvard-MIT Math Tournament, the largest collegiate math competition in the United States. The contest is held twice a year, in November and February.

IMO International Math Olympiad, the supreme high-school mathematics olympiad.

IMO Shortlist A list of about 30 problems prepared annually, from which the six problems of the IMO are selected by vote.

Putnam The William Lowell Putnam Mathematical Competition, an annual competition for undergraduate students studying in USA and Canada.

RMM Romanian Masters in Mathematics, an annual olympiad held in Romania in late February for teams with a strong performance at the International Mathematical Olympiad.

TSTST The embarrassingly named "Team Selection Test Selection Test". Held in June each year, the TSTST selects students for the USA Team Selection Test.

TST Abbreviation for Team Selection Test. Most countries use a TST as the final step in the selection of their team for the International Math Olympiad.

USAJMO USA Junior Math Olympiad, the junior version of the national math olympiad for the United States (for students in 10th grade and below).

USAMO USA Math Olympiad, the national math olympiad for the United States.

Contents

1 Notes on Proofs

This is a chapter detailing common logic mistakes in proofs, as well as containing some suggestions for how to present proofs more readably. It can be safely skipped by veterans with past proof experience. There are a small number of problems at the end to try to give you practice with these types of issues.

§1.1 Common proof mistakes

§1.1.1 "Find all" problem are always two-part problems

Any problem of the form "find all..." is always implicitly a two-part problem. (This includes functional equations and Diophantine equations, for example.)

To be more explicit, if you have are asked to find all x satisfying property $P(x)$, and you think the answer is a set S, then you must prove that

$$P(x) \quad \text{if and only if} \quad x \in S.$$

Note that this is an "if and only if", so there are *two* directions, not just one!

For any solution of this form, I strongly recommend that you structure your solution as follows:

- Start by writing "**We claim the answer is ...**" and state your conjectured answer.
- Then, say "**We prove these satisfy the conditions**", and do so. For example, in a functional equation with answer $f(x) = x^2$, you should plug this f back in and verify the equation is satisfied. Even if this verification is trivial, you must still explicitly include it, because it is part of the problem.
- Finally, say "**Now we prove these are the only ones**" and do so.

This is a common mistake because many standard high school curriculums do not make this distinction explicitly, if at all. Thus your instincts might be wrong, and so you will need to adjust slightly.

To give an example of what I mean, here's an example from middle school.

Example 1. Find all real numbers x such that $3x + 2 = 17$.

Bogus Solution. Note that

$$\begin{aligned} & 3x+2=17 \\ \implies & 3x=15 \\ \implies & x=5. \end{aligned}$$

Hence the answer is $x=5$.

But really what you have shown is that $3x+2=17 \implies x=5$. You haven't proven the other direction. Fortunately, in this case it's very easy to reverse all the steps you did; $x=5 \implies 3x=15 \implies 3x+2=17$. Put another way, here is a correct solution.

Solution 1. Note that

$$\begin{aligned} & 3x+2=17 \\ \iff & 3x=15 \\ \iff & x=5. \end{aligned}$$

Therefore the answer is $x=5$. ■

No big deal, right? However it's not always true that you can simply replace $\implies$ with $\iff$.

Example 2. Find all real numbers x such that $\sqrt{x+7}=x+1$.

This time, we see something different. Consider the solution:

Bogus Solution. Note that

$$\begin{aligned} & \sqrt{x+7}=x+1 \\ \implies & x+7=x^2+2x+1 \\ \implies & 0=x^2+x-6. \\ \implies & x=-3,2. \end{aligned}$$

Hence $x=-3$ or $x=2$.

This time, the first arrow (when we square both sides) is not reversible. We have proven that $\sqrt{x+7}=x+1 \implies x=-3,2$ but this time the converse is false, since $x=-3$ does not work.

If you follow my advice to structure your solutions by stating the answer, checking it, and then proving it is the only ones, you won't make this mistake.

Solution 2. The answer is $x=2$. Since $\sqrt{2+7}=3=2+1$, it works.

We now show this is the only solution. Note that

$$\begin{aligned} &\sqrt{x+7} = x+1 \\ \implies &x+7 = x^2+2x+1 \\ \implies &0 = x^2+x-6. \\ \implies &x = -3, 2. \end{aligned}$$

Hence $x = -3$ or $x = 2$. However, we can see that $x = -3$ does not work, since $\sqrt{-3+7} = 2 \neq -2 = (-3)+1$. Therefore $x = 2$ is the only solution, as claimed. ■

Now, here is a more serious example.

Example 3 (USAJMO 2011)**.** Find all positive integers n such that $2^n + 12^n + 2011^n$ is a perfect square.

Solution 3. The answer is $n = 1$ only.

This clearly works, since $2^1 + 12^1 + 2011^1 = 2025 = 45^2$.

Now we verify this is the only solution. If n is odd and $n > 1$, then taking modulo 4 we see the $2^n + 12^n + 2011^n \equiv 3 \pmod 4$, so it is not a square. If n is even, then taking modulo 3 we see the $2^n + 12^n + 2011^n \equiv 2 \pmod 3$, so it is not a square. Thus $n = 1$ is the only solution. ■

The subtle mistake one can make is to forget to do the calculation $2^1 + 12^1 + 2011^1 = 2025 = 45^2$. To see why this is necessary, compare this with a hypothetical different problem.

Example 4. Find all positive integers n such that $2^n + 12^n + 2023^n$ is a perfect square.

Solution 4. There are no such n at all.

First, $n = 1$ does not work since $2^1 + 12^1 + 2023^1 = 2037$, which is not a square.

If n is odd and $n > 1$, then taking modulo 4 we see the $2^n + 12^n + 2023^n \equiv 3 \pmod 4$, so it is not a square. If n is even, then taking modulo 3 we see the $2^n + 12^n + 2023^n \equiv 2 \pmod 3$, so it is not a square. ■

§1.1.2 Checking for reversibility

As you can see the previous logical mistake is due to not distinguishing between $P \implies Q$ and $Q \implies P$. Many of you have been taught the wrong instincts, and now you have to adjust. For "find all" problems the surest way to do this is to just do both directions explicitly in the way I suggested.

But this won't cover everything. I'm thinking in particular of the special case $Q = \texttt{true}$. Consider the following example.

Example 5. Suppose θ and η are angles in the interval $(0, \frac{1}{2}\pi)$. Verify the trig identity

$$\tan\left(\frac{\theta+\eta}{2}\right) = \frac{\sin\eta+\sin\theta}{\cos\eta+\cos\theta}.$$

The "high-school" proof again messes up the direction of the arrows.

Bogus Solution. Write

$$\begin{aligned}
\tan\left(\frac{\theta+\eta}{2}\right) &= \frac{\sin\eta+\sin\theta}{\cos\eta+\cos\theta} \qquad (\dagger)\\
\implies \frac{\sin(\theta+\eta)}{1+\cos(\theta+\eta)} &= \frac{\sin\eta+\sin\theta}{\cos\eta+\cos\theta}\\
\implies \frac{\sin\theta\cos\eta+\cos\theta\sin\eta}{1+\cos\theta\cos\eta-\sin\theta\sin\eta} &= \frac{\sin\eta+\sin\theta}{\cos\eta+\cos\theta}\\
\implies \sin\theta\left(\cos^2\eta+\sin^2\eta\right)+\sin\eta\left(\cos^2\theta+\sin^2\theta\right) &= \sin\theta+\sin\eta\\
\implies \sin\theta+\sin\eta &= \sin\theta+\sin\eta. \qquad \checkmark
\end{aligned}$$

The second line is by tangent half-angle formula.

What you've shown is the $(\dagger) \implies$ `true`. This isn't worth anything. I have a much easier proof that $(\dagger) \implies$ `true`: just multiply both sides by zero.

What you really want is `true` $\implies (\dagger)$, which you can again do by being careful that all arrows above are $\iff$ and not $\implies$. (The condition about the angles ensures that we do not have division by zero issues.)

§1.1.3 Optimization problems are always two-part problems

Along the same lines, some problems will ask you to "find the minimum (or maximum) value of X". **These problems are always two parts as well**, you need to prove a bound on X, and then show that bound can actually be achieved.

In such situations, I strongly recommend you write your solution as follows:

- Start by writing "**We claim the minimum/maximum is ...**".
- Then, say "**We prove that this is attainable**", and give the construction (or otherwise prove existence). Even if this verification is trivial, you must still explicitly include it, because it is part of the problem.
- Finally, say "**We prove this is a lower/upper bound**", and do so.

Here is a fun example.

Example 6 (USAMO 2010)**.** The 2010 positive real numbers a_1, a_2, ..., a_{2010} satisfy the inequality $a_i a_j \le i + j$ for all $1 \le i < j \le 2010$. Determine, with proof, the largest possible value of the product $a_1 a_2 \dots a_{2010}$.

This problem is quite difficult, and will be covered in a walkthrough later. For now, we show you how to *not* solve the problem by presenting three bogus solutions which get pairwise distinct answers!

Bogus Solution. We have $a_1 a_{2010} \le 2011$, $a_2 a_{2009} \le 2011$, ..., $a_{1005} a_{1006} \le 2011$, so $a_1 a_2 \dots a_{2010} \le 2011^{1005}$.

Bogus Solution. Multiplying all the possible $\binom{2010}{2}$ inequalities together gives

$$\prod_{n=1}^{2010} a_n \le \left(\prod_{1 \le i < j \le 2010} i + j \right)^{\frac{1}{2009}}.$$

Bogus Solution. We have $a_1 a_2 \le 3$, $a_3 a_4 \le 7$, and so on, thus

$$a_1 a_2 \dots a_{2010} \le 3 \cdot 7 \cdot \dots \cdot 4019.$$

Moreover, one can prove that this is the lowest possible bound of the form $(i_1 + j_1)(i_2 + j_2) \dots (i_{1005} + j_{1005})$, where i_1, ..., j_{1005} are a permutation of 1, ..., 2010. Thus this is the answer.

All of these solutions have correctly proven an upper bound on $\prod a_i$, but none of them have made any progress on showing that *there actually exists* a_i achieving that constant, which turns out to be the true difficulty of the problem.

§1.1.4 Be neat, be careful

This list isn't exhaustive. These are just the most common mistakes that more experienced students have learned to avoid. Yet there are plenty of problems that have their *own* pitfalls.

The best thing you can do about this is to be neat and be careful.

- If a solution has cases, give each case a separate bullet point and label clearly exactly what case it is doing.
- Write out more details on parts that you feel less confident in.
- If you have central claims in the problem, write them in full as explicit lemmas in the problem.

In short, be organized.

§1.2 Stylistic writing suggestions

§1.2.1 Deciding on the level of detail

One of the most common questions I get is: "how much detail do I need to include on a contest?". The answer is actually quite simple: **enough to convince the grader you know the solution**.[1] To put it one way, whenever you omit a detail, the grader has to decide whether you know how to do it and just did not write it, or whether you don't know how to do it and are just bluffing. So if you are ever unsure about how much to write, just ask yourself that.

In still other words, you should write your solution in such a way that a student who did *not* solve the problem could not plausibly write the same thing you did. This is especially important if you have a long computational solution, for example solving geometry with complex numbers. You cannot just skip over a page of calculation by saying that "simplifying, we find this is equal to …", because a student who did not solve the problem (i.e. was not actually able to do the calculation) is perfectly capable of writing the same thing.

§1.2.2 Never write wrong math

This is much more of a math issue than a style issue: you can lose all of your points for making false claims, because this is the easiest way to convince the grader that your solution is wrong.

As a special case, don't say something that is partially true and then say how to fix it later. At best this will annoy the grader; at worst they may get confused and think the solution is wrong.

§1.2.3 Emphasize the point where you cross the ocean

Solutions to olympiad problems often involve a few key ideas, with the rest of the solution being checking details. You want graders to immediately see all the key ideas in the solution: this way, they quickly have a high-level understanding of your approach.

Let me share a quote from Scott Aaronson:

> Suppose your friend in Boston blindfolded you, drove you around for twenty minutes, then took the blindfold off and claimed you were now in Beijing. Yes, you do see Chinese signs and pagoda roofs, and no, you can't immediately disprove him — but based on your knowledge of both cars and geography, isn't it more likely

[1] This is a slightly different standard than in many other places. For example, consider the official solutions to a contest. Here the reader knows the author already has the solution, and the reader is just trying to understand it. Whereas during a contest, the grader already knows the solution, and is interested in whether you know it.

> you're just in Chinatown? …**We start in Boston, we end up in Beijing, and at no point is anything resembling an ocean ever crossed.**

Olympiad solutions work the same way: a geometry solution might require a student to do some angle chasing, deduce that two triangles are congruent, and then finish by doing a little more angle chasing. In that case, you want to highlight the key step of proving the two triangles were congruent, so the grader sees it immediately and can say "okay, this student is using this approach".

Ways that you can highlight this are:

- Isolating crucial steps and claims as their own **lemmas**.[2]
- Using **claims** to say what you're doing. Rather than doing angle chasing and writing "blah blah blah, therefore $\triangle M_B I_B M \sim \triangle M_C I_C M$", consider instead "We claim $\triangle M_B I_B M \sim \triangle M_C I_C M$, proof".
- **Displaying** important equations. For example, notice how the line

$$\triangle M_B I_B M \sim \triangle M_C I_C M \tag{1.1}$$

  jumps out at the reader. You can even number such claims to reference them later, e.g. "by (1.1)". This is especially useful in functional equations.
- Just **say it**! Little hints like "the crucial claim is X" or "the main idea is Y" are immensely helpful. Don't make X and Y look like another intermediate step.

§1.2.4 Leave space

Most people don't leave enough space. This makes solutions hard to read. Things you can do to help with this:

- Skip a line after paragraphs. Use paragraph breaks more often than you already do.
- If you isolate a specific **lemma or claim** in your proof, then it should be on its own line, with some whitespace before and after it.
- Any time you do **casework**, you should always split cases into separate paragraphs or bullet points. Make it visually clear when each case begins and ends.

[2]This is often useful for another reason: breaking the proof into individual steps. The complexity of understanding a proof grows super-linearly in its length; therefore breaking it into smaller chunks is often a good thing.

- Display important equations, rather than squeezing them into paragraphs. If you have a long calculation, then do an aligned display[3] rather than squeezing it into a paragraph. For example, instead of writing $0 \le (a-b)^2 = (a+b)^2 - 4ab = (10-c)^2 - 4(25-c(a+b)) = (10-c)^2 - 4(25-c(10-c)) = c(20-3c)$, write instead

$$\begin{aligned} 0 \le (a-b)^2 &= (a+b)^2 - 4ab \\ &= (10-c)^2 - 4(25-c(a+b)) \\ &= (10-c)^2 - 4(25-c(10-c)) \\ &= c(20-3c). \end{aligned}$$

§1.3 Problems

Problem 7. Determine, with proof, the smallest positive integer c such that for any positive integer n, the decimal representation of the number $c^n + 2014$ has digits all less than 5.

Problem 8. The numbers 1, 2, ..., 10 are written on a board. Every minute, one can select three numbers a, b, c on the board, erase them, and write $\sqrt{a^2+b^2+c^2}$ in their place. This process continues until no more numbers can be erased. What is the largest possible number that can remain on the board at this point?

Problem 9 (Putnam 2017)**.** Find the smallest set S of positive integers such that

(a) $2 \in S$,

(b) $n \in S$ whenever $n^2 \in S$,

(c) $(n+5)^2 \in S$ whenever $n \in S$.

(The set S is "smallest" in the sense that S is contained in any other such set.)

Problem 10 (USAMO 2015)**.** Steve is piling $m \ge 1$ indistinguishable stones on the squares of an $n \times n$ grid. Each square can have an arbitrarily high pile of stones. After he finished piling his stones in some manner, he can then perform *stone moves*, defined as follows. Consider any four grid squares, which are corners of a rectangle, i.e. in positions (i,k), (i,l), (j,k), (j,l) for some $1 \le i,j,k,l \le n$, such that $i < j$ and $k < l$. A stone move consists of either removing one stone from each of (i,k) and (j,l) and moving them to (i,l) and (j,k) respectively, or removing one stone from each of (i,l) and (j,k) and moving them to (i,k) and (j,l) respectively.

Two ways of piling the stones are equivalent if they can be obtained from one another by a sequence of stone moves. How many different non-equivalent ways can Steve pile the stones on the grid?

[3]This is the `align*` environment, for those of you that like LaTeX.

§1.4 Solutions

Solution 7. The answer is $c = 10$. In what follows we say that a number is *good* if all its decimal digits are less than 5.

We first prove $c = 10$ is a working example for all n. When $n = 1, 2, 3$, we have 2024, 2114 and 3014, which are all good. When $n \geq 4$, we find that

$$10^n + 2014 = 1\underbrace{000\ldots000}_{n-4 \text{ zeros}}2014$$

which is good. This shows that $c = 10$ is works.

Next, we show that $c \geq 10$ is necessary.

- For $c = 1, 2, 3, 4, 5$, taking $n = 1$ gives the numbers 2015, 2016, ..., 2019, none of which are good.
- On the other hand, for $c = 6, 7, 8, 9$, taking $n = 2$ gives the numbers 2050, 2063, 2078, 2095, none of which are good.

Solution 8. The answer is $\sqrt{384} = 8\sqrt{6}$.

We begin by observing that the *sum of the squares of all numbers on the board is preserved*. Moreover, there are initially 10 numbers, and we erase 2 at a time, so at the end of the process there will be exactly two numbers, call them a and b. By our observation, these numbers are supposed to satisfy

$$a^2 + b^2 = 1^2 + 2^2 + \cdots + 10^2 = 385. \qquad (\star)$$

We now claim that $\sqrt{384}$ is achievable. Indeed, suppose we always avoid erasing the number 1 that was initially on the board. Then at the end of the process, one of the numbers on the board is $a = 1$; thus the other one is $\sqrt{384}$ by $(\star)$.

On the other hand, observe that since all initial numbers on the board are at least 1, every number that ever appears is at least 1 as well. Consequently, in $(\star)$ we always have $a \geq 1$. Thus $b \leq \sqrt{384}$, so this is indeed maximal.

Solution 9. The answer is that S contains the positive integers greater than 1 which are not divisible by 5.

First, we check this satisfies the properties.

(a) We have $2 \in S$ by construction.

(b) If $n > 1$ then $n^2 > 1$, and if $5 \nmid n$ then $5 \nmid n^2$.

(c) If $5 \nmid n$ then $5 \nmid (n+5)^2$ and moreover $(n+5)^2 > 1$.

Next, we check that any set S satisfying the property must contain all such integers claimed. Most solutions will involve some computation (and there isn't a real reason to try to optimize it too much).

The shortest solution is to compute

$$2 \in S \implies (2+5)^2 = 49 \in S \implies (49+5)^2 = 2916 \in S.$$

Thus by (b) and (c) together we have $2916 + 5k \in S$ for every integer k. Now if $n > 1$ and $5 \nmid n$ then $n^{16} \geq 65536 > 2916$ and $n^{16} \equiv 1 \pmod 5$. The end.

Solution 10. The answer is $\binom{m+n-1}{n-1}^2$. The main observation is that the ordered sequence of column counts (i.e. the number of stones in the first, second, etc. column) is invariant under stone moves, as does the analogous sequence of row counts. We call the pair (X, Y) of sequences the *signature* of the configuration.

We are far from done. This problem is a good test of mathematical maturity since the following steps are then necessary:

1. Check that signatures are invariant around moves (trivial)
2. Check conversely that two configurations are equivalent if they have the same signatures (the hard part of the problem), and
3. Show that each signature is realized by at least one configuration (not immediate, but pretty easy).

Most procedures to the second step are algorithmic in nature, but Ankan Bhattacharya gives the following far cleaner approach. Rather than having a grid of stones, we simply consider the multiset of ordered pairs (x, y). Then, the signatures correspond to the multisets of x and y coordinates, while *a stone move corresponds to switching two y-coordinates in different pairs*, say.

Then, the second part is completed just because transpositions generate any permutation. To be explicit, given two sets of stones, we can permute the labels so that the first set is (x_1, y_1), ..., (x_m, y_m) and the second set of stones is (x_1, y_1'), ..., (x_m, y_m'). Then we just induce the correct permutation on (y_i) to get (y_i').

The third part is obvious since given two multisets $X = \{x_1, \dots, x_m\}$ and $Y = \{y_1, \dots, y_m\}$ we just put stones at (x_i, y_i) for $i = 1, \dots, m$.

In that sense, the entire grid is quite misleading!

I

Algebra

2 Fundamentals of Inequalities

This chapter covers some basic theory for olympiad inequalities, but nothing super fancy.

For those who have not seen it before, we will make extensive use of the cyclic sum notation $\sum_{\text{cyc}}$ and the symmetric sum notation $\sum_{\text{sym}}$. For a problem involving n variables, these respectively mean to cycle through the n variables, and to go through all $n!$ permutations. To provide an example, in a three-variable problem we might write

$$\sum_{\text{cyc}} a^2 = a^2 + b^2 + c^2$$
$$\sum_{\text{cyc}} a^2 b = a^2 b + b^2 c + c^2 a$$
$$\sum_{\text{sym}} a^2 = a^2 + a^2 + b^2 + b^2 + c^2 + c^2$$
$$\sum_{\text{sym}} a^2 b = a^2 b + a^2 c + b^2 c + b^2 a + c^2 a + c^2 b.$$

§2.1 Brief warning for beginners

§2.1.1 On flipped inequalities

Two points:

- If you have $X \geq Y$ and $Y \leq Z$, it does not follow $X \geq Z$.
- If you have $X \geq Y$ and $Z \leq W$, you can't add the two; no comparison on $X + Z$ and $Y + W$.

This may sound obvious, but when you're doing a full-fledged olympiad inequality it can be easy to mess up signs.

If your solution flips an inequality somewhere, it is not "just" an error; often the error is fatal, meaning there is no way to repair it.

For example, suppose you are trying to prove that

$$a^3 + b^3 + c^3 \geq a^2 b + b^2 c + c^2 a$$

for $a, b, c > 0$. You might first write down $a^3 + b^3 + c^3 \geq 3abc$ by AM-GM. So you'd be happy if you could show that $3abc \geq a^2 b + b^2 c + c^2 a$.

Unfortunately this is false! And you are dead — it is *impossible* to complete this line of thought; you will have to abandon this approach completely and try something else.

§2.1.2 Writing chains of inequalities

If you're trying to prove $A \ge B$, a good way to style the proof is by a "chain" of inequalities

$$\begin{aligned} A &\ge \text{something} \\ &\ge \text{something else} \\ &\ge \dots \\ &\ge B. \end{aligned}$$

This way you will be less likely to make a mistake because it's clear which way everything is going.

I should also mention that the comments from Section 1.1.2 apply here as well. Be careful not to show ineq $\implies$ `true`. Either deduce the desired inequality as mentioned above, or else be very careful that all your steps are reversible, indicating this explicitly with $\iff$.

§2.2 Polynomial inequalities

§2.2.1 AM-GM and Muirhead

The most basic inequality to start out with is the following.

Theorem 2.1 (AM-GM)**.** *For nonnegative real numbers a_1, a_2, …, a_n we have*

$$\frac{a_1 + a_2 + \dots + a_n}{n} \ge \sqrt[n]{a_1 \dots a_n}.$$

Equality holds if and only if $a_1 = a_2 = \dots = a_n$.

The abbreviation stands from "Arithmetic Mean, Geometric Mean" (with the left-hand side being "arithmetic mean", the right-hand side being "geometric mean"). For example, this implies that for $a, b, c > 0$ we have

$$a^2 + b^2 \ge 2ab, \quad a^3 + b^3 + c^3 \ge 3abc.$$

The simplest problems can be solved by summing applications of AM-GM.

Example 11. For $a, b, c > 0$ prove that $a^2 + b^2 + c^2 \ge ab + bc + ca$.

Solution 11. We consider the following three applications of AM-GM:

$$\begin{aligned} a^2 + b^2 &\ge 2ab \\ b^2 + c^2 &\ge 2bc \\ c^2 + a^2 &\ge 2ca. \end{aligned}$$

Summing and dividing by 2 yields the result. ■

Example 12. For $a, b, c > 0$ prove that $a^4 + b^4 + c^4 \geq a^2bc + b^2ca + c^2ab$.

Solution 12. This is a little more subtle than the previous one, but it has the same shape. By AM-GM,

$$a^4 + a^4 + b^4 + c^4 \geq 4a^2bc$$
$$b^4 + b^4 + c^4 + a^4 \geq 4ab^2c$$
$$c^4 + c^4 + a^4 + b^4 \geq 4abc^2.$$

Summing these and dividing by 4 yields the correct result. ■

Exercise. If $a, b, c > 0$ prove that $a^3 + b^3 + c^3 \geq a^2b + b^2c + c^2a$.

Exercise. If $a, b, c > 0$ prove that $a^5 + b^5 + c^5 \geq a^3bc + b^3ca + c^3ab \geq abc(ab + bc + ca)$.

§2.2.2 Some vague cheerleading

You might already be picking up some connotations of the types of problems we consider:

- In a "stereotypical" symmetric inequality, both sides will be equal when we set all variables equal.
- Moreover, in the absence of other conditions, we often compare expressions which are the same degree, or *homogeneous*. For example when we write $a^2 + b^2 + c^2 \geq ab + bc + ca$, both sides are degree 2. (Notice that the AM-GM inequality itself has the same property!)

  There is a good reason for this: x^5 and x^3 are not comparable for generic $x > 0$, since the behaviors when x is very small and x is very large are different. So a non-homogeneous inequality like $a^2+b^2+c^2 \geq a^3+b^3+c^3$ will definitely not be true in general, since the behaviors if I take $a = b = c = 0.01$ and $a = b = c = 100$ will be different.

You may also already be picking up some intuition: **more "mixed" terms are smaller**. For example, for degree 3, the polynomial $a^3 + b^3 + c^3$ is biggest and $3abc$ is the smallest. Roughly, the more "mixed" polynomials are the smaller.

If you internalize this intuition well, you might already be able to see that

$$(a + b + c)^3 \geq a^3 + b^3 + c^3 + 24abc$$

must be true, just by looking. Indeed, it is homegeneous with equality when $a = b = c$. But more importantly, since upon expanding the LHS and cancelling $a^3 + b^3 + c^3$, we find that the RHS contains only the piddling term $24abc$. That means a straight AM-GM will suffice.

§2.2.3 Muirhead's inequality

In the case of a *symmetric* inequality, this intuition has actually been formalized, by the so-called Muirhead's inequality.

Definition 2.2. Suppose we have two sequences $x_1 \ge x_2 \ge \cdots \ge x_n$ and $y_1 \ge y_2 \ge \cdots \ge y_n$ such that

$$x_1 + x_2 + \cdots + x_n = y_1 + y_2 + \cdots + y_n,$$

and for $k = 1, 2, \ldots, n-1$,

$$x_1 + x_2 + \cdots + x_k \ge y_1 + y_2 + \cdots + y_k.$$

Then we say that (x_n) **majorizes** (y_n), written $(x_n) \succ (y_n)$.

Using the above, we have the following theorem.

Theorem 2.3 (Muirhead's inequality)**.** *If a_1, a_2, ..., a_n are nonnegative real numbers and the sequence $(x_1, \ldots, x_n)$ majorizes the sequence $(y_1, \ldots, y_n)$ then we have the inequality*

$$\sum_{sym} a_1^{x_1} a_2^{x_2} \ldots a_n^{x_n} \ge \sum_{sym} a_1^{y_1} a_2^{y_2} \ldots a_n^{y_n}.$$

For example, since $(5,0,0) \succ (3,1,1) \succ (2,2,1)$, Muirhead implies that

$$\begin{aligned} a^5 + a^5 + b^5 + b^5 + c^5 + c^5 &\ge a^3bc + a^3bc + b^3ca + b^3ca + c^3ab + c^3ab \\ &\ge a^2b^2c + a^2b^2c + b^2c^2a + b^2c^2a + c^2a^2b + c^2a^2b. \end{aligned}$$

From this we derive $a^5 + b^5 + c^5 \ge a^3bc + b^3ca + c^3ab \ge abc(ab + bc + ca)$, one of the earlier exercises.

Remark 2.4. It can be shown that, if one could prove an inequality by Muirhead, then one could also have proved it by repeated AM-GM with carefully chosen weights. However, it is much simpler to simply quote Muirhead directly, so that one does not need to refer to explicit weights.

Notice that Muirhead is *symmetric*, not *cyclic*. For example, even though $(3,0,0) \succ (2,1,0)$, Muirhead's inequality only gives that

$$2(a^3 + b^3 + c^3) \ge a^2b + a^2c + b^2c + b^2a + c^2a + c^2b$$

and in particular this does *not* imply that $a^3 + b^3 + c^3 \ge a^2b + b^2c + c^2a$. These situations must still be resolved by AM-GM.

§2.2.4 Non-homogeneous inequalities

Consider the following example.

Example 13. Let $a, b, c > 0$ and assume $abc = 1$. Prove that $a^2 + b^2 + c^2 \geq a + b + c$.

The inequality has a degree 2 right-hand side, and a degree 1 left-hand side. It also has a condition $abc = 1$. Both of these are undesirable, and the following solution shows how can deal with them.

Solution 13. AM-GM alone is hopeless here, because whenever we apply AM-GM, the left and right hand sides of the inequality all have the same degree. So we want to use the condition $abc = 1$ to force the problem to have the same degree. The trick is to notice that the given inequality can be rewritten as

$$a^2 + b^2 + c^2 \geq a^{1/3}b^{1/3}c^{1/3}\left(a + b + c\right).$$

Now the inequality is homogeneous.

An important point now is that, once written this way, **the restriction $abc = 1$ stops mattering**. Because observe that if we multiply a, b, c by any real number $k > 0$, all that happens is that both sides of the inequality are multiplied by k^2, which doesn't change anything. So if the inequality is true for all $abc = 1$, it is also true for all $abc = 8$ (by doubling each of a, b, c) or all $abc = 27$ (by tripling each of a, b, c), or indeed regardless of what abc equals. So we can treat this reduced problem without the condition, at which point it looks like the examples we did earlier.

In particular, $(2, 0, 0) \succ (\frac{4}{3}, \frac{1}{3}, \frac{1}{3})$, and so applying Muirhead's inequality solves the problem. ■

The importance of this problem is that it shows us how to eliminate a given condition by homogenizing the inequality; this is very important. (In fact, we will soon see that we can use this in reverse — we can impose an arbitrary condition on a homogeneous inequality.)

Exercise. Let $a, b, c > 0$ with $a + b + c = 1$. Show that

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} \leq 3 + 2 \cdot \frac{(a^3 + b^3 + c^3)}{abc}.$$

Exercise. Let $a, b, c > 0$ with $\frac{1}{a} + \frac{1}{b} + \frac{1}{c} = 1$. Prove that

$$(a + 1)(b + 1)(c + 1) \geq 64.$$

§2.3 Three polynomial tricks

§2.3.1 The special case of product 1

If an inequality has the condition $abc = 1$, one can also sometimes use the substitution $(a, b, c) = (x/y, y/z, z/x)$ which will transform it into a homogeneous inequality automatically.

§2.3.2 Ravi substitution

Sometimes, an inequality will refer to the a, b, c as the sides of a triangle. In that case, one can replace $(a, b, c) = (y + z, z + x, x + y)$ where $x, y, z > 0$ are real numbers. This is colloquially known as the Ravi substitution, in folklore.

§2.3.3 Schur's inequality

The following inequality, despite being polynomial in nature, cannot be proven using AM-GM easily, and so we record it here as a separate theorem.

Theorem 2.5 (Schur)**.** *Let a, b, c be nonnegative real numbers and let $r > 0$ be any positive real number. Then*

$$\sum_{cyc} a^r(a^2 + bc) \geq \sum_{cyc} a^{r+1}(b + c).$$

Equality occurs if $a = b = c$ or two of the variables are equal and the last is zero.

For example, the $r = 1$ case of this theorem says that

$$a^3 + b^3 + c^3 + 3abc \geq \sum_{\text{sym}} a^2 b$$

and is perhaps the most commonly used variant.

Proof. Assume without loss of generality $a \geq b \geq c$. Rewrite the inequality as

$$\begin{aligned}(a - b)\left[a^r(a - c) - b^r(b - c)\right] \\ + c^r(c - a)(c - b) \geq 0.\end{aligned}$$

Since $a \geq b$, we have $a^r(a - c) \geq b^r(b - c)$, and from this it's clear that each term on the left-hand side is nonnegative, as needed. □

§2.4 Eliminating radicals and fractions

§2.4.1 Weighted Power Mean

AM-GM has the following natural generalization.

Theorem 2.6 (Weighted Power Mean)**.** *Let a_1, …, a_n be positive real numbers. Let w_1, w_2, …, w_n be positive real numbers with $w_1 + w_2 + \cdots + w_n = 1$. For any real number r, we define*

$$\mathcal{P}(r) = \begin{cases} \left(w_1 a_1^r + w_2 a_2^r + \cdots + w_n a_n^r\right)^{1/r} & r \neq 0 \\ a_1^{w_1} a_2^{w_2} \ldots a_n^{w_n} & r = 0. \end{cases}$$

If $r > s$, then $\mathcal{P}(r) \geq \mathcal{P}(s)$. Equality occurs if and only if $a_1 = a_2 = \cdots = a_n$.

The quantity $\mathcal{P}(r)$ is called the rth power mean. Note that if we set all the weights equal, that is $w_1 = w_2 = \cdots = w_n = \frac{1}{n}$, then

$$\mathcal{P}(r) = \begin{cases} \left(\dfrac{a_1^r + a_2^r + \cdots + a_n^r}{n} \right)^{1/r} & r \neq 0 \\ \sqrt[n]{a_1 a_2 \ldots a_n} & r = 0. \end{cases}$$

Corollary 2.7 (QM-AM-GM-HM theorem)**.** *Let a_1, …, a_n be positive real numbers. Then*

$$\sqrt{\frac{a_1^2 + \cdots + a_n^2}{n}} \geq \frac{a_1 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 a_2 \ldots a_n} \geq \frac{n}{\frac{1}{a_1} + \cdots + \frac{1}{a_n}}.$$

Proof. Set $r \in \{2, 1, 0, -1\}$ we obtain the inequality □

Here "QM" and "HM" stand for "quadratic mean" and "harmonic mean" Here is an application of a $\frac{1}{3}$-power mean.

Example 14 (Taiwan TST 2014)**.** Let $a, b, c > 0$. Prove that

$$3(a+b+c) \geq 8\sqrt[3]{abc} + \sqrt[3]{\frac{a^3+b^3+c^3}{3}}.$$

Solution 14. By Power Mean with $r = 1$, $s = \frac{1}{3}$, and weights $\frac{1}{9} + \frac{8}{9} = 1$ we have the inequality

$$\left(\frac{1}{9}\sqrt[3]{\frac{a^3+b^3+c^3}{3}} + \frac{8}{9}\sqrt[3]{abc} \right)^3 \leq \frac{1}{9}\left(\frac{a^3+b^3+c^3}{3} \right) + \frac{8}{9}(abc).$$

Thus it is enough to prove $a^3+b^3+c^3+24abc \leq (a+b+c)^3$, which is clear. ■

§2.4.2 Cauchy and Hölder

We now present Hölder's inequality; we state the two-variable form for concreteness but the obvious generalization to any number of sequences is valid.

Theorem 2.8 (Hölder's inequality)**.** *Let p and q be positive real numbers. Let a_1, …, a_n, b_1, …, b_n be nonnegative real numbers. Then*

$$\left(\sum_{i=1}^n a_i \right)^p \left(\sum_{i=1}^n b_i \right)^q \geq \left(\sum_{i=1}^n \sqrt[p+q]{a_i^p b_i^q} \right)^{p+q}.$$

Proof. We will only address the case where the left hand side is not zero (since otherwise one of the sequences is entirely zero, and there is nothing to prove). By scaling the a_i's (since both sides have the same degree), we may as well assume that $\sum a_i = 1$. Similarly we assume $\sum a_i = \sum b_i = 1$. Then by AM-GM,

$$\sum_{i=1}^{n} \sqrt[p+q]{a_i^p b_i^q} \le \sum_{i=1}^{n} \frac{p \cdot a_i + q \cdot b_i}{p+q} = 1. \qquad \square$$

Hölder is often useful for eliminating radicals. The situation $p = q = 1$ gives the famous Cauchy inequality, which can be rewritten as

$$\frac{x_1^2}{y_1} + \frac{x_2^2}{y_2} + \cdots + \frac{x_n^2}{y_n} \ge \frac{(x_1 + x_2 + \cdots + x_n)^2}{y_1 + \cdots + y_n}.$$

This form is often called Titu's Lemma in the United States, where it is used to eliminate fractions.

Here are two examples for illustration. The first has a denominator; the second has both a denominator and a radical.

Example 15 (Nesbitt's inequality)**.** For $a, b, c > 0$ prove that

$$\frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b} \ge \frac{3}{2}.$$

Solution 15. By Cauchy,

$$\begin{aligned}\left(\sum_{\text{cyc}} \frac{a}{b+c}\right)^1 \left(\sum_{\text{cyc}} a(b+c)\right)^1 &\ge \left(\sum_{\text{cyc}} \sqrt{\frac{a}{b+c} \cdot a(b+c)}\right)^2 \\ &= (a+b+c)^2.\end{aligned}$$

Therefore, it is enough to prove that

$$(a+b+c)^2 \ge \frac{3}{2} \sum_{\text{cyc}} a(b+c)$$

which follows by expanding and applying Muirhead's inequality. ■

Remark 2.9. The solution above can also be rewritten to use Titu's lemma:

$$\sum_{\text{cyc}} \frac{a^2}{a(b+c)} \ge \frac{(a+b+c)^2}{\sum_{\text{cyc}} a(b+c)}.$$

Example 16. For $a, b, c > 0$ prove that

$$\frac{a}{\sqrt{b+c}} + \frac{b}{\sqrt{c+a}} + \frac{c}{\sqrt{a+b}} \ge \sqrt{\frac{3}{2}(a+b+c)}.$$

Solution 16. This time, we use Hölder with slightly changed weights in order to remove the square root:

$$\left(\sum_{\text{cyc}} \frac{a}{\sqrt{b+c}}\right)^2 \left(\sum_{\text{cyc}} a(b+c)\right) \geq (a+b+c)^3.$$

Again it is enough to prove $(a+b+c)^2 \geq \frac{3}{2}\sum_{\text{cyc}} a(b+c)$ which is true by expanding. ■

Exercise. Show that if one sets $b_i = 1$ for each i, then Hölder's inequality reduces to a power mean inequality with all weights equal.

§2.5 Inequalities in arbitrary functions

Let I be an open interval (for example $I = (0,\infty)$ or $I = (0,1)$) and let $f\colon (u,v) \to \mathbb{R}$ be a function and let $a_1, a_2, \ldots, a_n \in (u,v)$. Suppose that we fix $\frac{a_1+a_2+\cdots+a_n}{n} = a$ (if the inequality is homogeneous, we will often insert such a condition) and we want to prove that

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

is at least (or at most) $nf(a)$. In this section we will provide two methods for doing so.

Definition 2.10. We say that function f is **convex** if the second derivative f'' is nonnegative over all of (u,v). Similarly we say it is **concave** if $f''(x) \leq 0$ for all x. Note that f is convex if and only if $-f$ is concave.

§2.5.1 Jensen and Karamata

We have the following analog of AM-GM now.

Theorem 2.11 (Jensen's inequality)**.** *Let $f\colon I \to \mathbb{R}$ be a convex function. Then for any $a_1, \ldots, a_n \in I$ we have*

$$\frac{f(a_1) + \cdots + f(a_n)}{n} \geq f\left(\frac{a_1 + \cdots + a_n}{n}\right).$$

The reverse inequality holds when f is concave.

Exercise. Show that if one takes $I = (0,\infty)$ and f to be the natural logarithm, then Jensen reduces to AM-GM with all weights equal.

Just as Muirhead is repeated AM-GM, there is an analog of repeated Jensen; however its use is somewhat rarer.

Theorem 2.12 (Karamata's inequality)**.** *Let $f\colon I \to \mathbb{R}$ be convex. Suppose the sequence (x_n) majorizes (y_n), with each x_i and y_i in I. Then*

$$f(x_1) + \cdots + f(x_n) \geq f(y_1) + \cdots + f(y_n).$$

The reverse inequality holds when f is concave.

§2.5.2 Tangent line trick

Again fix $a = \frac{a_1+\cdots+a_n}{n}$. If f is not convex, we can sometimes still prove the inequality

$$f(x) \geq f(a) + f'(a)(x-a).$$

If this inequality manages to hold for all x, then simply summing the inequality will give us the desired conclusion. This method is called the *tangent line trick*.

Example 17 (Japanese Olympiad 1997)**.** Let a, b, c be positive reals. Prove that

$$\frac{(b+c-a)^2}{a^2+(b+c)^2} + \frac{(c+a-b)^2}{b^2+(c+a)^2} + \frac{(a+b-c)^2}{c^2+(a+b)^2} \geq \frac{3}{5}.$$

Solution 17. Since the inequality is homogeneous, we may assume WLOG that $a+b+c=3$. So the inequality we wish to prove is

$$\sum_{\text{cyc}} \frac{(3-2a)^2}{a^2+(3-a)^2} \geq \frac{3}{5}.$$

Let $f(x) = \frac{(3-2x)^2}{x^2+(3-x)^2}$. In an ideal world, f would be convex, and we could finish by applying Jensen's inequality. We do not live in an ideal world, and f is not convex.

Nonetheless, we can work around the issue by trying to prove that f lies above its tangent line at $x=1$. A computation gives that $f'(1) = -\frac{18}{25}$ and so we are motivated to try and prove

$$\frac{(3-2x)^2}{(3-x)^2+x^2} \geq -\frac{18}{25}(x-1) + \frac{1}{5}.$$

In fact, if we expand and factor the resulting inequality, we find that it actually is equivalent to

$$\frac{18(x-1)^2(2x+1)}{25(2x^2-6x+9)} \geq 0$$

which is obviously true. ∎

§2.6 Walkthroughs

Problem 18. If $abcd = 1$ for $a, b, c, d > 0$, prove that

$$a^4b + b^4c + c^4d + d^4a \geq a + b + c + d.$$

Walkthrough. There are two possible solutions I know of, one by Hölder and one by AM-GM. I find the latter much more natural.

(a) Homogenize the inequality to eliminate the condition (while keeping the inequality fifth-degree).

(b) Fill in the blanks in the following AM-GM:

$$? \cdot a^4b + ? \cdot b^4c + ? \cdot c^4d + ? \cdot d^4a \geq a^2bcd.$$

(c) Cyclically sum to finish.

Depending on how you did this, the number 51 might appear.

Problem 19 (IMO 2001)**.** Let a, b, c be positive reals. Prove that

$$\frac{a}{\sqrt{a^2+8bc}} + \frac{b}{\sqrt{b^2+8ca}} + \frac{c}{\sqrt{c^2+8ab}} \geq 1.$$

Walkthrough. There are a few ways to set up, but the general idea is to use Hölder in the form

$$\left(\sum_{\text{cyc}} \frac{a}{\sqrt{a^2+8bc}}\right)^p \left(\sum_{\text{cyc}} ?\right)^q \geq \left(\sum_{\text{cyc}} ?\right)^{p+q}$$

for some choice of weights p and q to eliminate the radicals and get a polynomial inequality.

(a) Pick a choice of weights $p, q > 0$ eliminate the radicals.

(b) Decide on values to fill in the ? above. (You probably want to eliminate the denominator, i.e. the left sum should be some multiple of $a^2 + 8bc$.)

(c) Try to prove the resulting inequality. Depending on what choices you made in (a) or (b), this may be relatively easy, or it may be impossible (because the inequality may not even be true.)

Problem 20 (IMO Shortlist 2009)**.** Let a, b, c be positive real numbers such that $\frac{1}{a} + \frac{1}{b} + \frac{1}{c} = a + b + c$. Prove that:

$$\frac{1}{(2a+b+c)^2} + \frac{1}{(a+2b+c)^2} + \frac{1}{(a+b+2c)^2} \leq \frac{3}{16}.$$

Walkthrough. This is sort of a canonical Jensen problem.
The first step is almost forced upon us.

(a) Homogenize the inequality to eliminate the constraint.

It's not 100% true that we *always* want to homogenize right away, although it is quite often a good start. Sometimes there is some reason not to homogenize. But this is not the case here. The condition $\frac{1}{a} + \frac{1}{b} + \frac{1}{c} = a + b + c$ is not even tangentially related to the inequality we want to prove, and is in any case an

abomination. So for this problem I think it would be hard to come up for a reason *not* to eliminate the constraint.

However, we will immediately turn around and recognize that if we set $a+b+c=3$, we can turn it into a sum of functions. And so we just follow through:

(b) De-homogenize the inequality in such a way that one can rewrite the inequality in the form $f(a)+f(b)+f(c) \le 0$ where $a+b+c=3$.

(c) Assuming you defined f correctly, show that (up to constant factors, depending on how you defined f),

$$f''(x) = \frac{96}{(x+3)^4} - \frac{2}{x^3}.$$

(d) Prove that f is concave over the interval $[0,3]$.

(e) Finish by Jensen.

Problem 21 (ELMO Shortlist 2013)**.** Let a, b, c be positive real numbers with $a+b+c=3$. Prove that

$$\frac{18}{(3-a)(4-a)} + \frac{18}{(3-b)(4-b)} + \frac{18}{(3-c)(4-c)} + 2(ab+bc+ca) \ge 15.$$

Walkthrough. This is a fairly token application of the so-called tangent line trick.

(a) Rewrite the inequality in the form $f(a)+f(b)+f(c) \ge 6$ for some function $f\colon (0,3) \to \mathbb{R}$ (where $a+b+c=3$).

(b) Check that $f(1)=2$, so equality holds when $a=b=c=1$.

The tangent line trick then leads us to conjecture that

$$f(x) \ge f'(1)(x-1)+2$$

is true for all real numbers $x \in (0,3)$. (Here $f'(1)$ is the derivative of f at 1.)

(c) Show that the inequality is valid for all real numbers $x \in (0,3)$.

(d) Sum up to finish.

§2.7 Problems

Problem 22 (Canadian Olympiad 2002)**.** Let a, b, c be positive reals. Prove that

$$\frac{a^3}{bc} + \frac{b^3}{ca} + \frac{c^3}{ab} \ge a+b+c.$$

Problem 23 (USAJMO 2012). For $a, b, c > 0$ prove that

$$\frac{a^3+3b^3}{5a+b}+\frac{b^3+3c^3}{5b+c}+\frac{c^3+3a^3}{5c+a}\geq\frac{2}{3}(a^2+b^2+c^2).$$

Problem 24 (IMO 2000). Let a, b, c be positive real numbers with $abc = 1$. Show that

$$\left(a-1+\frac{1}{b}\right)\left(b-1+\frac{1}{c}\right)\left(c-1+\frac{1}{a}\right)\leq 1.$$

Problem 25 (ELMO 2003). Let $x, y, z > 1$ be real numbers such that

$$\frac{1}{x^2-1}+\frac{1}{y^2-1}+\frac{1}{z^2-1}=1.$$

Prove that

$$\frac{1}{x+1}+\frac{1}{y+1}+\frac{1}{z+1}\leq 1.$$

Problem 26 (USAMO 2003). Let a, b, c be positive real numbers. Prove that

$$\frac{(2a+b+c)^2}{2a^2+(b+c)^2}+\frac{(2b+c+a)^2}{2b^2+(c+a)^2}+\frac{(2c+a+b)^2}{2c^2+(a+b)^2}\leq 8.$$

Problem 27 (USAMO 2017). Find the minimum possible value of

$$\frac{a}{b^3+4}+\frac{b}{c^3+4}+\frac{c}{d^3+4}+\frac{d}{a^3+4}$$

given that a, b, c, d are nonnegative real numbers such that $a+b+c+d=4$.

Problem 28 (USAMO 2004). Let a, b, c be positive reals. Prove that

$$(a^5-a^2+3)(b^5-b^2+3)(c^5-c^2+3)\geq(a+b+c)^3.$$

Problem 29 (TSTST 2012). Positive real numbers x, y, z satisfy $xyz+xy+yz+zx=x+y+z+1$. Prove that

$$\frac{1}{3}\left(\sqrt{\frac{1+x^2}{1+x}}+\sqrt{\frac{1+y^2}{1+y}}+\sqrt{\frac{1+z^2}{1+z}}\right)\leq\left(\frac{x+y+z}{3}\right)^{5/8}.$$

Problem 30 (IMO Shortlist 2003). Let n be a positive integer and let $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$ be two sequences of positive real numbers. Suppose $(z_2, \ldots, z_{2n})$ is a sequence of positive real numbers such that $z_{i+j}^2 \geq x_i y_j$ for all $1 \leq i, j \leq n$. Let $M = \max\{z_2, \ldots, z_{2n}\}$. Prove that

$$\left(\frac{M+z_2+z_3+\cdots+z_{2n}}{2n}\right)^2\geq\left(\frac{x_1+\cdots+x_n}{n}\right)\left(\frac{y_1+\cdots+y_n}{n}\right).$$

Problem 31 (ELMO 2013). Let a, b, c be positive reals satisfying $a+b+c=\sqrt[7]{a}+\sqrt[7]{b}+\sqrt[7]{c}$. Prove that $a^a b^b c^c \geq 1$.

§2.8 Solutions

Solution 18. We present two solutions.

First solution by weighted AM-GM By AM-GM,

$$\frac{23a^4b + 7b^4c + 11c^4d + 10d^4a}{51} \geq \sqrt[51]{a^{102}b^{51}c^{51}d^{51}} = a^2bcd = a.$$

You could find this solution by searching for weights w, x, y, z with sum 1 for which $w \cdot a^4b + x \cdot b^4c + y \cdot c^4d + z \cdot d^4a \geq a^2bcd$ holds; this amounts to the system of equations

$$\begin{aligned} 4w + x &= 2 \\ 4x + y &= 1 \\ 4y + z &= 1 \\ 4z + w &= 1 \end{aligned}$$

which when solved gives the weights above.

Second solution by Hölder By Hölder,

$$\left(\sum_{\text{cyc}} a^4b\right)\left(\sum_{\text{cyc}} a\right)\left(\sum_{\text{cyc}} c\right)\left(\sum_{\text{cyc}} d\right) \geq \left(\sum_{\text{cyc}} \sqrt[4]{a^4 \cdot abcd}\right)^4 = (a+b+c+d)^4.$$

Thus done.

Solution 19. By Holder, we have

$$\left(\sum_{\text{cyc}} \frac{a}{\sqrt{a^2+8bc}}\right)^2 \left(\sum_{\text{cyc}} a(a^2+8bc)\right) \geq (a+b+c)^3.$$

So it suffices to show $(a+b+c)^3 \geq a^3+b^3+c^3+24abc$ which is clear by expanding.

Solution 20. Homogenize to get rid of constraint:

$$\sum_{\text{cyc}} \left(\frac{16}{(2a+b+c)^2} - \frac{3}{a(a+b+c)}\right) \leq 0$$

To make this a sum of functions, we then *de-homogenize* with the condition $a+b+c=3$; thus we wish to show

$$\sum_{\text{cyc}} \left(\frac{16}{(a+3)^2} - \frac{1}{a}\right) \leq 0 \qquad a+b+c=3.$$

Let $f(x) = 16/(x+3)^2 - 1/x$, so $f(1) = 0$. Then

$$f''(x) = \frac{96}{(x+3)^4} - \frac{2}{x^3} \le 0$$

This is concave for $x \in [0,3]$ since for x in this interval we have $(x+3)^4 - 48x^3 = (x-3)(x^3 - 33x^2 - 45x - 27) \ge 0$. (In fact $f''(3) = 0$.) Consequently we are done as

$$f(a) + f(b) + f(c) \le 3f\left(\frac{a+b+c}{3}\right) = 3f(1) = 0$$

by Jensen.

Solution 21. Since $2(ab+bc+ca) = (a+b+c)^2 - (a^2+b^2+c^2) = 9 - (a^2+b^2+c^2)$, we can rewrite the given inequality as

$$\sum_{\text{cyc}} \left(\frac{18}{(3-c)(4-c)} - c^2 \right) \ge 6.$$

Using the tangent line trick lets us obtain the magical inequality

$$\frac{18}{(3-c)(4-c)} - c^2 \ge \frac{c+3}{2} \iff c(c-1)^2(2c-9) \le 0$$

and summing cyclically yields the result.

Solution 22. By the AM-GM inequality, we have

$$\frac{2\frac{a^3}{bc} + \frac{b^3}{ca} + \frac{c^3}{ab}}{4} \ge \sqrt[4]{\frac{a^3 \cdot a^3 \cdot b^3 \cdot c^3}{bc \cdot bc \cdot ca \cdot ab}} = a.$$

Thus we are done by summing cyclically. Alternatively, one can just quote Muirhead as the sequence $(3,-1,-1)$ majorizes $(1,0,0)$.

Solution 23. Apply Titu lemma to get

$$\sum_{\text{cyc}} \frac{a^3}{5a+b} = \sum_{\text{cyc}} \frac{a^4}{5a^2+ab} \ge \frac{(a^2+b^2+c^2)^2}{\sum_{\text{cyc}}(5a^2+ab)} \ge \frac{a^2+b^2+c^2}{6}$$

where the last step follows from the identity $\sum_{\text{cyc}}(5a^2+ab) \le 6(a^2+b^2+c^2)$.

Similarly,

$$\sum_{\text{cyc}} \frac{b^3}{5a+b} = \sum_{\text{cyc}} \frac{b^4}{5ab+b^2} \ge \frac{(a^2+b^2+c^2)^2}{\sum_{\text{cyc}}(5ab+b^2)} \ge \frac{a^2+b^2+c^2}{6}$$

using the fact that $\sum_{\text{cyc}} 5ab + b^2 \le 6(a^2+b^2+c^2)$.

Therefore, adding the first display to three times the second display implies the result.

Solution 24. Let $a = x/y$, $b = y/z$, $c = z/x$ for $x, y, z > 0$. Then the inequality rewrites as

$$(-x+y+z)(x-y+z)(x+y-z) \le xyz$$

which when expanded is equivalent to Schur's inequality. Alternatively, if one wants to avoid appealing to Schur, then the following argument works:

- At most one term on the left-hand side is negative; if that occurs we are done from $xyz > 0 > (-x+y+z)(x-y+z)(x+y-z)$.
- If all terms in the left-hand side are nonnegative, let us denote $m = -x+y+z \ge 0$, $n = x-y+z \ge 0$, $p = x+y-z \ge 0$. Then it becomes

  $$mnp \le \frac{(m+n)(n+p)(p+m)}{8}$$

  which follows by AM-GM.

Solution 25. We give two solutions. One is a tricky Cauchy-Schwarz application, the second is a straightforward Jensen.

First solution (Evan Chen) The key identity is

$$\left(\sum_{\text{cyc}} \frac{1}{x^2-1}\right)\left(\sum_{\text{cyc}} \frac{x-1}{x+1}\right) \ge \left(\sum_{\text{cyc}} \frac{1}{(x+1)}\right)^2.$$

which is of course Cauchy-Schwarz. Thus if we denote the sum in question by $S \ge 0$ we then have

$$1 \cdot (3-2S) \ge S^2 \implies S \le 1.$$

Second solution by Jensen (Ryan Kim) Let $a = \frac{1}{x^2-1}$, so $x = \sqrt{1+1/a}$, et cetera. Then $a+b+c = 1$ and we wish to show

$$\sum_{\text{cyc}} \frac{1}{\sqrt{1+\frac{1}{a}}+1} \le 1.$$

But the function $f(x) = \frac{1}{\sqrt{1+\frac{1}{x}}+1}$ is concave, and so we are done by Jensen inequality.

Solution 26. This is a canonical example of tangent line trick. Homogenize so that $a+b+c=3$. The desired inequality reads

$$\sum_{\text{cyc}} \frac{(a+3)^2}{2a^2+(3-a)^2} \le 8.$$

This follows from

$$f(x) = \frac{(x+3)^2}{2x^2+(3-x)^2} \le \frac{1}{3}(4x+4)$$

which can be checked as $\frac{1}{3}(4x+4)(2x^2+(3-x)^2)-(x+3)^2 = (x-1)^2(4x+3) \ge 0$.

Solution 27. The minimum $\frac{2}{3}$ is achieved at $(a,b,c,d)=(2,2,0,0)$ and cyclic permutations.

The problem is an application of the tangent line trick: we observe the miraculous identity

$$\frac{1}{b^3+4} \ge \frac{1}{4} - \frac{b}{12}$$

since $12-(3-b)(b^3+4) = b(b+1)(b-2)^2 \ge 0$. Moreover,

$$ab+bc+cd+da = (a+c)(b+d) \le \left(\frac{(a+c)+(b+d)}{2}\right)^2 = 4.$$

Thus

$$\sum_{\text{cyc}} \frac{a}{b^3+4} \ge \frac{a+b+c+d}{4} - \frac{ab+bc+cd+da}{12} \ge 1 - \frac{1}{3} = \frac{2}{3}.$$

Remark. The main interesting bit is the equality at $(a,b,c,d)=(2,2,0,0)$. This is the main motivation for trying tangent line trick, since a lower bound of the form $\sum a(1-\lambda b)$ preserves the unusual equality case above. Thus one takes the tangent at $b=2$ which miraculously passes through the point $(0,1/4)$ as well.

Solution 28. Observe that for all real numbers a, the inequality

$$a^5 - a^2 + 3 \ge a^3 + 2$$

holds. Then the problem follows by Hölder in the form

$$(a^3+1+1)(1+b^3+1)(1+1+c^3) \ge (a+b+c)^3.$$

Solution 29. The key is the identity

$$\begin{aligned}\frac{x^2+1}{x+1} &= \frac{(x^2+1)(y+1)(z+1)}{(x+1)(y+1)(z+1)} \\ &= \frac{x(xyz+xy+xz)+x^2+yz+y+z+1}{2(1+x+y+z)} \\ &= \frac{x(x+y+z+1-yz)+x^2+yz+y+z+1}{2(1+x+y+z)} \\ &= \frac{(x+y)(x+z)+x^2+(x-xyz+y+z+1)}{2(1+x+y+z)} \\ &= \frac{2(x+y)(x+z)}{2(1+x+y+z)} \\ &= \frac{(x+y)(x+z)}{1+x+y+z}.\end{aligned}$$

Remark. The "trick" can be rephrased as $(x^2+1)(y+1)(z+1) = 2(x+y)(x+z)$.

After this, straight Cauchy in the obvious way will do it (reducing everything to an inequality in $s = x+y+z$). One writes

$$\begin{aligned}\left(\sum_{\text{cyc}} \frac{\sqrt{(x+y)(x+z)}}{\sqrt{1+s}}\right)^2 &\le \frac{\left(\sum_{\text{cyc}} x+y\right)\left(\sum_{\text{cyc}} x+z\right)}{1+s} \\ &= \frac{4s^2}{1+s}\end{aligned}$$

and so it suffices to check that $\frac{4s^2}{1+s} \le 9(s/3)^{5/4}$, which is true because

$$(s/3)^5 \cdot 9^4 \cdot (1+s)^4 - (4s^2)^4 = s^5(s-3)^2(27s^2+14s+3) \ge 0.$$

Solution 30. For the record, this problem seems to be very difficult, but here's the very nice solution. We'll assume z_k are as small as possible.

The first step is to scale such that

$$\max\{x_1,\dots,x_n\} = \max\{y_1,\dots,y_n\} = 1 \implies M = 1.$$

Here is an example picture, with M and z_k bolded (the x_i are columns, the y_j are rows, hence the diagonals correspond to fixing $i+j$ and the z_k =

$\max_{i+j=k} x_i y_j$).

$M = \mathbf{1}$	0.81	1.00	0.49	0.16
1.00	**0.90**	**1.00**	0.70	0.40
0.36	0.54	0.60	0.42	0.24
0.64	**0.72**	**0.80**	**0.56**	0.32
0.25	0.45	0.50	**0.35**	**0.20**

After this we claim that:

Claim. We have

$$M + z_2 + z_3 + \cdots + z_{2n} \geq x_1 + \cdots + x_n + y_1 + \cdots + y_n$$

In fact, one can bijectively pair each of the $2n$ terms on the right-hand side to a term on the left-hand side exceeding it.

Proof. Enough to prove that for a given $0 \leq r \leq 1$, at least as many terms at least r on the left-hand side compared to the right-hand side. To this end, let

$$I = \{i \mid x_i \geq r\}$$
$$J = \{j \mid y_j \geq r\}.$$

Thus the right-hand side has $|I| + |J|$ terms exceeding r. But the left-hand side has at least $1 + |I + J|$ (the 1 coming from $M = 1$). From the well-known fact that

$$|I + J| \geq |I| + |J| - 1$$

for sets I and J, we are done. □

Solution 31. This problem admits several approaches; here are a few.

First solution (original) By weighted AM-GM we have that

$$1 = \sum_{\text{cyc}} \left(\frac{\sqrt[7]{a}}{a+b+c} \right) = \sum_{\text{cyc}} \left(\frac{a}{a+b+c} \cdot \frac{1}{\sqrt[7]{a^6}} \right) \geq \left(\frac{1}{a^a b^b c^c} \right)^{\frac{6/7}{a+b+c}}.$$

Rearranging yields $a^a b^b c^c \geq 1$.

Second solution (chronodecay) From $e^t \geq 1 + t$ for $t = \log x^{-\frac{6}{7}}$, we find

$$\frac{6}{7} \log x \geq 1 - x^{-\frac{6}{7}}.$$

Thus

$$\frac{6}{7} \sum_{\text{cyc}} a \log a \geq \sum_{\text{cyc}} a - a^{\frac{1}{7}} = 0.$$

---

3 Functional Equations

This chapter is concerned with functional equations, which typically ask you to find all functions satisfying a certain property. For many problems, there is an obvious solution that works, but the main difficulty is to prove that those are all solutions.

§3.1 Definitions

I need to define a function first.

Definition 3.1. Let X and Y be sets. A **function** $f\colon X \to Y$ is an assignment of a value in Y for each $x \in X$; we denote this value $f(x) \in Y$.

§3.1.1 On the generality of functions

Beginners are often surprised how general this definition is. Here are some examples of functions $f\colon \mathbb{R} \to \mathbb{R}$.

$$f(x) = \lfloor x \rfloor$$
$$f(x) = \exp(\sin(x))$$
$$f(x) = \begin{cases} 1 & x \in \mathbb{Z} \\ 0 & x \notin \mathbb{Z} \end{cases}$$
$$f(x) = \begin{cases} 1/q & x = p/q \text{ in lowest terms} \\ 0 & x \notin \mathbb{Q} \end{cases}$$
$$f(x) = \text{number of sloths with age } \leq x.$$

and so on. There's also no restriction on "closed forms": In particular,

- A function need not be a polynomial.
- A function need not be increasing.
- A function need not be continuous.
- A function need not be differentiable.
- The graph of the function need not be well behaved.

Exercise (For experts). Show that there are infinitely many functions which cannot be expressed in LaTeX in any way.

Essentially, anything that can't be proved using manipulations in some way is likely wrong. Any argument that appeals to pictures or graphs for proofs is definitely wrong (helpful as they may be for intuition).

§3.1.2 Special types of functions

In solving functional equations, the following adjectives are convenient.

Definition 3.2. A function $f\colon X \to Y$ is **injective** if it is "one-to-one" in the following sense: if $f(x) = f(x')$ then $x = x'$. In other words, for any $y \in Y$, there is *at most* one $x \in X$ such that $f(x) = y$.

Definition 3.3. A function $f\colon X \to Y$ is **surjective** if it is "onto" in the following sense: for any $y \in Y$ there is *at least* one $x \in X$ such that $f(x) = y$.

Definition 3.4. A function $f\colon X \to Y$ is **bijective** if it is both injective and surjective. In other words, for each $y \in Y$, there is *exactly* one $x \in X$ such that $f(x) = y$.

Here are some examples.

- There's a function from living humans to $\mathbb{Z}_{\geq 0}$ by taking every human to their age in years (rounded to the nearest integer). This function is *not injective*, because for example there are many people with age 20. This function is also *not surjective*: no one has age 10000.
- There's also a function taking every American citizen to their social security number (SSN), which we view as a function from citizens to $\mathbb{Z}_{\geq 0}$. This is also *not surjective* (no one has SSN equal to 3), but at least it *is injective* (no two people have the same SSN).

Here is a common situation in which you get such hypotheses.

Definition 3.5. A function $f\colon X \to X$ is called an **involution** if $f(f(x)) = x$ for every $x \in X$.

Lemma 3.6. *If $f\colon X \to X$ is an involution, then f is a bijection.*

Proof. To see f is injective, note that if $f(a) = f(b)$ then $a = f(f(a)) = f(f(b)) = b$. And f is clearly surjective, since it maps $f(a)$ to a for each a. □

If you have never seen these concepts before, don't worry about them yet; it will become clear with examples why these are useful notions.

§3.2 First example

For concreteness, let me start off with a standard example, that shows a lot of the types of things that often come up in these sorts of problems.

Example 32 (Kyrgyzstan Olympiad 2012)**.** Find all functions $f\colon \mathbb{R} \to \mathbb{R}$ such that

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

for all $x, y \in \mathbb{R}$.

Before I begin solving the problem, I want to make two initial remarks on "finding the answer", which apply to nearly every problem.

Guessing the answer Clearly, $f(x) = +x$ works. But there's actually a second solution: $f(x) = -x$. In general, a "garden-variety" functional equation will have $f(x) = x$ as a solution, but sometimes also $f(x) = 0$, $f(x) = kx$, $f(x) = x + c$, or even $f(x) = kx + c$. So therefore, I recommend **at the start of every problem that you start by seeing which linear functions work**, and to just keep these in your head.

(If it is not too much trouble, try also checking degree n polynomials in general. This is often easier than it seems, since degrees usually end up not matching except for finitely many n.)

For this problem, it looks like $f(x) = \pm x$ is a solution, so we just need to keep in mind that we need to allow for this case.[1]

Verifying it Officially, every functional equation is a *two-directional problem* (as warned in Section 1.1.1.) If we think the answer is $f(x) = \pm x$, then we need to check that these do indeed work, and more importantly prove they are the *only* solutions.

Solution 32. We claim the answers are $f(x) = \pm x$. Obviously they work, so we will now prove they are the only ones.

Well, one can simply start off by plugging stuff in, and grabbing whatever low-hanging fruit we can. Usually, the first thing I try is setting all zeros; this is often helpful, and in general your first attempts should try to make a lot of terms vanish. When we do this here, we get

$$f(f(0)^2 + f(0)) = 0.$$

The inner term is pretty messy, but let me for now just denote it u, i.e. we have some u such that $f(u) = 0$. This is still useful, because we can use it to make things disappear! By plugging in $x = u$ we obtain that

$$f(f(y)) = y$$

and so f is an *involution*; hence a bijection by Lemma 3.6.

Of course, this is not all it gives us. In the given equation, we can now put $x = f(t)$ in order to replace all the $f(x)$'s with $f(f(t)) = t$'s (thus paradoxically we're decreasing the number of nested terms by adding an extra f into the

[1] In general, the set of solutions you find also motivates which claims may be helpful to prove. For example, if $f(x) = x$ and $f(x) = 2 - x$ then you can't hope to prove $f(0) = 0$ or $f(xy) = f(x)f(y)$. But maybe we can get $f(1) = 1$?

given!). This gives us:

$$\begin{aligned} f(f(x)^2 + f(y)) &= xf(x) + y && \text{given} \\ f(f(f(t))^2 + f(y)) &= (f(t))f(f(t)) + y && \text{put } x = f(t) \\ f(t^2 + f(y)) &= f(t) \cdot t + y && \text{since } f(f(t)) = t \\ &= f(f(t)^2 + f(y)) && \text{by given.} \end{aligned}$$

We arrive at the conclusion that

$$f(t^2 + f(y)) = f(f(t)^2 + f(y)).$$

But **since f is injective**, we can now conclude that

$$t^2 + f(y) = f(t)^2 + f(y) \implies f(t)^2 = t^2$$

for every t!

There's still a little more to go, even though this looks like almost what we want — this is the so-called **pointwise trap**. If we are careful, we find that the statement we have proved is

$$f(t) \in \{-t, t\} \text{ for every } t.$$

This is different from our claim that f is one of the two linear functions we noticed! There are *infinitely* many other functions still in contention, like $f(t) = |t|$. The issue is that $f(t)$ might change signs as t varies. (Ankan Bhattacharya has the following to say: if a person is either happy or unhappy at any particular time, does that mean they are always happy or always unhappy?)

So, we need to rule out these unruly functions. This turns out to not be so hard. Suppose that $f(a) = +a$ and $f(b) = -b$ for now, for some nonzero a and b. Substituting these into the given (for x and y) gives that

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

The left-hand side should either equal $a^2 - b$ or $b - a^2$. However these then give $b = 0$ and $a = 0$, respectively. This contradiction completes the proof. ∎

§3.3 Second example (or non-example)

Our second example is a USAJMO problem, for which we begin by presenting solutions that *don't* work, illustrating some of the pitfalls earlier.

Example 33 (USAJMO 2015)**.** Find all functions $f\colon \mathbb{Q} \to \mathbb{Q}$ such that

$$f(x) + f(t) = f(y) + f(z)$$

for all rational numbers $x < y < z < t$ that form an arithmetic progression.

Bogus Solution. Let $f(x) = a_0 + a_1 x + \cdots + a_n x^n$. Then…

This is an instant zero. You can't assume f is a polynomial.

Bogus Solution. Let d be the common difference of the arithmetic progressions; then the given rearranges to

$$f(x+3d) - f(x+2d) = f(x+d) - f(x)$$

so the function has constant slope. Thus it is linear.

This doesn't make sense since one can't talk about slopes of nonlinear functions. You might try to use a derivative instead, but one does not know that f is differentiable. Basically, any solution of this shape is not going to work.

Bogus Solution. Define $g(x) = f(x+1) - f(x)$. Then from the givens,

$$f(x) + f(x+3) = f(x+1) + f(x+2) \implies g(x+2) = g(x)$$

so g is constant. Thus, $f(x+1) - f(x) = c$ for some constant c. Thus f is linear.

There is no reason that $g(x) = g(x+2)$ means g is constant: for example, consider $g(x) = \{x\}$ the fractional part of x. Similarly, there is no reason $f(x+1) - f(x) = c$ implies linear; consider $f(x) = \lfloor x \rfloor$.

Bogus Solution. Note that for positive rational numbers $a, d > 0$ we have

$$\begin{aligned} f(a) + f(a+3d) &= f(a+d) + f(a+2d) \\ f(a-d) + f(a+2d) &= f(a) + f(a+d) \\ \implies f(a-d) + f(a+3d) &= 2f(a+d). \end{aligned}$$

This is enough to imply

$$f(x) + f(y) = 2f\left(\frac{x+y}{2}\right)$$

for distinct rational numbers x and y; but clearly this holds when $x = y$ and so the relation holds whenever $x < y$.

This shows that given two points on the graph of f, the midpoint also lies on the graph. This implies f is linear.

The first paragraph is correct (and the right way to start), but the part about the graph of f doesn't make sense. The graphs of functions can be arbitrarily weird; they don't need to be continuous in any way. In general, I've never heard of any reasonable way to make "graphical" arguments work. It might help to just never try to use them.

We now give a correct solution, but only for contrast with the preceding solution. We present it temporarily, with no motivation, since the correct motivation will come from later results. Thus it may make sense to *not* read the following solution, but merely to look at it and convince yourself that you *could* read it if you wanted to.

Solution 33. Let $d > 0$ be a positive integer, and let n be an integer. Consider the two equations

$$f\left(\frac{2n-1}{2d}\right) + f\left(\frac{2n+2}{2d}\right) = f\left(\frac{2n}{2d}\right) + f\left(\frac{2n+1}{2d}\right)$$
$$f\left(\frac{2n-2}{2d}\right) + f\left(\frac{2n+1}{2d}\right) = f\left(\frac{2n-1}{2d}\right) + f\left(\frac{2n}{2d}\right)$$

Summing them and simplifying implies that

$$f\left(\frac{n-1}{d}\right) + f\left(\frac{n+1}{d}\right) = 2f\left(\frac{n}{d}\right)$$

or equivalently $f\left(\frac{n}{d}\right) - f\left(\frac{n-1}{d}\right) = f\left(\frac{n+1}{d}\right) - f\left(\frac{n}{d}\right)$. This implies that on the set of rational numbers with denominator dividing d, the function f is linear.

In particular, we should have $f\left(\frac{n}{d}\right) = f(0) + \frac{n}{d}(f(1) - f(0))$ since $\frac{n}{d}$, 0, 1 have denominators dividing d. This is the same as saying $f(q) = f(0) + q(f(1) - f(0))$ for any $q \in \mathbb{Q}$, which is what we wanted to prove. ∎

§3.4 Four techniques for motivating substitutions

As you saw from the first example, a lot of functional equations involve extensive substitutions, which can seem almost random at first glance. In fact, many substitutions *are* simply the result of extensive trial and error.

Nonetheless, there is some method to this madness. For example, one heuristic might be to make substitutions which cause any many terms to vanish as possible. In this section we present four less obvious techniques which can help with finding the correct substitutions.

§3.4.1 Forced cancellation

This is best done by example; the following one was invented by David Yang.

Example 34. Find all functions $f\colon \mathbb{R} \to \mathbb{R}$ such that

$$f(x^2 + y) = f(x^{27} + 2y) + f(x^4)$$

for all $x, y \in \mathbb{R}$.

Solution 34. For this problem, we claim the only answer is the constant function $f = 0$, which evidently works. As usual our first move is to take the all-zero setting, which gives $f(0) = 0$.

Now, let's step back: can we do anything that will make lots of terms go away? There's actually a very artificial choice that will do wonders. It is motivated by the following battle cry:

> *"DURR WE WANT STUFF TO CANCEL."*

So we do the most blithely stupid thing possible. *See that* $x^2 + y$ *and* $x^{27} + 2y$ *up there?* Let's make them equal in the rudest way possible:

$$x^2 + y = x^{27} + 2y \iff y = x^2 - x^{27}.$$

Plugging in this choice of y, this gives us $f(x^4) = 0$, so f is zero on all nonnegatives.

All that remains is to get f zero on all reals. The easiest way to do this is put $y = 0$ since this won't hurt the already positive x^2 and x^4 terms there. ■

This is a common trick: see if you can make a substitution that will kill off two terms. We will see this technique in Problem 41.

§3.4.2 The fff trick

The situation $f(f(x)) = x$ is great. However, sometimes we will run into problems where $f(f(x))$ might be something else. In this case, considering $f(f(f(x)))$ in two different ways can often be helpful. Here is an artificial example showing the technique; Problem 41 will give another example.

Example 35. Find all strictly increasing functions $f\colon \mathbb{Z} \to \mathbb{Z}$ such that $f(f(x)) = x + 2$ for all integers x.

Solution 35. The answer is $f(x) = x + 1$ only, which obviously works.

We now consider $f(f(f(x)))$ in two ways. On the other hand, it should be equal to $f(x + 2)$, by replacing *two inner* f's to the statement.

On the other hand, by replacing x with $f(x)$ in the given, we should also have $f(f(f(x))) = f(x) + 2$. In summary,

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

We are now essentially done. Indeed, $f(x) < f(x+1) < f(x+2) = f(x)+2$ and all three expressions must be integers, so this can only occur if $f(x+1) = f(x) + 1$. In other words, $f(x) = x + f(0)$ for all integers x. Finally, checking we find that only $f(x) = x + 1$ works. ■

§3.4.3 Symmetry

If significant parts of your functional equation are symmetric with respect to x and y, then swapping x and y can yield good information. Here is an artificial example.

Example 36. Find all functions $f\colon \mathbb{R} \to \mathbb{R}$ such that

$$xf(x) + y^2 + f(xy) = f(x+y)^2 - f(x)f(y).$$

for all real numbers x and y.

Solution 36. The answer is $f(x) = x$ only, which works.

To prove that is all, we compare the result of swapping x and y:

$$\begin{aligned} xf(x) + y^2 + f(xy) &= f(x+y)^2 - f(x)f(y) \\ yf(y) + x^2 + f(xy) &= f(x+y)^2 - f(x)f(y). \end{aligned}$$

Subtracting the two now gives $xf(x)+y^2 = yf(y)+x^2$, or equivalently $xf(x)-x^2 = yf(y)-y^2$. Thus $xf(x)-x^2$ is equal to some fixed constant c. Taking $x = 0$ we get $c = 0$. Thus $xf(x) = x^2$ and so $f(x) = x$ for every $x \neq 0$.

Finally, put $x = y = 0$ in the original given to get $f(0) = 0$ as well. ∎

§3.4.4 Isolated parts

Sometimes, a variable is "isolated" in such a way that you can read off injectivity or surjectivity.

For injectivity, one can for example try to express y as a function of $f(y)$, no matter how ugly. For example, suppose f is a nonzero function satisfying the ugly condition

$$f(x + 2xf(y)^2) = yf(x) + f(f(y) + 1).$$

Fix any value x_0 such that $f(x_0) \neq 0$. Then one can write

$$y = \frac{f(x_0 + 2x_0 f(y)^2) - f(f(y)+1)}{f(x_0)}.$$

Observe that the right-hand side, as hideous as it is, is determined by the value of $f(y)$. In other words, given the value of $f(y)$ we can find y. This is just saying that f is injective.

Proving surjectivity can often be done in similar spirit. For example, suppose we have an equation like

$$f(f(y) + x^3 f(x)) = y + f(x^2)^2.$$

Again fix some value of x. Then varying y, the right-hand side takes all real values, while the left-hand side is of the form $f(\text{something})$. Thus we conclude right away that f is surjective!

Incidentally, both techniques work well in the given equation of Example 32, so feel free to try it out!

§3.5 Cauchy's equation

One common situations that arises is the functional equation $f(x+y) = f(x)+f(y)$. This entire section is dedicated to the study of this equation.

§3.5.1 Cauchy's equation over $\mathbb{Q}$

We begin with the case where the function is $\mathbb{Q} \to \mathbb{Q}$. We highly encourage the reader to try these examples on their own before reading the solutions; they are good practice problems!

Example 37 (Cauchy's functional equation over $\mathbb{Q}$)**.** Find all functions $f \colon \mathbb{Q} \to \mathbb{Q}$ satisfying

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

for all $x, y \in \mathbb{Q}$.

Solution 37. As before we begin by examining which functions we think the answers are. Trying out the most general $f(x) = kx + c$, we find that $c = 0$ but k can be anything. So our guess is that the answer is $f(x) = kx$.

We now prove this guess is right. First of all, all such functions clearly work.

Now, to prove the reverse, observe we have "one degree of freedom": the family of solutions has a free variable. So it makes sense to set, say, $k = f(1)$ and try to solve everything else in terms of k.

We begin now by setting $x = y = 0$ to derive $f(0) = 0$. Then, we can put $x = 1$, $y = 1$ to get $f(2) = f(1) + f(1) = 2k$. Now, $(x, y) = (2, 1)$ gives $f(3) = 3k$, and so on, so by induction we get $f(n) = kn$ for any integer $n \geq 1$.

What about the negative integers? Well, by putting $x = -y$ we get $f(x) + f(-x) = 0$, and so in fact f is odd. Thus the result $f(n) = kn$ holds for the negative integers.

We're still stuck with the problem of getting all of $\mathbb{Q}$. As a thought experiment, let's see what we can do to get $f(\frac{1}{2})$. We have that

$$f\left(\frac{1}{2}\right) + f\left(\frac{1}{2}\right) = f(1) = k$$

whence $f(\frac{1}{2}) = \frac{1}{2}k$. And now the path is clear for general p/q: we have

$$\underbrace{f(p/q) + \cdots + f(p/q)}_{q} = f(p) = kp$$

and hence $f(p/q) = k \cdot p/q$. Thus, we conclude that $f(x) = kx$ for all x. ∎

Remark 3.7. Notice how the choice of $\mathbb{Q}$ as domain is critical: this all works out because we are able to do induction in order to get the function f over $\mathbb{Z}$ inputs, and then over $\mathbb{Q}$. This fails if $f \colon \mathbb{R} \to \mathbb{R}$, as the next section shows.

In contrast, the choice of codomain is irrelevant, we run into no problem if we repeat this proof for $f \colon \mathbb{Q} \to \mathbb{R}$.

Example 38 (Jensen's functional equation over $\mathbb{Q}$)**.** Find all functions $f\colon \mathbb{Q} \to \mathbb{Q}$ satisfying

$$f(x) + f(y) = 2f\left(\frac{x+y}{2}\right).$$

for all $x, y \in \mathbb{Q}$.

Solution 38. This time, our preliminary checks reveal that $f(x) = kx + c$ works for any k and c.[2] We prove these are the only solutions.

So now we do the following trick: we can shift the function f by c without changing the function. To be clear, this means that we rewrite the given as

$$(f(x) - f(0)) + (f(y) - f(0)) = 2\left(f\left(\frac{x+y}{2}\right) - f(0)\right).$$

If we now let $g(x) = f(x) - f(0)$, then we derive

$$g(x) + g(y) = 2g\left(\frac{x+y}{2}\right)$$

so this is the same functional equation; but now, we know $g(0) = 0$.

So, setting $(x, y) = (t, 0)$ gives $g(t) = 2g(t/2)$. We might try the same trick as before with Cauchy, say setting $(x, y) = (1, 2)$ to get

$$g(1) + g(2) = 2g(3/2)$$

which seems non-useful until we remember we have $g(t) = 2g(t/2)$. Indeed, the given functional equation can be rewritten as

$$g(x) + g(y) = 2g\left(\frac{x+y}{2}\right) = g\left(x+y\right)$$

with $t = x + y$. So g is Cauchy!

Therefore, g must be linear, and so f must be linear too. ■

We now remark that the earlier equation

$$f\left(\frac{n-1}{d}\right) + f\left(\frac{n+1}{d}\right) = 2f\left(\frac{n}{d}\right)$$

which we used in our solution to USAJMO 2015/4 is more or less the same as the preceding example, if we drop the constraint that n and d are integers (and allow them to be any positive rational numbers).

(Sidenote: Beginners should *not* worry about remembering the name or statement of Example 38; the name is only included for completeness.)

[2] In a vague sense, the fact that c is free to vary is manifested in the fact that plugging in all zeros yields the tautology $0 = 0$.

§3.5.2 Cauchy's equation over $\mathbb{R}$

As I alluded to earlier, the situation becomes very different if we replace $\mathbb{Q}$ by $\mathbb{R}$, since induction is no longer valid. Actually, over $\mathbb{R}$, we get new pathological (or just "bad") solutions to Cauchy's equation that weren't there before. Such functions are discussed more carefully in the next chapter.

In general, however, if you end up with Cauchy's Functional Equation, then often a judicious use of some other known equation will work. The relation $f(x+y) = f(x)+f(y)$ is very powerful, and usually just using the multiplicative structure a little bit will get you what you need.

One common criteria is the following theorem (which we will not prove).

Theorem 3.8 (Cauchy + Continuous $\implies$ Linear)**.** *Suppose $f : \mathbb{R} \to \mathbb{R}$ satisfies $f(x+y) = f(x) + f(y)$. Then $f(qx) = qf(x)$ for any $q \in \mathbb{Q}$.*

Moreover, f is linear if any of the following are true:

- *f is continuous in any interval.*
- *f is bounded (either above or below) in any nontrivial interval.*
- *There exists (a, b) and $\varepsilon > 0$ such that $(x-a)^2 + (f(x)-b)^2 > \varepsilon$ for every x (i.e. the graph of f omits some disk, however small).*

Here's an example of how it can be used.

Example 39. Find all functions $f\colon \mathbb{R} \to \mathbb{R}$ such that for any $x, y \in \mathbb{R}$, we have both $f(x+y) = f(x) + f(y)$ and $f(xy) = f(x)f(y)$.

Solution 39. We claim $f(x) = x$ and $f(x) = 0$ are the only solutions (which both work). According to the theorem, to prove f is linear it suffices to show f is nonnegative over some nontrivial interval. Now,

$$f(t^2) = f(t)^2 \geq 0$$

for any t, meaning f is bounded below on $[0, \infty)$ and so we conclude $f(x) = cx$ for some c. Then $cxy = (cx)(cy)$ implies $c \in \{0, 1\}$, as claimed. ■

In general, as far as olympiad contexts, the most common ways to get from additive to linear are:

- Being able to prove bounded conditions (such as $f \geq 0$), or
- The problem *gives* you that the function f is continuous[3], inviting you to quote the above theorem.

[3]It is extremely rare that you need to prove continuity yourself; in fact I personally cannot think of any examples off-hand.

§3.6 Walkthroughs

Problem 40 (USAMO 2002)**.** Determine all functions $f : \mathbb{R} \to \mathbb{R}$ such that

$$f(x^2 - y^2) = xf(x) - yf(y)$$

for all pairs of real numbers x and y.

Walkthrough. This is a classic example of getting down to a Cauchy equation, and then pushing just a little harder.

(a) Find all linear solutions and show there are no higher-degree polynomial ones.

(b) Show that f is odd and hence $f(0) = 0$.

(c) Show that f is additive and $f(x^2) = xf(x)$.

(d) Optionally: prove that the problem statement is *equivalent* to the relations in (c). Hence we can more or less ignore the given equation now.

(e) Prove that f is linear, by inserting $x = a + b$ into $f(x^2) = xf(x)$.

Problem 41 (IMO 2017)**.** Solve over $\mathbb{R}$ the functional equation

$$f\left(f(x)f(y)\right) + f(x + y) = f(xy).$$

Walkthrough. This problem is sort of divided into two parts. One is the "standard" part, which is not easy *per se*, but which experienced contestants won't find surprising. However, the argument in the final part is quite nice and conceptual, and much less run-of-the-mill.

We begin with some standard plug/chug.

(a) Find all three linear solutions and convince yourself there are no other polynomial solutions.

(b) Check that if f is a solution, then so is $-f$.

(c) Show there exists z such that $f(z) = 0$. (We'll find the exact value later; for now just show it exists.)

(d) Show that if $f(0) = 0$ then $f \equiv 0$. So we henceforth assume $f(0) \neq 0$.

(e) Using the cancellation trick, prove that if $f(z) = 0$ (and $f(0) \neq 0$) for some z, then $z = 1$. Then show that $f(0) = \pm 1$.

From (b) and (e), we assume $f(0) = 1$, $f(1) = 0$ in what follows, and will try to show $f(x) \equiv 1 - x$. This lets us plug in some more stuff.

(f) Show that $f(x + 1) = f(x) - 1$ and compute f on all integer values.

(g) Show that $f(f(x)) = 1 - f(x)$. Thus if f was surjective we would be done. However, this seems hard to arrange, since the original equation has everything wrapped in f's.

(h) Using the triple involution trick, prove that $f(1 - f(x)) = f(x)$. Thus if f was injective, we would also be done.

So we will now prove f is injective: this is the nice part. Assume $f(a) = f(b)$; we will try to prove $a = b$.

(i) Show that if N is a sufficiently large integer, then we can find x and y such that $x + y = a + N$ and $xy = b + N$. Use this to prove that $f(f(x)f(y)) = 0$ for that pair (x, y) and hence thus $f(x)f(y) = 1$.

(j) The previous part shows us how we might think about using the cancellation trick. However, it is basically useless since $f(x)f(y) = 1$ is not really a useful condition.

However, modify the approach of (i) so that instead the conclusion ends up as $f(x)f(y) = 0$ instead. Deduce that $1 \in \{x, y\}$ in that case.

(k) Using the argument in (j) prove that $a = b$.

Some historical lore about this problem: this was shortlisted as A6, and in my opinion too hard for the P2 position, despite being nice for a functional equation. Most countries did poorly, with USA and China having only two solves, but the Korean team had an incredibly high five solves. However, an unreasonably generous 4 points was awarded for progress up to part (h), thus cancelling a lot of the advantage from the Korean team. Thus I was relieved that the Korean team still finished first.

Problem 42 (USAMO 2018)**.** Find all functions $f\colon (0,\infty) \to (0,\infty)$ such that

$$f\left(x + \frac{1}{y}\right) + f\left(y + \frac{1}{z}\right) + f\left(z + \frac{1}{x}\right) = 1$$

for all $x, y, z > 0$ with $xyz = 1$.

Walkthrough. This is long and technical, but easier than its length might make it appear.

(a) Eliminate the condition $xyz = 1$ by writing $x = a/b$ and so on. (This shouldn't involve any cube roots or high degrees. You'll see $\frac{a+b}{c}$ appear if you do this right.)

(b) Guess a nonconstant solution after this substitution.

(c) Using your answer to (b), find a family of solutions.

(d) By making substitutions, reduce the problem to solving the functional equation
$$g(a) + g(b) + g(c) = 1 \qquad \forall a + b + c = 1$$
for $g \colon (0,1) \to (0,1)$.

At this point, your intuition should be that this feels like Jensen's functional equation, and a priori any additive function should work. The good news is that this is where the condition $g \geq 0$ gets used: you also know that any additive function which is bounded works. So, we'll need a carefully choreographed ballet of manipulations in order to get to the point we want.

(e) Show that g satisfies Jensen's functional equation over the interval $(0, 1/2)$.

We define $h \colon [0,1] \to \mathbb{R}$ by
$$h(t) = g\left(\frac{2t+1}{8}\right) - (1-t)g(1/8) - tg(3/8).$$
This function mimics g across $[1/8, 3/8]$.

(f) Show that $h(0) = h(1) = h(1/2) = 0$.

(g) Prove that h can be extended "modulo 1" to a function $\widetilde{h} \colon \mathbb{R} \to \mathbb{R}$. (You may need $h(1/2) = 0$ for this.)

(h) Prove that $\widetilde{h}$ satisfies Jensen's functional equation over all of $\mathbb{R}$ and deduce that $\widetilde{h}$ is additive.

(i) Conclude that h is zero everywhere, and hence g is linear over $[1/8, 3/8]$.

We write $g(x) = kx + \ell$ for $x \in [1/8, 3/8]$ where k, ℓ are constants. Now to carry this back:

(j) Use (e) to prove that $g(x) = kx + \ell$ over $[0, 1/8]$.

(k) Prove that $k + 3\ell = 1$.

(l) Use (d) to show that $g(x) = kx + \ell$ over $[3/8, 1]$.

(m) Find the range of acceptable values of k, and write down the final answer.

§3.7 Problems

Problem 43 (IMO 2008)**.** Find all functions f from the positive reals to the positive reals such that
$$\frac{f(w)^2 + f(x)^2}{f(y^2) + f(z^2)} = \frac{w^2 + x^2}{y^2 + z^2}$$
for all positive real numbers w, x, y, z satisfying $wx = yz$.

Problem 44 (IMO 2010). Find all functions $f\colon \mathbb{R} \to \mathbb{R}$ such that for all $x, y \in \mathbb{R}$,

$$f(\lfloor x \rfloor y) = f(x) \lfloor f(y) \rfloor .$$

Problem 45 (IMO 2009). Find all functions $f\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ such that for positive integers a and b, the numbers

$$a, \qquad f(b), \qquad f(b + f(a) - 1)$$

are the sides of a non-degenerate triangle.

Problem 46 (USAMO 2000). Call a real-valued function f *very convex* if

$$\frac{f(x) + f(y)}{2} \geq f\left(\frac{x+y}{2}\right) + |x - y|$$

holds for all real numbers x and y. Prove that no very convex function exists.

Problem 47 (IMO Shortlist 2015). Determine all functions $f\colon \mathbb{Z} \to \mathbb{Z}$ with the property that

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

holds for all $x, y \in \mathbb{Z}$.

Problem 48 (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 49 (IMO Shortlist 2016). Find all functions $f : (0, \infty) \to (0, \infty)$ such that for any $x, y \in (0, \infty)$,

$$xf(x^2)f(f(y)) + f(yf(x)) = f(xy)\left(f(f(x^2)) + f(f(y^2))\right).$$

Problem 50 (ELMO Shortlist 2013). Find all $f\colon \mathbb{R} \to \mathbb{R}$ such that for all $x, y \in \mathbb{R}$,

$$f(x) + f(y) = f(x + y) \qquad \text{and} \qquad f(x^{2013}) = f(x)^{2013}.$$

Problem 51 (TSTST 2013). Let $\mathbb{N}$ be the set of positive integers. Find all functions $f : \mathbb{N} \to \mathbb{N}$ that satisfy the equation

$$f^{abc-a}(abc) + f^{abc-b}(abc) + f^{abc-c}(abc) = a + b + c$$

for all $a, b, c \geq 2$. (Here f^k means f applied k times.)

§3.8 Solutions

Solution 40. The answer is $f(x) = cx$, $c \in \mathbb{R}$ (these obviously work).

First, by putting $x = 0$ and $y = 0$ respectively we have

$$f(x^2) = xf(x) \quad \text{and} \quad f(-y^2) = -yf(y).$$

From this we deduce that f is odd, in particular $f(0) = 0$. Then, we can rewrite the given as $f(x^2 - y^2) + f(y^2) = f(x^2)$. Combined with the fact that f is odd, we deduce that f is additive (i.e. $f(a + b) = f(a) + f(b)$).

Remark (Philosophy)**.** At this point we have $f(x^2) \equiv xf(x)$ and f additive, and everything we have including the given equation is a direct corollary of these two. So it makes sense to only focus on these two conditions.

Then

$$\begin{aligned} & f((x+1)^2) = (x+1)f(x+1) \\ \implies & f(x^2) + 2f(x) + f(1) = (x+1)f(x) + (x+1)f(1) \end{aligned}$$

which readily gives $f(x) = f(1)x$.

Solution 41. The only solutions are $f(x) = 0$, $f(x) = x - 1$ and $f(x) = 1 - x$, which clearly work.

Note that

- If f is a solution, so is $-f$.
- Moreover, if $f(0) = 0$ then setting $y = 0$ gives $f \equiv 0$. So henceforth we assume $f(0) > 0$.

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

Proof. For the forwards direction, if $f(z) = 0$ and $z \neq 1$ one may put $(x, y) = \left(z, z(z-1)^{-1}\right)$ (so that $x + y = xy$) we deduce $f(0) = 0$ which is a contradiction.

For the reverse, $f(f(0)^2) = 0$ by setting $x = y = 0$, and use the previous part. We also conclude $f(1) = 0$, $f(0) = 1$. □

Claim. If f is injective, we are done.

Proof. Setting $y = 0$ in the original equation gives $f(f(x)) = 1 - f(x)$. We apply this three times on the expression $f^3(x)$:

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

Hence $1 - f(x) = x$ or $f(x) = 1 - x$. □

Remark. The result $f(f(x)) + f(x) = 1$ also implies that surjectivity would solve the problem.

Claim. f is injective.

Proof. Setting $y = 1$ in the original equation gives $f(x+1) = f(x) - 1$, and by induction

$$f(x+n) = f(x) - n. \tag{3.1}$$

Assume now $f(a) = f(b)$. By using (3.1) we may shift a and b to be large enough that we may find x and y obeying $x+y = a+1$, $xy = b$. Setting these gives

$$\begin{aligned} f(f(x)f(y)) &= f(xy) - f(x+y) = f(b) - f(a+1) \\ &= f(b) + 1 - f(a) = 1 \end{aligned}$$

from which we conclude

$$f\left(f(x)f(y) + 1\right) = 0.$$

Hence by the first claim we have $f(x)f(y) + 1 = 1$, so $f(x)f(y) = 0$. Applying the first claim again gives $1 \in \{x, y\}$. But that implies $a = b$. □

Remark. Jessica Wan points out that for any $a \neq b$, at least one of $a^2 > 4(b-1)$ and $b^2 > 4(a-1)$ is true. So shifting via (3.1) is actually unnecessary for this proof.

Remark. One can solve the problem over $\mathbb{Q}$ using only (3.1) and the easy parts. Indeed, that already implies $f(n) = 1 - n$ for all n. Now we induct to show $f(p/q) = 1 - p/q$ for all $0 < p < q$ (on q). By choosing $x = 1 + p/q$, $y = 1 + q/p$, we cause $xy = x + y$, and hence $0 = f\left(f(1+p/q)f(1+q/p)\right)$ or $1 = f(1+p/q)f(1+q/p)$.

By induction we compute $f(1+q/p)$ and this gives $f(p/q+1) = f(p/q) - 1$.

Solution 42. The main part of the problem is to show all solutions are linear. As always, let $x = b/c$, $y = c/a$, $z = a/b$ (classical inequality trick). Then the problem becomes

$$\sum_{\text{cyc}} f\left(\frac{b+c}{a}\right) = 1.$$

Let $f(t) = g(\frac{1}{t+1})$, equivalently $g(s) = f(1/s - 1)$. Thus $g\colon (0,1) \to (0,1)$ which satisfies $\sum_{\text{cyc}} g\left(\frac{a}{a+b+c}\right) = 1$, or equivalently

$$\boxed{g(a) + g(b) + g(c) = 1} \qquad \forall a + b + c = 1.$$

The rest of the solution is dedicated to solving this equivalent functional equation in g. It is a lot of technical details and I will only outline them (with apologies to the contestants who didn't have that luxury).

Claim. The function g is linear.

Proof. This takes several steps, all of which are technical. We begin by proving g is linear over $[1/8, 3/8]$.

- First, whenever $a+b \le 1$ we have
$$1 - g(1-(a+b)) = g(a) + g(b) = 2g\left(\frac{a+b}{2}\right).$$
Hence g obeys Jensen's functional equation over $(0, 1/2)$.
- Define $h\colon [0,1] \to \mathbb{R}$ by $h(t) = g(\frac{2t+1}{8}) - (1-t)\cdot g(1/8) - t \cdot g(3/8)$, then h satisfies Jensen's functional equation too over $[0,1]$. We have also arranged that $h(0) = h(1) = 0$, hence $h(1/2) = 0$ as well.
- Since
$$h(t) = h(t) + h(1/2) = 2h(t/2 + 1/4) = h(t+1/2) + h(0) = h(t+1/2)$$
for any $t < 1/2$, we find h is periodic modulo $1/2$. It follows one can extend $\widetilde{h}$ by
$$\widetilde{h}\colon \mathbb{R} \to \mathbb{R} \qquad \text{by} \qquad \widetilde{h}(t) = h(t - \lfloor t \rfloor)$$
and still satisfy Jensen's functional equation. Because $\widetilde{h}(0) = 0$, it's well-known this implies $\widetilde{h}$ is additive (because $\widetilde{h}(x+y) = 2\widetilde{h}\left((x+y)/2\right) = \widetilde{h}(x) + \widetilde{h}(y)$ for any real numbers x to y).

But $\widetilde{h}$ is bounded below on $[0,1]$ since $g \ge 0$, and since $\widetilde{h}$ is also additive, it follows (well-known) that $\widetilde{h}$ is linear. Thus h is the zero function. So, the function g is linear over $[1/8, 3/8]$; thus we may write $g(x) = kx + \ell$, valid for $1/8 \le x \le 3/8$.

Since $3g(1/3) = 1$, it follows $k + 3\ell = 1$.

For $0 < x < 1/8$ we have $g(x) = 2g(0.15) - g(0.3 - x) = 2(0.15k + \ell) - (k(0.3-x) + \ell) = kx + \ell$, so g is linear over $(0, 3/8)$ as well. Finally, for $3/8 < x < 1$, we use the given equation

$$1 = g\left(\frac{1-x}{2}\right) + g\left(\frac{1-x}{2}\right) + g(x) \implies g(x) = 1 - 2\left(k \cdot \frac{1-x}{2} + \ell\right) = kx + \ell$$

since $\frac{1-x}{2} < \frac{5}{16} < \frac{3}{8}$. Thus g is linear over all. □

Putting this back in, we deduce that $g(x) = kx + \frac{1-k}{3}$ for some $k \in [-1/2, 1]$, and so

$$f(x) = \frac{k}{x+1} + \frac{1-k}{3}$$

for some $k \in [-1/2, 1]$. All such functions work.

Solution 43. The answers are $f(x) \equiv x$ and $f(x) \equiv 1/x$. These work, so we show they are the only ones.

First, setting (t, t, t, t) gives $f(t^2) = f(t)^2$. In particular, $f(1) = 1$. Next, setting $(t, 1, \sqrt{t}, \sqrt{t})$ gives

$$\frac{f(t)^2 + 1}{2f(t)} = \frac{t^2 + 1}{2t}$$

which as a quadratic implies $f(t) \in \{t, 1/t\}$.

Now assume $f(a) = a$ and $f(b) = 1/b$. Setting $(\sqrt{a}, \sqrt{b}, 1, \sqrt{ab})$ gives

$$\frac{a + 1/b}{f(ab) + 1} = \frac{a + b}{ab + 1}.$$

One can check the two cases on $f(ab)$ each imply $a = 1$ and $b = 1$ respectively. Hence the only answers are those claimed.

Solution 44. The only solutions are $f(x) \equiv c$, where $c = 0$ or $1 \le c < 2$. It's easy to see these work.

Plug in $x = 0$ to get $f(0) = f(0) \lfloor f(y) \rfloor$, so either

$$1 \le f(y) < 2 \quad \forall y \qquad \text{or} \qquad f(0) = 0$$

In the first situation, plug in $y = 0$ to get $f(x) \lfloor f(0) \rfloor = f(0)$, thus f is constant. Thus assume henceforth $f(0) = 0$.

Now set $x = y = 1$ to get

$$f(1) = f(1) \lfloor f(1) \rfloor$$

so either $f(1) = 0$ or $1 \le f(1) < 2$. We split into cases:

- If $f(1) = 0$, pick $x = 1$ to get $f(y) \equiv 0$.
- If $1 \le f(1) < 2$, then $y = 1$ gives

  $$f(\lfloor x \rfloor) = f(x)$$

  from $y = 1$, in particular $f(x) = 0$ for $0 \le x < 1$. Choose $(x, y) = \left(2, \frac{1}{2}\right)$ to get $f(1) = f(2) \left\lfloor f\left(\frac{1}{2}\right) \right\rfloor = 0$.

Solution 45. The only function is the identity function (which works). We prove it is the only one.

Let $P(a, b)$ denote the given statement.

Claim. We have $f(1) = 1$, and $f(f(n)) = n$. (In particular f is a bijection.)

Proof. Note that

$$P(1,b) \implies f(b) = f(b+f(1)-1).$$

Otherwise, the function f is periodic modulo $N = f(1) - 1 \geq 1$. This is impossible since we can fix b and let a be arbitrarily large in some residue class modulo N.

Hence $f(1) = 1$, so taking $P(n,1)$ gives $f(f(n)) = n$. □

Claim. Let $\delta = f(2) - 1 > 0$. Then for every n,

$$f(n+1) = f(n) + \delta \quad \text{or} \quad f(n-1) = f(n) + \delta$$

Proof. Use

$$P(2, f(n)) \implies n - 2 < f(f(n) + \delta) < n + 2.$$

Let $y = f(f(n) + \delta)$, hence $n - 2 < y < n + 2$ and $f(y) = f(n) + \delta$. But, remark that if $y = n$, we get $\delta = 0$, contradiction. So $y \in \{n+1, n-1\}$ and that is all. □

We now show f is an arithmetic progression with common difference $+\delta$. Indeed we already know $f(1) = 1$ and $f(2) = 1 + \delta$. Now suppose $f(1) = 1$, ..., $f(n) = 1 + (n-1)\delta$. Then by induction for any $n \geq 2$, the second case can't hold, so we have $f(n+1) = f(n) + \delta$, as desired.

Combined with $f(f(n)) = n$, we recover that f is the identity.

Solution 46. For $C \geq 0$, we say a function f is *C-convex*

$$\frac{f(x) + f(y)}{2} \geq f\left(\frac{x+y}{2}\right) + C\,|x - y|\,.$$

Suppose f is C-convex. Let $a < b < c < d < e$ be any arithmetic progression, such that $t = |e - a|$. Observe that

$$f(a) + f(c) \geq 2f(b) + C \cdot \frac{1}{2}t$$
$$f(c) + f(e) \geq 2f(d) + C \cdot \frac{1}{2}t$$
$$f(b) + f(d) \geq 2f(c) + C \cdot \frac{1}{2}t$$

Adding the first two to twice the third gives

$$f(a) + f(e) \geq 2f(c) + 2C \cdot t.$$

So we conclude C-convex function is also $2C$-convex. This is clearly not okay for $C > 0$.

Solution 47. The answer is $f(x) \equiv x + 1$ and $f(x) \equiv -1$, which both work.

Claim. There exists a with $f(a) = -1$, and thus we have $f(x+1) = f(f(x))$ for all x.

Proof. By setting $(x, y) = (1000, f(1000))$ we see -1 is in the range of f, and choosing y with $f(y) = -1$ gives the latter claim. □

We now outline two approaches.

First approach using images In what follows, let

$$S \coloneqq \{f(y) + 1 \mid y \in \mathbb{Z}\}.$$

Note $0 \in S$.

Now replacing $f(f(x))$ with $f(x+1)$ in the given, then using the S notation, we obtain

$$f(x + s) = f(x) + s$$

for all $s \in S$. Thus S is closed under addition/subtraction, meaning $S = n\mathbb{Z}$ for some $n \geq 0$ (namely $n = \gcd S$).

We now consider three cases.

- First, if $n = 0$ then $f \equiv -1$.
- If $n = 1$, then $f(x) \equiv x + 1$.
- Finally we contend $n > 1$ is impossible. Indeed, this means that $f(x) \equiv -1 \pmod{n}$ for all x. Thus we may select $a \equiv 0 \pmod{n}$ and $b \equiv 1 \pmod{n}$ such that $f(a) = f(b)$ whence

  $$f(a + 1) = f(f(a)) = f(f(b)) = f(b + 1).$$

  Then continuing we find that $f(a + 2) = f(b + 2)$ and so on, so for sufficiently large x, we have $f(x) = f(x + d)$ where $d = |b - a| \equiv \pm 1 \pmod{n}$. Then $f(x) = f(x + nd) = f(x) + nd$ which is a contradiction.

Second approach using direct substitution Taking $P(f(t) - 1, t)$ gives

$$\begin{aligned} f(-1) + 1 &= f(f(f(t) - 1)) - f(t) \\ &= f([f(t) - 1] + 1) - f(t) = f(f(t)) - f(t) \\ &= f(t + 1) - f(t). \end{aligned}$$

Here we have used $f(x + 1) = f(f(x))$ from the Claim twice. Thus f is linear and it is not hard to see that the claimed solutions are the only linear ones.

Solution 48. Let a, b, c denote the values $f(0)$, $g(0)$ and $h(0)$. Notice that by putting $y = 0$, we can get that

$$\begin{aligned} f(x+a) &= g(x) + c \\ g(x+b) &= h(x) + a \\ h(x+c) &= f(x) + b. \end{aligned}$$

Thus the given equation may be rewritten in the form

$$f(x + f(y)) = [f(x+a) - c] + [f(y-c) + b].$$

At this point, we may set $x = y - c - f(y)$ and cancel the resulting equal terms to obtain

$$c - b = f(y + a - c - f(y)).$$

Since f is injective, this implies that $y + a - c - f(y)$ is constant, so that $y - f(y)$ is constant. Thus, f is linear, and $f(y) \equiv y + a$. Similarly, $g(x) \equiv x + b$ and $h(x) \equiv x + c$.

Finally, we just need to notice that upon placing $x = y = 0$ in all the equations, we get $2a = b + c$, $2b = c + a$ and $2c = a + b$, whence $a = b = c$.

So, the family of solutions is $f(x) = g(x) = h(x) = x + c$, where c is an arbitrary real. One can easily verify these solutions are valid.

Solution 49. The answer is $f(x) = 1/x$ only which works.

First start in bughouse mode, and grab the low-hanging fruit. Letting $P(x, y)$ be as usual; noting the left-hand side is asymmetric but the right-hand side is symmetric, we will usually consider $P(x, y)$ in tandem with $P(y, x)$ (so-called "symmetry tricky").

- Put $P(1, 1)$ to conclude $\boxed{f(1) = 1}$.
- Put $P(x, 1)$ and $P(1, x)$ to get

  $$xf(x^2) + f(f(x)) = f(f(x)) + f(x) = f(x)\left[1 + f(f(x^2))\right].$$

  Comparing the left and middle, and the middle and right gives two important corollaries:

  $$\boxed{\frac{f(x)}{x} = f(x^2)} \quad \text{and} \quad \boxed{\frac{f(f(x))}{f(x)} = f(f(x^2))}. \qquad (\heartsuit)$$

  (The second could also be motivated by the cancellation trick.)

In fact ($\heartsuit$) is enough to get the following.

Claim. If f is injective, we are done.

Proof. Using (♡) thrice, we evaluate $f(f(x)^2)$ in two ways:

$$f\left(\frac{f(x)}{x}\right) = f(f(x^2)) = \frac{f(f(x))}{f(x)} = f(f(x)^2)$$

so if f is injective we could conclude $f(x)/x = f(x)^2$, meaning $f(x) = 1/x$. □

Motivated by our quest for injectivity we now use (♡) to **eliminate all squares** which gives

$$f(x)f(f(y)) + f(yf(x)) = f(xy)\left(\frac{f(f(x))}{f(x)} + \frac{f(f(y))}{f(y)}\right). \qquad (♠)$$

We are now ready to prove:

Claim. f is injective.

Proof. Assume $c = f(a) = f(b)$. First, taking (a, b) and (b, a) in (♠) gives

$$f(bf(b)) = f(bc) \overset{(♠)}{=} (\text{function of } c) \overset{(♠)}{=} f(ac) = f(af(a)).$$

We remember just $f(af(a)) = f(bf(b))$. Using this, in (♠) we put (a, a) and (b, b) and compare the two; we find that $f(a^2) = f(b^2)$.

But $f(a^2) = f(a)/a$ and $f(b^2) = f(b)/b$. So $a = b$. □

Solution 50. The answer is $f(x) \equiv +x$, $f(x) \equiv -x$, and $f(x) \equiv 0$.

First slick approach (Daniel Xia) By scaling, assume $f(1) = 1$ (if $f(1) = 0$, the proof if similar). Notice that

$$f\left((1+x)^{2013} + (1-x)^{2013}\right) = (f(1) + f(x))^{2013} + (f(1) - f(x))^{2013} = P(f(x))$$

where $P(t) = (1+t)^{2013} + (1-t)^{2013}$ is a polynomial of degree 2012 (the leading term $4026t^{2012}$). Since P has even degree, it is bounded below. So f is bounded on a nontrivial interval (actually the image of P), hence linear.

Second approach (official solution) Let $n = 2013$ be odd. Set $c = f(1)$, so $f(q) = cq$ for any rational number q.

We know $f(q) = cq$ for any $q \in \mathbb{Q}$. Thus for any $x \in \mathbb{R}$ and $q \in \mathbb{Q}$ we have

$$\sum_k \binom{n}{k} q^k f(x^{n-k}) = f\left((x+q)^n\right) = f(x+q)^n = \sum_k \binom{n}{k} f(x)^{n-k}(qc)^k.$$

For any particular x, both left and right hand side are *polynomials in* q, so the coefficients must agree for each x.

Now matching q^1 terms, $f(x^{n-1}) = cf(x)^{n-1}$ for all x. Consequently, f has fixed sign over nonnegative reals, and is thus linear; this concludes the proof.

Remark. Over $\mathbb{C}$ this problem is false: actually, there exist so called "wild automorphisms" of the complex numbers, i.e. functions that are both additive and multiplicative.

Solution 51. The answer is $f(n) = n - 1$ for $n \geq 3$ with $f(1)$ and $f(2)$ arbitrary; check these work.

Lemma. *We have $f^{t^2-t}(t^2) = t$ for all t.*

Proof. We say $1 \leq k \leq 8$ is good if $f^{t^9-t^k}(t^9) = t^k$ for all t. First, we observe that

$$f^{t^9-t^3}(t^9) = t^3 \quad \text{and} \quad f^{t^3-t}(t^3) = t \implies f^{t^9-t}(t^9) = t.$$

so $k = 1$ and $k = 3$ are good. Then taking $(a, b, c) = (t, t^4, t^4)$, $(a, b, c) = (t^2, t^3, t^4)$ gives that $k = 4$ and $k = 2$ are good, respectively. The lemma follows from this $k = 1$ and $k = 2$ being good. □

Now, letting $t = abc$ we combine

$$\begin{aligned} f^{t-a}(a) + f^{t-b}(b) + f^{t-c}(c) &= a + b + c \\ f^{t^2-ab}(t^2) + f^{t^2-t}(t^2) + f^{t^2-c}(t^2) &= ab + t + c \\ \implies \left[f^{t-a}(t) - a\right] + \left[f^{t-b}(t) - b\right] &= \left[f^{t-ab}(t) - ab\right] \end{aligned}$$

by subtracting and applying the lemma repeatedly. In other words, we have proven the second lemma:

Lemma. *Let t be fixed, and define $g_t(n) = f^{t-n}(t) - n$ for $n < t$. If $a, b \geq 2$ and $ab \mid t$, $ab < t$, then $g_t(a) + g_t(b) = g_t(ab)$.*

Now let $a, b \geq 2$ be arbitrary, and let $p > q > \max\{a, b\}$ be primes. Suppose $s = a^p b^q$ and $t = s^2$; then

$$pg_t(a) + qg_t(b) = g_t\left(a^p b^q\right) = g_t(s) = f^{s^2-s}(s) - s = 0.$$

Now

$$q \mid g_t(a) > -a \quad \text{and} \quad p \mid g_t(b) > -b \implies g_t(a) = g_t(b) = 0.$$

and so we conclude $f^{t-a}(t) = a$ and $f^{t-b}(t) = b$ for $a, b \geq 2$.

In particular, if $a = n$ and $b = n + 1$ then we deduce $f(n + 1) = n$ for all $n \geq 2$, as desired.

Remark. If you let $c = (ab)^2$ after the first lemma, you recover the 2-variable version!

---

4 Monstrous Functional Equations

§4.1 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 Cauchy equation

$$f \colon \mathbb{R} \to \mathbb{R} \qquad f(x+y) = f(x) + f(y).$$

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!

§4.2 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 $f \colon \mathbb{Z} \to \mathbb{Z}$ and kind of find a way to relate $f(n+1)$ to $f(n)$, but there is still some degree of freedom left. The first walkthrough Problem 53 will give you some practice with this.
- **You have some values of f down, but others seem out of reach**. This usually happens when $f \colon \mathbb{R} \to \mathbb{R}$. 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 Problem 54.
- **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 $f(x) = x$.

  As suggested last chapter, it may be worth it to **begin by checking for all linear or polynomial 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 $f : \mathbb{Z} \to \mathbb{Z}$ 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 $f : \mathbb{R} \to \mathbb{R}$ 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.

§4.3 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 exact axioms[1] (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 ≤ 2, that is,
$$W = \left\{ ax^2 + bx + c \mid a, b, c \in \mathbb{R} \right\},$$
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
$$v = t_1 e_1 + t_2 e_2 + \cdots + t_m e_m \tag{4.1}$$
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:

[1] Reminder for experts: it's an abelian group under addition with a 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 (4.1) 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.

§4.4 Cauchy's equation over $\mathbb{R}$

We are now ready to address.

Example 52 (Cauchy's functional equation over $\mathbb{R}$)**.** Describe all functions $f\colon \mathbb{R} \to \mathbb{R}$ satisfying $f(x+y) = f(x) + f(y)$.

Solution 52. 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

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

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 $f : \mathbb{R} \to \mathbb{R}$, I asked for solutions of $f : \mathbb{Q}[\sqrt{2}] \to \mathbb{R}$? You might at this point be able to guess a solution:

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

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

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

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*.

For us, we will just state the technical result. If you're interested in the details, see the chapter on Zorn's Lemma in [Che19]. For this handout I will just state the result.

Proposition 4.5 (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**.*

As a metaphor, you can almost think of this as saying something like

$$\mathbb{R} = \left\{ a_1 + \sqrt{2}a_2 + \sqrt{3}a_3 + a_4\pi + a_5 e + \cdots \mid a_1, \cdots \in \mathbb{Q} \right\}.$$

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 α. Then, declare

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

Those of you very 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 (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, ….)

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.

§4.4.1 Back to earth

As a reminder, for actual contests, we will almost never really have to deal with the full pathology. In particular, Theorem 3.8 gives us one way of finishing once we have gotten our additivity, and in many cases even this is not necessary (see the solution to Problem 40).

§4.5 Walkthroughs

Problem 53 (Gabriel Dospinescu)**.** Find all $f\colon \mathbb{Z}_{\geq 0} \to \mathbb{R}$ satisfying

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

for nonnegative integers x and y.

Walkthrough. This is an example of a functional equation with a couple instructive properties: (i) you really see what it means when you get an f-recursion that pins down f mostly, but not completely, and (ii) why linear algebra intuition can be really helpful. It's one of my favorite examples.

First, let's find as many solutions as we can.

(a) Check that the set of solutions forms a real vector space.

(b) Find all polynomial solutions. There are nonlinear solutions!

(c) Check that $f(x) = x^2 \bmod 3$ is a solution.

(d) Guess another non-polynomial solution, not in the span of the solutions in (b) or (c). This is a lot easier to do than it seems; many people get it on their first guess.

(e) Collate the previous parts to show that the space of solutions has dimension at least five.

At this point, we want to start bounding the set of solutions above.

(f) Try plugging in $x = 0$ or $x = 1$. Why does this make you sad?

(g) Plug in $(x, y) = (n, 2)$ to get a recursion for $f(2n+1)$ in terms of smaller f-inputs (for n large enough).

(h) Let's keep pushing. Express $f(4n+1)$ in two different ways, and use this to find a recursion for $f(2n+2)$ in terms of smaller f-inputs (for n large enough).

(i) Use (g) and (h) to show that the set of solutions is at most six-dimensional: i.e. there exist six initial values a_1, ..., a_6 for which the values $f(a_1)$, ..., $f(a_6)$ give at most one function f.

(j) At this point we have a crisis: we have a five-dimensional space of solutions found, but an upper bound of six dimensions. Is the answer five or six dimensional? So far we have only used the given equation when $x = 2$ or $x = 4$.

Well, here's one way to try and tell: pick your six favorite values in (i), and pump the recursion to get a function f. If f doesn't satisfy the conditions, then you know there is some relation you haven't found yet. But if it does, then there's a good chance we're missing a sixth solution.

(k) Depending on what you found in (j), do *either*:

- Find a sixth solution you did not find in (e), showing that there is indeed a six-dimensional space of solutions.
- Find a relation between the six initial values in (i), showing that the solutions you found in (e) were the only ones.

Possible hint: think about the value of $f(0)$, and part (f).

Problem 54 (IMO Shortlist 2004)**.** Find all functions $f : \mathbb{R} \to \mathbb{R}$ which obey

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

for all real numbers x and y.

Walkthrough. Fasten your seat belts!

(a) Find all linear solutions, and show there are no nonlinear polynomial solutions.

(b) There is no reason not to work with $x + y$ and xy as variables here, so show that the problem is the same as

$$f(s^2 + g(p)) = f(s)^2 \qquad \forall s^2 \geq 4p$$

where $g(p) = 2f(p) - 2p$.

(c) Prove that $|f(x)| = |f(-x)|$ for all x, and f is eventually nonnegative.

Suppose that $f(x) \not\equiv x + c$, so g is nonconstant. We will use this to prove that f is eventually constant. Intuitively this makes sense: the equation in (b) gives two incompatible equations if the g values are different. The details require significant care, though. (You are welcome to secretly skip the next few parts, as the fun portion is after this.)

(d) Show that there is a constant $c \neq 0$ such that if $u^2 - v^2 = c$, then $f(u) = f(v)$ for $u, v \gg 0$.

(e) For u and v as in (d), prove that $g(v) - g(u) = \frac{k}{u+v}$ for some constant k.

(f) The trick is now to use u and v as values of p: show that for s large enough in terms of u and v we have

$$f(s)^2 = f\left(s^2 + g(u)\right) = f\left(s^2 + g(v)\right).$$

(g) Let u and v vary now in an interval $[M, 3M]$, where M is huge. Show there is a constant δ with the property: if $a^2 - b^2 = [\delta, 2\delta]$ and $a > b > 12M$ then $f(a) = f(b)$.

(h) Deduce f is eventually constant.

So assume eventually f is some constant k. Now for the fun part!

(i) Prove that $k = 0$ or $k = 1$.

(j) Show that $f(s)^2 = k$ for every s by taking p to be a large negative constant in the original equation.

(k) Deduce that if $k = 0$ then $f \equiv 0$.

Henceforth, assume $k = 1$.

(l) Prove that $f(0) = +1$ (probably by contradiction). Conclude that $f(x) = +1$ for $x \geq 2$.

(m) If $f(t) = -1$, try to pick x and y such that $x^2 + y^2 - 2 = xy = t$, which would give a contradiction. You will find this is only possible for certain t. Which ones?

(n) The result in (j) is no surprise: come up with an example of a function for which $f = 1$ eventually but $f(-10000) = -1$.

(o) Figure out the set of solutions to the original problem. You should find the number of solutions has cardinality $2^{|\mathbb{R}|}$, the so-called hypercontinuum!

§4.6 Problems

Problem 55 (HMMT November 2015)**.** Consider functions $f\colon \mathbb{Z} \to \mathbb{Z}$ satisfying

$$f(f(x) + 2x + 20) = 15.$$

An integer n is called *undetermined* if $f(n)$ could take any value, i.e. for every integer y, some function f as above satisfies $f(n) = y$. Which integers are undetermined?

Problem 56 (IMO 2012)**.** Find all functions $f : \mathbb{Z} \to \mathbb{Z}$ 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 57 (USA TST 2015)**.** Let $f\colon \mathbb{Q} \to \mathbb{Q}$ 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 58. Let $f\colon \mathbb{R} \to \mathbb{R}$ be a function such that $f(x+y) = f(x) + f(y)$ and $f(f(x)) = x$ for all real numbers x and y. Must f be linear?

Problem 59 (IMO Shortlist 2001)**.** Find all functions $f\colon \mathbb{R} \to \mathbb{R}$ satisfying

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

for all real numbers x and y.

Problem 60 (ELMO Shortlist 2013)**.** Let $\mathbb{N}$ denote the set of positive integers, and for a function f, let $f^k(n)$ denote the function f applied k times. Call a function $f : \mathbb{N} \to \mathbb{N}$ *saturated* if

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

for every positive integer n. Find all positive integers m for which the following holds: every saturated function f satisfies $f^{2014}(m) = m$.

Problem 61 (EGMO 2014)**.** Solve over $\mathbb{R}$ the functional equation

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

Problem 62 (IMO 2015)**.** Solve the functional equation

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

for $f\colon \mathbb{R} \to \mathbb{R}$.

Problem 63 (IMO 1998)**.** Classify all functions $f\colon \mathbb{N} \to \mathbb{N}$ satisfying the identity

$$f(n^2 f(m)) = mf(n)^2.$$

§4.7 Solutions

Solution 53. First, note the set of solutions is a real vector space and the following are six linearly independent solutions:

- $f(x) \equiv 1$
- $f(x) \equiv x$
- $f(x) \equiv x^2$
- $f(x) \equiv x \pmod 2$ (or equivalently $f(x) \equiv x^2 \pmod 4$)
- $f(x) \equiv x^2 \pmod 3$ (which is motivated from $x^2 \pmod 4$ working!)
- $f(0) = 1$ and $f(x) = 0$ for $x > 0$.

So we only need to show the solution set has dimension at most 6. Plug in $(x, y) = (n, 2)$ to get

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

Plug in $(x, y) = (2n, 2)$ and $(x, y) = (n, 4)$ gives

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

One can check that these two recursions let us determine f from the values of f at 0, 1, 2, 3, 4, 6.

Solution 54. The answer is $f(x) \equiv x$, $f(x) \equiv 0$ and

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

where S is any subset of $(-\infty, -\frac{2}{3})$.

It will be more economical to write the given condition as

$$f(s)^2 = f\left(s^2 + g(p)\right) \qquad \forall s^2 \geq 4p$$

where $g(x) = 2(f(x) - x)$. (Here $s = x + y$ and $p = xy$.)

First, note that if g is a constant function, it follows $f(x) \equiv x + f(0)$ in which case we derive immediately that $f(x) \equiv x$. So let us assume g is not a constant function in the sequel.

First, note that by taking $p = 0$ we derive the following claim.

Claim 4.6. $|f(-x)| = |f(x)|$ for every real number x, and $f(x) \geq 0$ for $x \gg 0$ (in fact for $x \geq g(0)$).

Now we claim that:

Claim 4.7. The function f is eventually constant, that is $f(x) = k$ for $x \gg 0$.

Proof. Since g is not constant, pick $c_1 = g(p_1)$, $c_2 = g(p_2)$ with $c_1 > c_2$. Then we remark that

$$\text{If } u > v \gg 0 \text{ and } u^2 - v^2 = c_1 - c_2 \quad \text{then} \quad f(u) = f(v).$$

since $f(u)^2 = f(u^2 + c_2) = f(v^2 + c_1) = f(v)^2$ for $u > v > 4\max(p_1, p_2)$; and if $u, v > g(0)$ as well then the f are nonnegative. (To be explicit, the $u > v \gg 0$ dependence is $u > v > \max(4p_1, 4p_2, g(0))$.)

Now the trick is to use u and v as values of p; we have

$$g(v) - g(u) = 2\left((f(v) - f(u)) + u - v\right) = 2(u - v) = \frac{2(c_1 - c_2)}{u + v}.$$

So let u and v vary with $u, v \in [M, 3M]$ and $u^2 - v^2 = c_1 - c_2$ for large $M > \max(4p_1, 4p_2, g(0))$. Then $g(v) - g(u)$ achieves all values between $[\delta, 2\delta]$ for some small δ. Then by repeating the argument with these u and v in place of p_1 and p_2, we find

$$\text{If } a > b > L \text{ and } a^2 - b^2 \in [\delta, 2\delta] \quad \text{then} \quad f(a) = f(b)$$

where $L = 12M$ ensures $L = 12M \geq \max(4u, 4v, g(0))$.

Then f is eventually constant on the interval $[\sqrt{L^2 + \delta}, \sqrt{L^2 + 2\delta}]$, since it equals $f(L)$ there. Similarly, it is constant on the intervals $[\sqrt{L^2 + 2\delta}, \sqrt{L^2 + 3\delta}]$, $[\sqrt{L^2 + 3\delta}, \sqrt{L^2 + 4\delta}]$, and so on. Therefore the function f is eventually constant. □

Remark. It is necessary to introduce the upper cap $3M$ on the values of u and v. Otherwise the threshold values of L may become arbitrarily large as $u, v \to \infty$.

So, we may assume $f(x) = k$ for $x \gg 0$. From large s and $p = 0$ we conclude $k^2 = k$, so

$$k = 0 \quad \text{or} \quad k = 1.$$

Claim 4.8. For every $s \in \mathbb{R}$, $f(s)^2 = k$.

Proof. The idea is to make $g(p)$ large for $p < 0$ (since this enables us to select any s).

Let $p = -N$ for a large value $N \gg 0$. Then $g(p) = 2(N + f(-N)) \geq 2N - 2$, so

$$f(s)^2 = f(s^2 + 2g(p)) = f(\text{large}) = k. \qquad \square$$

Now if $k = 0$ we are done; we have $f \equiv 0$. So we are left with the case where

$$f(x) \in \{-1, 1\}$$

for every x.

Claim 4.9. Assume $f(x) = \pm 1$ for every x. If $f(t) = -1$ then $t < -\frac{2}{3}$.

Proof. Now the condition becomes

$$f(x^2 + y^2 + 2f(xy)) = 1 \qquad \forall x, y \in \mathbb{R}.$$

Note that $f(0) = +1$ since otherwise we can put $(x, y) = (\sqrt{2}, 0)$. Hence for $x \geq 2$ we have $f(x) = 1$.

Now assume for contradiction that $-\frac{2}{3} \leq t \leq 2$. Then we can find x and y such that

$$x^2 + y^2 - 2 = xy = t.$$

since $t + 2 \geq |2t|$. But then $f(t) = +1$, contradiction. □

This concludes the proof.

Solution 55. The answer is that -35 is the only undetermined integer.

First, suppose n is undetermined. Then, there should be an f such that $f(n) = -n - 20$. Putting this into the given then gives

$$15 = f(f(n) + 2n + 20) = f((-n - 20) + 2n + 20) = f(n) = -n - 20$$

so we must have $n = -35$.

Conversely, for any integer y, we construct the function

$$f(x) = \begin{cases} y & x = -35 \\ 15 & x \neq -35. \end{cases}$$

This function f apparently satisfies the given condition. Therefore, -35 is indeed undetermined.

Solution 56. Answer: 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 $x \equiv 2 \pmod 4$.

These can be painfully seen to work. (It's more natural to think of these as $f(x) = x^2$, $f(x) = x^2 \pmod 4$, $f(x) = x^2 \pmod 8$, and multiples thereof.)

Set $a = b = c = 0$ to get $f(0) = 0$. Then set $c = 0$ to get $f(a) = f(-a)$, so f is even. Now

$$f(a)^2 + f(b)^2 + f(a+b)^2 = 2f(a+b)\left(f(a) + f(b)\right) + 2f(a)f(b)$$

or

$$\left(f(a+b) - \left(f(a) + f(b)\right)\right)^2 = 4f(a)f(b).$$

Hence $f(a)f(b)$ is a perfect square for all $a, b \in \mathbb{Z}$. So there exists a λ such that $f(n) = \lambda g(n)^2$, where $g(n) \geq 0$. From here we recover

$$\boxed{g(a+b) = \pm g(a) \pm g(b)}.$$

Also $g(0) = 0$.

Let $k = g(1) \neq 0$. We now split into cases on $g(2)$:

- $g(2) = 0$. Put $b = 2$ in original to get $g(a+2) = \pm g(a) = +g(a)$.
- $g(2) = 2k$. Cases on $g(4)$:
  - $g(4) = 0$, then we get $(g(n))_{n\geq 0} = (0, 1, 2, 1, 0, 1, 2, 1, \dots)$. This works.
  - $g(4) = 4k$. This only happens when $g(1) = k$, $g(2) = 2k$, $g(3) = 3k$, $g(4) = 4k$. Then
    - $g(5) = \pm 3k \pm 2k = \pm 4k \pm k$.
    - $g(6) = \pm 4k \pm 2k = \pm 5k \pm k$.
    - …

    and so by induction $g(n) = nk$.

Solution 57. No, such a constant need not exist.

One possible solution is as follows: define a sequence by $x_0 = 1$ and

$$\begin{aligned}
2x_1 &= x_0 \\
2x_2 &= x_1 + 1 \\
2x_3 &= x_2 \\
2x_4 &= x_3 + 1 \\
2x_5 &= x_4 \\
2x_6 &= x_5 + 1 \\
&\vdots
\end{aligned}$$

Set $f(2^{-k}) = x_k$ and $f(2^k) = 2^k$ for $k = 0, 1, \ldots$. Then, let

$$f\left(a \cdot 2^k + \frac{b}{c}\right) = af\left(2^k\right) + \frac{b}{c}$$

for odd integers a, b, c. One can verify this works.

A second shorter solution (given by the proposer) is to set, whenever $\gcd(p, q) = 1$ and $q > 0$,

$$f\left(\frac{p}{q}\right) = \frac{p}{q}\left(1! + 2! + \cdots + q!\right).$$

Remark. Silly note: despite appearances, $f(x) = \lfloor x \rfloor$ is not a counterexample since one can take $c = 0$.

Solution 58. The answer is no. Fix $(e_\alpha)_\alpha$ a Hamel basis of $\mathbb{R}$ (viewed as a $\mathbb{Q}$-vector space). A function sending

$$f(e_\alpha) = \pm e_\alpha$$

with f extended linearly will then work, for any choice of $\pm$ signs for each α. The linear solutions correspond to always picking $+$ or always picking $-$, but there are uncountably many other solutions obtained by varying the signs across α.

Solution 59. Answer: any function f of the form

$$f(x) = \begin{cases} Cx & x \in G \\ 0 & x \notin G \end{cases}$$

where G is a multiplicatively closed subgroup of the real numbers, and C is an arbitrary real number.

One can check mechanically that such solutions work. Now we prove they are the only ones. Indeed, let $C = f(1)$ and put $y = 1$ into the given to get

$$f(x)\left(f(x) - C\right) = C(x-1)f(x)$$

whence we obtain that $f(x) \in \{0, Cx\}$ for each x. In particular $f(0) = 0$.

Now assume $C \neq 0$ else $f \equiv 0$ follows. Let $G = \{x \mid f(x) \neq 0 \iff f(x) = Cx\}$. We make the following remarks:

- $1 \in G$.
- If $a \in G$, then $1/a \in G$ as well. Indeed $a = -1$ is clear and otherwise substitute $(x, y) = (a, 1/a)$.

- If $a \neq b \in G$, then $ab \in G$ as well, just by setting $(x, y) = (a, b)$.
- If $a \in G$, then $a^2 \in G$. To see this, put $x = a^2$ and $y = 1/a$ to get

$$Ca \cdot \left(f(a^2) - \frac{C}{a} \right) = \left(a^2 - \frac{1}{a} \right) \cdot f(a^2) \cdot Ca.$$

  Then $f(a^2) \neq 0$.

Thus G contains 1, and is closed under multiplication and division; so it is a group.

Solution 60. The answer is all n dividing 2014.

First, one observes that f is both surjective and injective, so f is a permutation of the positive integers; and moreover all the cycles have finite length. Actually, we claim that the functions f are exactly those for which in any cycle C of f, all elements are divisible by the length of C. It's easy to see this is sufficient; we prove it is necessary.

Let C be a cycle with length d. Then the condition is that $d \mid f^{f(n)}(n)$ for any $n \in C$. Consequently, we have that

$$d \mid f(n) \implies d \mid n.$$

But obviously d divides some element of C, and so we conclude d divides all elements of C, as desired.

Solution 61. A key motivation throughout the problem is that the left-hand side is asymmetric while the right-hand side is symmetric. Thus any time we plug in two values for x and y we will also plug in the opposite pair. Once f is injective this will basically kill the problem.

First, we prove the following.

Lemma. *There is a unique $z \in \mathbb{R}$ such that $f(z) = 0$.*

Proof. Clearly by putting $y = -f(x)$ such z exists. Now, suppose $f(u) = f(v) = 0$. Then:

- Plug $(x, y) = (u, v)$ gives $f(v^2) = uv$.
- Plug $(x, y) = (v, u)$ gives $f(u^2) = uv$.
- Plug $(x, y) = (u, u)$ gives $f(u^2) = u^2$.
- Plug $(x, y) = (v, v)$ gives $f(v^2) = v^2$.

Consequently $u^2 = uv = v^2$ which yields $u = v$. □

Next let $(x,y) = (z,0)$ and $(x,y) = (0,z)$ to get

$$\begin{aligned} & f\left(2zf(0)\right) = f\left(z^2 + f(0)^2\right) = 0 \\ \implies & 2zf(0) = z^2 + f(0)^2 = z \\ \implies & \boxed{f(0) = z \in \left\{0, \frac{1}{2}\right\}}. \end{aligned}$$

We now set to prove:

Lemma. *The function f is injective.*

Proof. By putting $(x,y) = (x,z)$ and $(x,y) = (z,x)$ we get

$$f\left(f(x)^2 + z^2\right) = f\left(2zf(x) + x^2\right) = x(z + f(x)).$$

Now suppose $f(x_1) = f(x_2)$ but $x_1 \neq x_2$. This can only happen if $f(x_1) = f(x_2) = -z$. And now

$$f(x_i)^2 + z^2 = 2zf(x_i) + x_i^2 = z \qquad i = 1, 2.$$

Solving, we have $x_i = \pm 1$, $z = \frac{1}{2}$, (since $z = 0$ is not permissible). Thus we have "almost injectivity".

Now plug in $(x,y) = (-1,0)$ and $(x,y) = (0,-1)$ in the original and equate in order to obtain $f(-\frac{3}{4}) = f(\frac{5}{4})$, which contradicts the work above. □

Finally we may use the symmetry trick in full to obtain

$$y^2 + 2xf(y) + f(x)^2 = x^2 + 2yf(x) + f(y)^2. \qquad (\heartsuit)$$

In particular, by setting $y = 0$ we obtain

$$f(x)^2 = (z - x)^2.$$

Two easy cases remain:

- In the $z = 0$ case simply note that $(\heartsuit)$ gives $2xf(y) = 2yf(x)$, so for $x \neq 0$ the value $f(x)/x$ is constant and hence $f(x) \equiv \pm x$ follows.
- In the $z = \frac{1}{2}$ case $(\heartsuit)$ becomes $(2f(y) + 1)\, x = (2f(x) + 1)\, y$ and hence we're done again by the same reasoning.

Solution 62. The answers are $f(x) \equiv x$ and $f(x) \equiv 2 - x$. Obviously, both of them work.

Let $P(x,y)$ be the given assertion. We also will let $S = \{t \mid f(t) = t\}$ be the set of fixed points of f.

- From $P(0,0)$ we get $f(f(0)) = 0$.
- From $P(0, f(0))$ we get $2f(0) = f(0)^2$ and hence $f(0) \in \{0, 2\}$.
- From $P(x, 1)$ we find that $x + f(x+1) \in S$ for all x.

We now solve the case $f(0) = 2$.

Claim. If $f(0) = 2$ then $f(x) \equiv 2 - x$.

Proof. Let $t \in S$ be any fixed point. Then $P(0, t)$ gives $2 = 2t$ or $t = 1$; so $S = \{1\}$. But we also saw $x + f(x+1) \in S$, which implies $f(x) \equiv 2 - x$. □

Henceforth, assume $f(0) = 0$.

Claim. If $f(0) = 0$ then f is odd.

Proof. Note that $P(1, -1) \implies f(1) + f(-1) = 1 - f(1)$ and $P(-1, 1) \implies f(-1) + f(-1) = -1 + f(1)$, together giving $f(1) = 1$ and $f(-1) = -1$. To prove f odd we now obtain more fixed points:

- From $P(x, 0)$ we find that $x + f(x) \in S$ for all $x \in \mathbb{R}$.
- From $P(x-1, 1)$ we find that $x - 1 + f(x) \in S$ for all $x \in \mathbb{R}$.
- From $P(1, f(x) + x - 1)$ we find $x + 1 + f(x) \in S$ for all $x \in \mathbb{R}$.

Finally $P(x, -1)$ gives f odd. □

To finish from f odd, notice that

$$\begin{aligned} P(x, -x) &\implies f(x) + f(-x^2) = x - xf(x) \\ P(-x, x) &\implies f(-x) + f(-x^2) = -x + xf(-x) \end{aligned}$$

which upon subtracting gives $f(x) \equiv x$.

Solution 63. Let $\mathcal{P}$ be the set of primes, and let $g \colon \mathcal{P} \to \mathcal{P}$ be any involution on them. Extend g to a completely multiplicative function on $\mathbb{N}$. Then $f(n) = dg(n)$ is a solution for any $d \in \mathbb{N}$ which is fixed by g.

It's straightforward to check these all work, since $g \colon \mathbb{N} \to \mathbb{N}$ is an involution on them. So we prove these are the only functions.

Let $d = f(1)$.

Claim. We have $df(n) = f(dn)$ and $d \cdot f(ab) = f(a)f(b)$.

Proof. Let $P(m, n)$ denote the assertion in the problem statement. Off the bat,

- $P(1, 1)$ implies $f(d) = d^2$.

- $P(n,1)$ implies $f(f(n)) = d^2 n$. In particular, f is injective.
- $P(1,n)$ implies $f(dn^2) = f(n)^2$.

Then

$$\begin{aligned} f(a)^2 f(b)^2 &= f(da^2) f(b)^2 && \text{by third bullet} \\ &= f(b^2 f(f(da^2))) && \text{by problem statement} \\ &= f(b^2 \cdot d^2 \cdot da^2) && \text{by second bullet} \\ &= f(dab)^2 && \text{by third bullet} \\ \implies f(a)f(b) &= f(dab). \end{aligned}$$

This implies the first claim by taking $(a,b) = (1,n)$. Then $df(a) = f(da)$, and so we actually have $f(a)f(b) = df(ab)$. □

Claim. All values of f are divisible by d.

Proof. We have

$$\begin{aligned} f(n^2) &= \frac{1}{d} f(n)^2 \\ f(n^3) &= \frac{f(n^2)f(n)}{d} = \frac{f(n)^3}{d^2} \\ f(n^4) &= \frac{f(n^3)f(n)}{d} = \frac{f(n)^4}{d^3} \end{aligned}$$

and so on, which implies the result. □

Then, define $g(n) = f(n)/d$. We conclude that g is completely multiplicative, with $g(1) = 1$. However, $f(f(n)) = d^2 n$ also implies $g(g(n)) = n$, i.e. g is an involution. Moreover, since $f(d) = d^2$, $g(d) = d$.

All that remains is to check that g must map primes to primes to finish the description in the problem. This is immediate; since g is multiplicative and $g(1) = 1$, if $g(g(p)) = p$ then $g(p)$ can have at most one prime factor, hence $g(p)$ is itself prime.

Remark. The IMO problem actually asked for the least value of $f(1998)$. But for instruction purposes, it is probably better to just find all f. Since $1998 = 2 \cdot 3^3 \cdot 37$, this answer is $2^3 \cdot 3 \cdot 5 = 120$, anyways.

---

5 Selected Algebra from USA TST

§5.1 Problems

Problem 64 (USAMO 2018)**.** Let a, b, c be positive real numbers such that $a+b+c = 4\sqrt[3]{abc}$. Prove that

$$2(ab+bc+ca) + 4\min(a^2, b^2, c^2) \geq a^2+b^2+c^2.$$

Problem 65 (TSTST 2018)**.** For an integer $n > 0$, denote by $\mathcal{F}(n)$ the set of integers $m > 0$ for which the polynomial $p(x) = x^2 + mx + n$ has an integer root.

(a) Let S denote the set of integers $n > 0$ for which $\mathcal{F}(n)$ contains two consecutive integers. Show that S is infinite but

$$\sum_{n \in S} \frac{1}{n} \leq 1.$$

(b) Prove that there are infinitely many positive integers n such that $\mathcal{F}(n)$ contains three consecutive integers.

Problem 66 (TSTST 2018)**.** Let $S = \{1, \dots, 100\}$, and for every positive integer n define

$$T_n = \{(a_1, \dots, a_n) \in S^n \mid a_1 + \cdots + a_n \equiv 0 \pmod{100}\}.$$

Determine which n have the following property: if we color any 75 elements of S red, then at least half of the n-tuples in T_n have an even number of coordinates with red elements.

Problem 67 (USA TST 2016)**.** Let p be a prime number. Let $\mathbb{F}_p$ denote the integers modulo p, and let $\mathbb{F}_p[x]$ be the set of polynomials with coefficients in $\mathbb{F}_p$. Define $\Psi \colon \mathbb{F}_p[x] \to \mathbb{F}_p[x]$ by

$$\Psi\left(\sum_{i=0}^{n} a_i x^i\right) = \sum_{i=0}^{n} a_i x^{p^i}.$$

Prove that for nonzero polynomials $F, G \in \mathbb{F}_p[x]$,

$$\Psi(\gcd(F,G)) = \gcd(\Psi(F), \Psi(G)).$$

Problem 68 (TSTST 2017)**.** Consider solutions to the equation

$$x^2 - cx + 1 = \frac{f(x)}{g(x)}$$

where f and g are nonzero polynomials with nonnegative real coefficients. For each $c > 0$, determine the minimum possible degree of f, or show that no such f, g exist.

Problem 69 (USA TST 2017)**.** Let $P, Q \in \mathbb{R}[x]$ be relatively prime nonconstant polynomials. Show that there can be at most three real numbers λ such that $P + \lambda Q$ is the square of a polynomial.

Problem 70 (USA TST 2018)**.** Alice and Bob play a game. First, Alice secretly picks a finite set S of lattice points in the Cartesian plane. Then, for every line ℓ in the plane which is horizontal, vertical, or has slope $+1$ or -1, she tells Bob the number of points of S that lie on ℓ. Bob wins if he can then determine the set S.

Prove that if Alice picks S to be of the form

$$S = \left\{(x, y) \in \mathbb{Z}^2 \mid m \le x^2 + y^2 \le n\right\}$$

for some positive integers m and n, then Bob can win. (Bob does not know in advance that S is of this form.)

§5.2 Solutions

Solution 64. WLOG let $c = \min(a, b, c) = 1$ by scaling. The given inequality becomes equivalent to

$$4ab + 2a + 2b + 3 \geq (a+b)^2 \qquad \forall a + b = 4(ab)^{1/3} - 1.$$

Now, let $t = (ab)^{1/3}$ and eliminate $a + b$ using the condition, to get

$$4t^3 + 2(4t-1) + 3 \geq (4t-1)^2 \iff 0 \leq 4t^3 - 16t^2 + 16t = 4t(t-2)^2$$

which solves the problem.

Equality occurs only if $t = 2$, meaning $ab = 8$ and $a + b = 7$, which gives

$$\{a, b\} = \left\{ \frac{7 \pm \sqrt{17}}{2} \right\}$$

with the assumption $c = 1$. Scaling gives the curve of equality cases.

Solution 65. We prove the following.

Claim. The set S is given explicitly by $S = \{x(x+1)y(y+1) \mid x, y > 0\}$.

Proof. Note that $m, m+1 \in \mathcal{F}(n)$ if and only if there exist integers $q > p \geq 0$ such that

$$\begin{aligned} m^2 - 4n &= p^2 \\ (m+1)^2 - 4n &= q^2. \end{aligned}$$

Subtraction gives $2m + 1 = q^2 - p^2$, so p and q are different parities. We can thus let $q - p = 2x + 1$, $q + p = 2y + 1$, where $y \geq x \geq 0$ are integers. It follows that

$$\begin{aligned} 4n &= m^2 - p^2 \\ &= \left(\frac{q^2 - p^2 - 1}{2} \right)^2 - p^2 = \left(\frac{q^2 - p^2 - 1}{2} - p \right) \left(\frac{q^2 - p^2 - 1}{2} + p \right) \\ &= \frac{q^2 - (p^2 + 2p + 1)}{2} \cdot \frac{q^2 - (p^2 - 2p + 1)}{2} \\ &= \frac{1}{4}(q - p - 1)(q - p + 1)(q + p - 1)(q + p + 1) = \frac{1}{4}(2x)(2x+2)(2y)(2y+2) \\ \implies n &= x(x+1)y(y+1). \end{aligned}$$

Since $n > 0$ we require $x, y > 0$. Conversely, if $n = x(x+1)y(y+1)$ for positive x and y then $m = \sqrt{p^2 + 4n} = \sqrt{(y-x)^2 + 4n} = 2xy + x + y = x(y+1) + (x+1)y$ and $m + 1 = 2xy + x + y + 1 = xy + (x+1)(y+1)$. Thus we conclude the main claim. □

From this, part (a) follows as

$$\sum_{n\in S} n^{-1} \le \left(\sum_{x\ge 1} \frac{1}{x(x+1)}\right)\left(\sum_{y\ge 1} \frac{1}{y(y+1)}\right) = 1\cdot 1 = 1.$$

As for (b), retain the notation in the proof of the claim. Now $m+2 \in S$ if and only if $(m+2)^2 - 4n$ is a square, say r^2. Writing in terms of p and q as parameters we find

$$\begin{aligned} r^2 &= (m+2)^2 - 4n = m^2 - 4n + 4m + 4 = p^2 + 2 + 2(2m+1) \\ &= p^2 + 2(q^2-p^2) + 2 = 2q^2 - p^2 + 2 \\ \iff 2q^2 + 2 &= p^2 + r^2 \qquad (\dagger) \end{aligned}$$

with $q > p$ of different parity and $n = \frac{1}{16}(q-p-1)(q-p+1)(q+p-1)(q+p+1)$.

Note that (by taking modulo 8) we have $q \not\equiv p \equiv r \pmod 2$, and so there are no parity issues and we will always assume $p < q < r$ in $(\dagger)$. Now, for every q, the equation $(\dagger)$ has a canonical solution $(p,r) = (q-1, q+1)$, but this leaves $n = 0$. Thus we want to show for infinitely many q there is a third way to write $2q^2+2$ as a sum of squares, which will give the desired p.

To do this, choose large integers q such that q^2+1 is divisible by at least three distinct 1 mod 4 primes. Since each such prime can be written as a sum of two squares, using Lagrange identity, we can deduce that $2q^2+2$ can be written as a sum of two squares in at least three different ways, as desired.

Remark. We can see that $n = 144$ is the smallest integer such that $\mathcal{F}(n)$ contains three consecutive integers and $n = 15120$ is the smallest integer such that $\mathcal{F}(n)$ contains four consecutive integers. It would be interesting to determine whether the number of consecutive elements in $\mathcal{F}(n)$ can be arbitrarily large or is bounded.

Solution 66. We claim this holds exactly for n even.

First solution by generating functions Define

$$R(x) = \sum_{s \text{ red}} x^s, \qquad B(x) = \sum_{s \text{ blue}} x^s.$$

(Here "blue" means "not-red", as always.) Then, the number of tuples in T_n with exactly k red coordinates is exactly equal to

$$\binom{n}{k} \cdot \frac{1}{100} \sum_{\omega} R(\omega)^k B(\omega)^{n-k}$$

where the sum is over all 100th roots of unity. So, we conclude the number of tuples in T_n with an even (resp odd) number of red elements is exactly

$$\begin{aligned}
X &= \frac{1}{100} \sum_{\omega} \sum_{k \text{ even}} \binom{n}{k} R(\omega)^k B(\omega)^{n-k} \\
Y &= \frac{1}{100} \sum_{\omega} \sum_{k \text{ odd}} \binom{n}{k} R(\omega)^k B(\omega)^{n-k} \\
\implies X - Y &= \frac{1}{100} \sum_{\omega} (B(\omega) - R(\omega))^n \\
&= \frac{1}{100} \left[(B(1) - R(1))^n + \sum_{\omega \neq 1} (2B(\omega))^n \right] \\
&= \frac{1}{100} \left[(B(1) - R(1))^n - (2B(1))^n + 2^n \sum_{\omega} B(\omega)^n \right] \\
&= \frac{1}{100} \left[(B(1) - R(1))^n - (2B(1))^n \right] + 2^n Z \\
&= \frac{1}{100} \left[(-50)^n - 50^n \right] + 2^n Z.
\end{aligned}$$

where

$$Z \coloneqq \frac{1}{100} \sum_{\omega} B(\omega)^n \geq 0$$

counts the number of tuples in T_n which are all blue. Here we have used the fact that $B(\omega) + R(\omega) = 0$ for $\omega \neq 1$.

We wish to show $X - Y \geq 0$ holds for n even, but may fail when n is odd. This follows from two remarks:

- If n is even, then $X - Y = 2^n Z \geq 0$.
- If n is odd, then if we choose the coloring for which s is red if and only if $s \not\equiv 2 \pmod 4$; we thus get $Z = 0$. Then $X - Y = -\frac{2}{100} \cdot 50^n < 0$.

Second solution by strengthened induction and random coloring We again prove that n even work. Let us define

$$T_n(a) = \{(a_1, \dots, a_n) \in S^n \mid a_1 + \cdots + a_n \equiv a \pmod{100}\}.$$

Also, call an n-tuple good if it has an even number of red elements. We claim that $T_n(a)$ also has at least 50% good tuples, by induction.

This follows by induction on $n \geq 2$. Indeed, the base case $n = 2$ can be checked by hand, since $T_2(a) = \{(x, a - x) \mid x \in S\}$. With the stronger claim, one can check the case $n = 2$ manually and proceed by induction to go from $n - 2$ to n, noting that

$$T_n(a) = \bigsqcup_{b+c=a} T_{n-2}(b) \oplus T_2(c)$$

where $\oplus$ denotes concatenation of tuples, applied set-wise. The concatenation of an $(n-2)$-tuple and 2-tuple is good if and only if both or neither are good. Thus for each b and c, if the proportion of $T_{n-2}(b)$ which is good is $p \geq \frac{1}{2}$ and the proportion of $T_2(c)$ which is good is $q \geq \frac{1}{2}$, then the proportion of $T_{n-2}(b) \oplus T_2(c)$ which is good is $pq + (1-p)(1-q) \geq \frac{1}{2}$, as desired. Since each term in the union has at least half the tuples good, all of $T_n(a)$ has at least half the tuples good, as desired.

It remains to fail all odd n. We proceed by a suggestion of Yang Liu and Ankan Bhattacharya by showing that if we pick the 75 elements *randomly*, then any particular tuple in S^n has strictly less than 50% chance of being good. This will imply (by linearity of expectation) that T_n (or indeed any subset of S^n) will, for some coloring, have less than half good tuples.

Let $(a_1, \dots, a_n)$ be such an n-tuple. If any element appears in the tuple more than once, keep *discarding pairs* of that element until there are zero or one; this has no effect on the good-ness of the tuple. If we do this, we obtain an m-tuple $(b_1, \dots, b_m)$ with no duplicated elements where $m \equiv n \equiv 1 \pmod 2$. Now, the probability that any element is red is $\frac{3}{4}$, so the probability of being good is

$$\begin{aligned}\sum_{k \text{ even}}^{m} \binom{m}{k} \left(\frac{3}{4}\right)^k \left(-\frac{1}{4}\right)^{m-k} &= \frac{1}{2}\left[\left(\frac{3}{4}+\frac{1}{4}\right)^m - \left(\frac{3}{4}-\frac{1}{4}\right)^m\right] \\ &= \frac{1}{2}\left[1 - \left(\frac{1}{2}\right)^m\right] < \frac{1}{2}.\end{aligned}$$

Remark (Adam Hesterberg)**.** Here is yet another proof that n even works. Group elements of T_n into equivalence classes according to the $n/2$ sums of pairs of consecutive elements (first and second, third and fourth, …). For each such pair sum, there are at least as many monochrome pairs with that sum as nonmonochrome ones, since every nonmonochrome pair uses one of the 25 non-reds. The monochromaticity of the pairs is independent.

If $p_i \leq \frac{1}{2}$ is the probability that the ith pair is nonmonochrome, then the probability that k pairs are nonmonochrome is the coefficient of x^k in $f(x) = \prod_i (xp_i + (1-p_i))$. Then the probability that evenly many pairs are nonmonochrome (and hence that evenly many coordinates are red) is the sum of the coefficients of even powers of x in f, which is $(f(1)+f(-1))/2 = (1+\prod_i(1-2p_i))/2 \geq \frac{1}{2}$, as desired.

Solution 67. Observe that Ψ is also a linear map of $\mathbb{F}_p$ vector spaces, and that $\Psi(xP) = \Psi(P)^p$ for any $P \in \mathbb{F}_p[x]$. (In particular, $\Psi(1) = x$, not 1, take caution!)

First solution (Ankan Bhattacharya) We start with:

Claim. If $P \mid Q$ then $\Psi(P) \mid \Psi(Q)$.

Proof. Set $Q = PR$, where $R = \sum_{i=0}^{k} r_i x^i$. Then

$$\Psi(Q) = \Psi\left(P \sum_{i=0}^{k} r_i x^i\right) = \sum_{i=0}^{k} \Psi\left(P \cdot r_i x^i\right) = \sum_{i=0}^{k} r_i \Psi(P)^{p^i}$$

which is divisible by $\Psi(P)$. □

This already implies

$$\Psi(\gcd(F, G)) \mid \gcd(\Psi(F), \Psi(G)).$$

For the converse, by Bezout there exists $A, B \in \mathbb{F}_p[x]$ such that $AF + BG = \gcd(F, G)$, so taking Ψ of both sides gives

$$\Psi(AF) + \Psi(BG) = \Psi\left(\gcd(F, G)\right).$$

The left-hand side is divisible by $\gcd(\Psi(F), \Psi(G))$ since the first term is divisible by $\Psi(F)$ and the second term is divisible by $\Psi(G)$. So $\gcd(\Psi(F), \Psi(G)) \mid \Psi(\gcd(F, G))$ and noting both sides are monic we are done.

Second solution Here is an alternative (longer but more conceptual) way to finish without Bezout lemma. Let $\beth \subseteq \mathbb{F}_p[x]$ denote the set of polynomials in the image of Ψ, thus $\Psi \colon \mathbb{F}_p[x] \to \beth$ is a bijection on the level of sets.

Claim. If $A, B \in \beth$ then $\gcd(A, B) \in \beth$.

Proof. It suffices to show that if A and B are monic, and $\deg A > \deg B$, then the remainder when A is divided by B is in $\beth$. Suppose $\deg A = p^k$ and $B = x^{p^{k-1}} - c_2 x^{p^{k-2}} - \cdots - c_k$. Then

$$\begin{aligned} x^{p^k} &\equiv \left(c_2 x^{p^{k-2}} + c_3 x^{p^{k-3}} + \cdots + c_k\right)^p \pmod{B} \\ &\equiv c_2 x^{p^{k-1}} + c_3 x^{p^{k-2}} \cdots + c_k \pmod{B} \end{aligned}$$

since exponentiation by p commutes with addition in $\mathbb{F}_p$. This is enough to imply the conclusion. The proof if $\deg B$ is smaller less than p^{k-1} is similar. □

Thus, if we view $\mathbb{F}_p[x]$ and $\beth$ as partially ordered sets under polynomial division, then gcd is the "greatest lower bound" or "meet" in both partially ordered sets. We will now prove that Ψ is an *isomorphism* of the posets. We have already seen that $P \mid Q \implies \Psi(P) \mid \Psi(Q)$ from the first solution. For the converse:

Claim. If $\Psi(P) \mid \Psi(Q)$ then $P \mid Q$.

Proof. Suppose $\Psi(P) \mid \Psi(Q)$, but $Q = PA + B$ where $\deg B < \deg P$. Thus $\Psi(P) \mid \Psi(PA) + \Psi(B)$, hence $\Psi(P) \mid \Psi(B)$, but $\deg \Psi(P) > \deg \Psi(B)$ hence $\Psi(B) = 0 \implies B = 0$. □

This completes the proof.

Remark. In fact $\Psi \colon \mathbb{F}_p[x] \to \beth$ is a ring isomorphism if we equip $\beth$ with function composition as the ring multiplication. Indeed in the proof of the first claim (that $P \mid Q \implies \Psi(P) \mid \Psi(Q)$) we saw that

$$\Psi(RP) = \sum_{i=0}^{k} r_i \Psi(P)^{p^i} = \Psi(R) \circ \Psi(P).$$

Solution 68. First, if $c \geq 2$ then we claim no such f and g exist. Indeed, one simply takes $x = 1$ to get $f(1)/g(1) \leq 0$, impossible.

For $c < 2$, let $c = 2\cos\theta$, where $0 < \theta < \pi$. We claim that f exists and has minimum degree equal to n, where n is defined as the smallest integer satisfying $\sin n\theta \leq 0$. In other words

$$n = \left\lceil \frac{\pi}{\arccos(c/2)} \right\rceil.$$

First we show that this is necessary. To see it, write explicitly

$$g(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-2} x^{n-2}$$

with each $a_i \geq 0$, and $a_{n-2} \neq 0$. Assume that n is such that $\sin(k\theta) \geq 0$ for $k = 1, \ldots, n-1$. Then, we have the following system of inequalities:

$$\begin{aligned}
a_1 &\geq 2\cos\theta \cdot a_0 \\
a_0 + a_2 &\geq 2\cos\theta \cdot a_1 \\
a_1 + a_3 &\geq 2\cos\theta \cdot a_2 \\
&\vdots \\
a_{n-5} + a_{n-3} &\geq 2\cos\theta \cdot a_{n-4} \\
a_{n-4} + a_{n-2} &\geq 2\cos\theta \cdot a_{n-3} \\
a_{n-3} &\geq 2\cos\theta \cdot a_{n-2}.
\end{aligned}$$

Now, multiply the first equation by $\sin\theta$, the second equation by $\sin 2\theta$, et cetera, up to $\sin((n-1)\theta)$. This choice of weights is selected since we have

$$\sin(k\theta) + \sin((k+2)\theta) = 2\sin((k+1)\theta)\cos\theta$$

so that summing the entire expression cancels nearly all terms and leaves only

$$\sin((n-2)\theta)\, a_{n-2} \geq \sin((n-1)\theta) \cdot 2\cos\theta \cdot a_{n-2}$$

and so by dividing by a_{n-2} and using the same identity gives us $\sin(n\theta) \le 0$, as claimed.

This bound is best possible, because the example

$$a_k = \sin\left((k+1)\theta\right) \ge 0$$

makes all inequalities above sharp, hence giving a working pair (f, g).

Remark. Calvin Deng points out that a cleaner proof of the lower bound is to take $\alpha = \cos\theta + i\sin\theta$. Then $f(\alpha) = 0$, but by condition the imaginary part of $f(\alpha)$ is apparently strictly positive, contradiction.

Remark. Guessing that $c < 2$ works at all (and realizing $c \ge 2$ fails) is the first part of the problem.

The introduction of trigonometry into the solution may seem magical, but is motivated in one of two ways:

- Calvin Deng points out that it's possible to guess the answer from small cases: For $c \le 1$ we have $n = 3$, tight at $\frac{x^3+1}{x+1} = x^2 - x + 1$, and essentially the "sharpest $n = 3$ example". A similar example exists at $n = 4$ with $\frac{x^4+1}{x^2+\sqrt{2}x+1} = x^2 - \sqrt{2}x + 1$ by the Sophie-Germain identity. In general, one can do long division to extract an optimal value of c for any given n, although c will be the root of some polynomial.

  The thresholds $c \le 1$ for $n = 3$, $c \le \sqrt{2}$ for $n = 4$, $c \le \frac{1+\sqrt{5}}{2}$ for $n = 5$, and $c \le 2$ for $n < \infty$ suggest the unusual form of the answer via trigonometry.

- One may imagine trying to construct a polynomial recursively / greedily by making all inequalities above hold (again the "sharpest situation" in which f has few coefficients). If one sets $c = 2t$, then we have

  $$a_0 = 1, \quad a_1 = 2t, \quad a_2 = 4t^2 - 1, \quad a_3 = 8t^3 - 4t, \quad \dots$$

  which are the Chebyshev polynomials of the second type. This means that trigonometry is essentially mandatory. (One may also run into this when by using standard linear recursion techniques, and noting that the characteristic polynomial has two conjugate complex roots.)

Remark. Mitchell Lee notes that an IMO longlist problem from 1997 shows that if $P(x)$ is any polynomial satisfying $P(x) > 0$ for $x > 0$, then $(x+1)^n P(x)$ has nonnegative coefficients for large enough n. This show that f and g at least exist for $c \le 2$, but provides no way of finding the best possible $\deg f$.

Meghal Gupta also points out that showing f and g exist is possible in the following way:

$$\left(x^2 - 1.99x + 1\right)\left(x^2 + 1.99x + 1\right) = \left(x^4 - 1.9601x^2 + 1\right)$$

and so on, repeatedly multiplying by the "conjugate" until all coefficients become positive. To my best knowledge, this also does not give any way of actually minimizing $\deg f$, although Ankan Bhattacharya points out that this construction is actually optimal in the case where n is a power of 2.

Remark. It's pointed out that Matematicheskoe Prosveshchenie, issue 1, 1997, page 194 contains a nearly analogous result, available at `https://mccme.ru/free-books/matpros/pdf/mp-01.pdf` with solutions presented in `https://mccme.ru/free-books/matpros/pdf/mp-05.pdf`, pages 221–223; and `https://mccme.ru/free-books/matpros/pdf/mp-10.pdf`, page 274.

Solution 69. This is true even with $\mathbb{R}$ replaced by $\mathbb{C}$, and it will be necessary to work in this generality.

First solution using transformations We will prove the claim in the following form:

Claim. Assume $P, Q \in \mathbb{C}[x]$ are relatively prime. If $\alpha P + \beta Q$ is a square for four different choices of the ratio $[\alpha : \beta]$ then P and Q must be constant.

Call pairs (P, Q) as in the claim *bad*; so we wish to show the only bad pairs are pairs of constant polynomials. Assume not, and take a bad pair with $\deg P + \deg Q$ minimal.

By a suitable Möbius transformation, we may transform (P, Q) so that the four ratios are $[1 : 0]$, $[0 : 1]$, $[1 : -1]$ and $[1 : -k]$, so we find there are polynomials A and B such that

$$A^2 - B^2 = C^2$$
$$A^2 - kB^2 = D^2$$

where $A^2 = P + \lambda_1 Q$, $B^2 = P + \lambda_2 Q$, say. Of course $\gcd(A, B) = 1$.

Consequently, we have $C^2 = (A + B)(A - B)$ and $D^2 = (A + \mu B)(A - \mu B)$ where $\mu^2 = k$. Now $\gcd(A, B) = 1$, so $A + B$, $A - B$, $A + \mu B$ and $A - \mu B$ are squares; id est (A, B) is bad. This is a contradiction, since $\deg A + \deg B < \deg P + \deg Q$.

Second solution using derivatives (by Zack Chroman) We will assume without loss of generality that $\deg P \neq \deg Q$; if not, then one can replace (P, Q) with $(P + cQ, Q)$ for a suitable constant c.

Then, there exist $\lambda_i \in \mathbb{C}$ and polynomials R_i for $i = 1, 2, 3, 4$ such that

$$\begin{aligned} & P + \lambda_i Q = R_i^2 \\ \implies & P' + \lambda_i Q' = 2 R_i R_i' \\ \implies & R_i \mid Q'(P + \lambda_i Q) - Q(P' + \lambda_i Q') = Q'P - QP'. \end{aligned}$$

On the other hand by Euclidean algorithm it follows that R_i are relatively prime to each other. Therefore

$$R_1R_2R_3R_4 \mid Q'P - QP'.$$

However, we have

$$\sum_1^4 \deg R_i \geq \frac{3\max(\deg P, \deg Q) + \min(\deg P, \deg Q)}{2} \geq \deg P + \deg Q > \deg(Q'P - QP'$$

This can only occur if $Q'P - QP' = 0$ or $(P/Q)' = 0$ by the quotient rule! But P/Q can't be constant, the end.

Remark. The result is previously known; see e.g. Lemma 1.6 of `http://math.mit.edu/~ebelmont/ec-notes.pdf` or Exercise 6.5.L(a) of Vakil's notes.

Solution 70. Clearly Bob can compute the number N of points.

The main claim is that:

Claim. Fix m and n as in the problem statement. Among all sets $T \subseteq \mathbb{Z}^2$ with N points, the set S is the *unique* one which maximizes the value of

$$F(T) := \sum_{(x,y)\in T} (x^2+y^2)(m+n-(x^2+y^2)).$$

Proof. Indeed, the different points in T do not interact in this sum, so we simply want the points (x,y) with x^2+y^2 as close as possible to $\frac{m+n}{2}$ which is exactly what S does. □

As a result of this observation, it suffices to show that Bob has enough information to compute $F(S)$ from the data given. (There is no issue with fixing m and n, since Bob can find an upper bound on the magnitude of the points and then check all pairs (m,n) smaller than that.) The idea is that he knows the full distribution of each of X, Y, $X+Y$, $X-Y$ and hence can compute sums over T of any power of a single one of those linear functions. By taking linear combinations we can hence compute $F(S)$.

Let us make the relations explicit. For ease of exposition we take $Z = (X,Y)$ to be a uniformly random point from the set S. The information is precisely the individual distributions of X, Y, $X+Y$, and $X-Y$. Now compute

$$\begin{aligned}\frac{F(S)}{N} &= \mathbb{E}\left[(m+n)(X^2+Y^2) - (X^2+Y^2)^2\right] \\ &= (m+n)\left(\mathbb{E}[X^2]+\mathbb{E}[Y^2]\right) - \mathbb{E}[X^4] - \mathbb{E}[Y^4] - 2\mathbb{E}[X^2Y^2].\end{aligned}$$

On the other hand,

$$\mathbb{E}[X^2Y^2] = \frac{\mathbb{E}[(X+Y)^4] + \mathbb{E}[(X-Y)^4] - 2\mathbb{E}[X^4] - 2\mathbb{E}[Y^4]}{12}.$$

Thus we have written $F(S)$ in terms of the distributions of X, Y, $X-Y$, $X+Y$ which completes the proof.

Remark (Mark Sellke).

- This proof would have worked just as well if we allowed arbitrary $[0,1]$-valued weights on points with finitely many weights non-zero. There is an obvious continuum generalization one can make concerning the indicator function for an annulus. It's a simpler but fun problem to characterize when just the vertical/horizontal directions determine the distribution.

- An obstruction to purely combinatorial arguments is that if you take an octagon with points $(\pm a, \pm b)$ and $(\pm b, \pm a)$ then the two ways to pick every other point (going around clockwise) are indistinguishable by Bob. This at least shows that Bob's task is far from possible in general, and hints at proving an inequality.

- A related and more standard fact (among a certain type of person) is that given a probability distribution μ on $\mathbb{R}^n$, if I tell you the distribution of *all* 1-dimensional projections of μ, that determines μ uniquely. This works because this information gives me the Fourier transform $\hat{\mu}$, and Fourier transforms are injective.

  For the continuum version of this problem, this connection gives a much larger family of counterexamples to any proposed extension to arbitrary non-annular shapes. Indeed, take a fast-decaying smooth function $f\colon \mathbb{R}^2 \to \mathbb{R}$ which vanishes on the four lines

  $$x = 0,\ y = 0,\ x+y = 0,\ x-y = 0.$$

  Then the Fourier transform $\hat{f}$ will have mean 0 on each line ℓ as in the problem statement. Hence the positive and negative parts of $\hat{f}$ will not be distinguishable by Bob.

II

Combinatorics

6 Graph Theory Terminology

Graph theory is a pretty field of mathematics which is remarkably accessible even without much prerequisites, but which is still an active field of research. It is, in my opinion, one of the best fields of post-high-school math to learn first if you wish to get comfortable with proof-based problems, for the following reasons:

- You will be able to get used to logical thinking and real proofs.
- The field is not overly technical, making it more intuitive.
- Nevertheless, you will still be proving results that feel concrete, interesting, and substantial.

Despite this, we will not use any serious graph theory in this textbook. Instead, **we are only going to use some basic terminology for convenience**, and will not use any deep theorems (or barely theorems at all). The purpose of this chapter is to provide a list of some common terms that we will be using. Some other terms may appear in solutions to more difficult problems, which you are responsible for looking up yourself.

You should spend at most 15 minutes reading this chapter as there is nothing to "get". (And if you are already familiar with graph theory, everything here will be review.)

§6.1 Textbook references

If you are interested in learning more graph theory (beyond the pitifully narrow scope of this book), here are some textbook recommendations.

- As a high school student, I liked Chartrand and Zhang, *A First Course in Graph Theory*.
- Another classical standard textbook is Diestel, *Graph Theory*.

In general, a lot of textbooks on graph theory are quite similar to each other, so I think you can't really go wrong with most of them.

§6.2 Graphs

A **graph** is a collection of **vertices** together with some number of **edges** which connect pairs of different vertices. In this book, repeated edges aren't allowed.[1]

Figure 6.1 shows a finite simple graph with six vertices and seven edges.

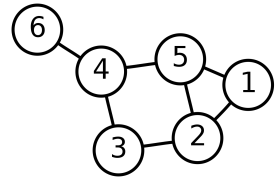


Figure 6.1: An example of a graph, with six vertices and seven edges. Image taken from Wikipedia, explicitly in the public domain.

The graph is **finite** if there are finitely many vertices, and assumption you can make throughout this textbook (although we will sometimes include the word "finite" for emphasis anyways).

§6.3 Degree

The **degree** of a vertex is the number of edges it touches, or put another way, the number of **neighbors** it has. So in Figure 6.1

- Vertex 1 has degree two (it has neighbors 2 and 5)
- Vertex 2 has degree three (it has neighbors 1, 3, and 5)
- Vertex 3 has degree two (it has neighbors 2 and 4)
- Vertex 4 has degree three (it has neighbors 3, 5 and 6)
- Vertex 5 has degree three (it has neighbors 1, 2, and 4)
- Vertex 6 has degree one (its only neighbor is 4).

Remark. The sum of the degrees is $2+3+2+3+3+1 = 14$, an even number. In the next chapter, the first example (Example 71) will ask you to prove this sum is an even number for *any* finite graph G. This is called the "handshake lemma", after a silly real-world interpretation where some people are shaking hands. If you are new to this, you can try to prove this now, or at least draw some more examples of graphs to convince yourself that this is true.

[1] We write "simple" to emphasize this, sometimes, but you can ignore this word each time you see it.

§6.4 Paths, cycles, connectedness

A **path** is exactly what it sounds like: a sequence of vertices which are adjacent and get you from one vertex to another. So,

$$1 \to 5 \to 4 \to 6$$

is an example of a path in Figure 6.1. We don't allow repeated vertices in a path.

A **cycle** is exactly what it sounds like: a path where you start and end at the same vertex (and, other than this, no repeated vertices are allowed). So

$$2 \to 3 \to 4 \to 5 \to 2$$

is an example of a cycle in Figure 6.1.

A graph is **connected** if there is a path between any two vertices in the graph (again, what it sounds like). So the above graph is connected. If a graph is not connected, it instead consists of several **connected components**; you can guess what this means. For example, Figure 6.2 is a graph which is not connected, but has three connected components.

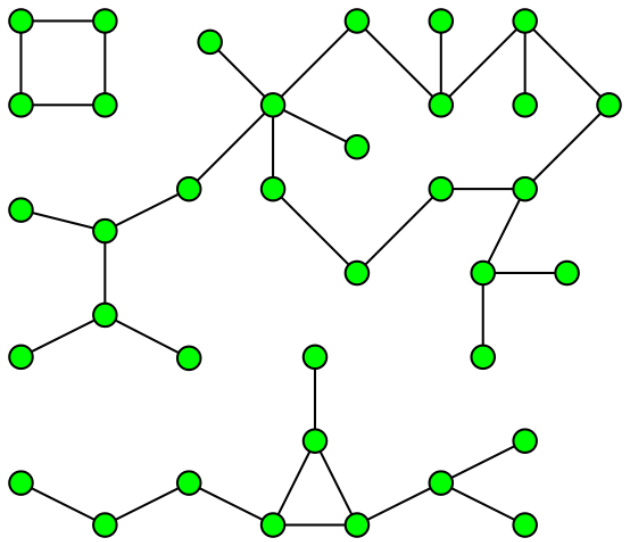


Figure 6.2: An example of a disconnected graph. It has three connected components. Image taken from Wikipedia, explicitly in public domain.

7 Global

This chapter is dedicated to the idea that you can often extract nontrivial information about a problem by looking at the entire structure at once.

§7.1 A simple example, the handshake lemma

The simplest example that many of you have already heard of is:

Example 71 (Handshake lemma)**.** Let G be a finite simple graph. Then the sum of all the degrees of vertices of G is an even number.

Solution 71. We claim that the sum of the degrees is equal to twice the number of edges. In fact both quantities are equal to the cardinality of the set

$$\{(v, e) \mid v \text{ vertex}, e \text{ edge containing } v\}.$$

In particular, this is even. ■

This example is a little surprising because this seems to actually be the shortest way to show that the degree sum is even, despite the fact that we were not initially interested in edges or the set above.

We will see some more examples in the walkthrough.

§7.2 Expected value

However, for the theoretical part of this chapter, I want to introduce a useful notion that will allow us to let us capture the "boiler-plate" of having to construct a double-indexed set, like the "subset of $V \times E$" above. That is the notion of *expected value* and *linearity of expectation*. Here are the blueprints.

§7.2.1 Definitions and notation

Nothing tricky here, just setting up notation. I'll try not to be overly formal.

A **random variable** is just a quantity that we take to vary randomly. For example, the outcome of a standard six-sided dice roll, say D_6, is a random variable. We can now discuss the **probability** of certain events, which we'll denote $\mathbb{P}(\bullet)$. For instance, we can write

$$\mathbb{P}(D_6 = 1) = \mathbb{P}(D_6 = 2) = \cdots = \mathbb{P}(D_6 = 6) = \frac{1}{6}$$

or $\mathbb{P}(D_6 = 0) = 0$ and $\mathbb{P}(D_6 \geq 4) = \frac{1}{2}$.

We can also discuss the **expected value** of a random variable X, which is the "average" value. The formal definition is

$$\mathbb{E}[X] \coloneqq \sum_x \mathbb{P}(X = x) \cdot x.$$

But an example for our dice roll D_6 makes this clearer:

$$\mathbb{E}[D_6] = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \cdots + \frac{1}{6} \cdot 6 = 3.5.$$

In natural language, we just add up all the outcomes weighted by probability they appear.

§7.2.2 Another motivating example

It is an unspoken law that any introduction to expected value begins with the following classical example.

Example 72. At MOP, there are n people, each of who has a name tag. We shuffle the name tags and randomly give each person one of the name tags. Let S be the number of people who receive their own name tag. Prove that the expected value of S is 1.

This result might seem surprising, as one might intuitively expect $\mathbb{E}[S]$ to depend on the choice of n.

Solution 72. For simplicity, let us call a person a *fixed point* if they receive their own name tag.[1] Thus S is just the number of fixed points, and we wish to show that $\mathbb{E}[S] = 1$. If we're interested in the expected value, then according to our definition we should go through all $n!$ permutations, count up the total number of fixed points, and then divide by $n!$ to get the average. Since we want $\mathbb{E}[S] = 1$, we expect to see a total of $n!$ fixed points.

Let us begin by illustrating the case $n = 4$ first, calling the people W, X, Y, Z.

[1] This is actually a term used to describe points which are unchanged by a permutation. So the usual phrasing of this question is "what is the expected number of fixed points of a random permutation?"

	W	X	Y	Z	Σ
1	**W**	**X**	**Y**	**Z**	4
2	**W**	**X**	Z	Y	2
3	**W**	Y	X	**Z**	2
4	**W**	Y	Z	X	1
5	**W**	Z	X	Y	1
6	**W**	Z	**Y**	X	2
7	X	W	**Y**	**Z**	2
8	X	W	Z	Y	0
9	X	Y	W	**Z**	1
10	X	Y	Z	W	0
11	X	Z	W	Y	0
12	X	Z	**Y**	W	1
13	Y	W	X	**Z**	1
14	Y	W	Z	X	0
15	Y	**X**	W	**Z**	2
16	Y	**X**	Z	W	1
17	Y	Z	W	X	0
18	Y	Z	X	W	0
19	Z	W	X	Y	0
20	Z	W	**Y**	X	1
21	Z	**X**	W	Y	1
22	Z	**X**	**Y**	W	2
23	Z	Y	W	X	0
24	Z	Y	X	W	0
Σ	6	6	6	6	24

We've listed all $4! = 24$ permutations, and indeed we see that there are a total of 24 fixed points, which I've bolded in red. Unfortunately, if we look at the rightmost column, there doesn't seem to be a pattern, and it seems hard to prove that this holds for larger n.

However, suppose that *rather than trying to add by rows, we add by columns.* There's a very clear pattern if we try to add by the columns: we see a total of 6 fixed points in each column. Indeed, the six fixed W points correspond to the $3! = 6$ permutations of the remaining letters X, Y, Z. Similarly, the six fixed X points correspond to the $3! = 6$ permutations of the remaining letters W, Y, Z.

This generalizes very nicely: if we have n letters, then each letter appears as a fixed point $(n-1)!$ times.

Thus the expected value is

$$\mathbb{E}[S] = \frac{1}{n!}\left(\underbrace{(n-1)! + (n-1)! + \cdots + (n-1)!}_{n \text{ times}}\right) = \frac{1}{n!} \cdot n \cdot (n-1)! = 1.$$

Cute, right? Now let's bring out the artillery. ■

§7.2.3 Linearity of expectation

The crux result of this section is the following theorem.

Theorem 7.1 (Linearity of Expectation)**.** *Given any random variables X_1, X_2, …, X_n, we always have*

$$\mathbb{E}[X_1 + X_2 + \cdots + X_n] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n].$$

This theorem is obvious if the X_1, X_2, …, X_n are independent of each other – if I roll 100 dice, I expect an average of 350. Duh. The wonderful thing is that this holds even if the variables are not independent. And the basic idea is just the double-counting we did in the earlier example: even if the variables depend on each other, if you look only at the expected value, you can still add just by columns. The proof of the theorem is just a bunch of sigma signs which say exactly the same thing, so I won't include it.

Anyways, that means we can now nuke our original problem. The trick is to define **indicator variables** as follows: for each $i = 1, 2, \ldots, n$ let

$$S_i \coloneqq \begin{cases} 1 & \text{if person } i \text{ gets his own name tag} \\ 0 & \text{otherwise.} \end{cases}$$

Obviously,

$$S = S_1 + S_2 + \cdots + S_n.$$

Moreover, it is easy to see that $\mathbb{E}[S_i] = \mathbb{P}(S_i = 1) = \frac{1}{n}$ for each i: if we look at any particular person, the probability they get their own name tag is simply $\frac{1}{n}$. Therefore,

$$\mathbb{E}[S] = \mathbb{E}[S_1] + \mathbb{E}[S_2] + \cdots + \mathbb{E}[S_n] = \underbrace{\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}}_{n \text{ times}} = 1.$$

Now that was a lot easier! By working in the context of expected value, we get a framework where the "double-counting" idea is basically automatic. In other words, linearity of expectation lets us only focus on small, local components when computing an expected value, without having to think about why it works.

Here is another example which captures the same boiler-plate.

Example 73 (HMMT 2006)**.** At a nursery, 2006 babies sit in a circle. Suddenly, each baby randomly pokes either the baby to its left or to its right. What is the expected value of the number of unpoked babies?

Solution 73. Number the babies 1, 2, …, 2006. Define

$$X_i \coloneqq \begin{cases} 1 & \text{if baby } i \text{ is unpoked} \\ 0 & \text{otherwise.} \end{cases}$$

We seek $\mathbb{E}[X_1+X_2+\cdots+X_{2006}]$. Note that any particular baby has probability $\left(\frac{1}{2}\right)^2 = \frac{1}{4}$ of being unpoked (if both its neighbors miss). Hence $\mathbb{E}[X_i] = \frac{1}{4}$ for each i, and

$$\begin{aligned}\mathbb{E}[X_1 + X_2 + \cdots + X_{2006}] &= \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_{2006}] \\ &= 2006 \cdot \frac{1}{4} \\ &= \frac{1003}{2}\end{aligned}$$

which is the answer. ■

Seriously, this should feel like cheating.

§7.3 The so-called pigeonhole principle

In its simplest form, we can use expected value to show existence as follows: suppose we know that the average score of the USAMO 2014 was 12.51. Then there exists a contestant who got at least 13 points, and a contestant who got at most 12 points.

This is isomorphic to the pigeonhole principle, but the probabilistic phrasing is far more robust.

Example 74 (International Math Competition 2002)**.** An olympiad has six problems and 200 contestants. The contestants are very skilled, so each problem is solved by at least 120 of the contestants. Prove that there exist two contestants such that each problem is solved by at least one of them.

Solution 74. We randomly pick two contestants (possibly even the same contestant). Note that the probability they both miss the first problem is at most $\left(\frac{2}{5}\right)^2 = \frac{4}{25}$. So the expected value of the number of problems that both miss is at most $6 \cdot \frac{4}{25} < 1$. Therefore there certainly *exists* a pair of students who together miss less than one problem, which is what we wanted. ■

You will also see applications of pigeonhole that are the same in spirit, but which are not formulated in terms of probability.

§7.4 Walkthroughs

Problem 75 (Canadian Olympiad 2006)**.** In a tournament on $2k+1$ vertices, find the minimum and maximum possible number of directed triangles.

Walkthrough. The minimum bound is not that interesting.

(a) Give an example of a tournament with no directed triangle. This finds the minimum.

It's the maximum that we'll be most interested in. In order to count it, it will actually be parametrize our target in terms of degrees.

(b) Rephrase the "maximum" problem in terms of the number of *non-cyclic* triplets.

(c) By double-counting, find an expression for the number of non-cyclic triplets in terms of the outdegrees of the vertices. (Possible hint: every non-cyclic triplet can be labeled vwx with $v \to w$, $v \to x$.)

Thus we are reduced to an algebraic calculation.

(d) Use Jensen's inequality to show there are at least $(2k+1)\binom{k}{2}$ non-cyclic triplets.

(e) Give an example where equality holds; thus the maximum is $\binom{2k+1}{3} - (2k+1)\binom{k}{2}$.

Problem 76 (IMO Shortlist 2016)**.** Let $n \geq 5$ be a positive integer such that $\gcd(n, 6) = 1$. We color the vertices of a regular n-gon either red, blue, or black, such that each color is used on an odd number of vertices. Prove that there exists an isosceles triangle whose vertices are all different colors.

Walkthrough. This is almost a canonical double-counting problem, in that if you decide to try and write down some equations which count the data in two ways, then there is only really one thing you can write down, and unsurprisingly it works. The problem leaves some visible clues this is what you should be doing, such as:

(a) Use the condition $\gcd(n, 6) = 1$ to show that every segment is the side of three distinct isosceles triangles.

This is one big hint that a double-counting approach will work, as is the fact that each color is used an odd number of times. In fact, the solution here will get us that the number of rainbow isosceles triangles is odd. This hints the only obstructions are "global mod 2", whatever that means.

Let Y denote the number of monochromatic isosceles triangles, and X the remaining isosceles triangles. Let a, b, c denote the number of vertices of each color.

(b) Find $X + Y$ in terms of n.

(c) Express $N = \binom{a}{2} + \binom{b}{2} + \binom{c}{2}$ in terms of X and Y. (There is essentially only one way to do this.)

(d) Using the answer to (c), show that $N \equiv X + Y \pmod 2$.

(e) Use this to derive a contradiction.

A solution using $ab+bc+ca$ instead of $\binom{a}{2}+\binom{b}{2}+\binom{c}{2}$ is possible (and cleaner). It's actually not hard, now that we have this solution, to establish more.

(f) Show more strongly that the number of rainbow isosceles triangles is odd.

(g) Show that the condition $\gcd(n, 6) = 1$ can be dropped entirely — even without it, the number of rainbow isosceles triangles is still odd. (Of course, the problem is only interesting for n odd.)

§7.5 Problems

Problem 77 (HMMT February 2013)**.** Values $a_1, \ldots, a_{2013}$ are chosen independently and at random from the set $\{1, \ldots, 2013\}$. What is the expected number of distinct values in the set $\{a_1, \ldots, a_{2013}\}$?

Problem 78 (AIME 1985)**.** In a tournament each player played exactly one game against each of the other players. In each game the winner was awarded 1 point, the loser got 0 points, and each of the two players earned 1/2 point if the game was a tie. After the completion of the tournament, it was found that exactly half of the points earned by each player were earned against the ten players with the least number of points. (In particular, each of the ten lowest scoring players earned half of her/his points against the other nine of the ten.) What was the total number of players in the tournament?

Problem 79 (Bay Area Olympiad 2013)**.** For a positive integer $n > 2$, consider the $n-1$ fractions $\frac{2}{1}$, $\frac{3}{2}$, ..., $\frac{n}{n-1}$. The product of these fractions equals n, but if you reciprocate (i.e. turn upside down) some of the fractions, the product will change. For which n can the product be made into 1?

Problem 80 (ELMO 2015)**.** Let m, n, and x be positive integers. Prove that

$$\sum_{i=1}^{n} \min\left(\left\lfloor \frac{x}{i} \right\rfloor, m\right) = \sum_{i=1}^{m} \min\left(\left\lfloor \frac{x}{i} \right\rfloor, n\right).$$

Problem 81 (Russia 1996)**.** In the Duma there are 1600 delegates, who have formed 16000 committees of 80 people each. Prove that one can find two committees having no fewer than four common members.

Problem 82 (IMO 1998)**.** In a competition, there are a contestants and b judges, where $b \geq 3$ is an odd integer. Each judge rates each contestant as either "pass" or "fail". Suppose k is a number such that for any two judges, their ratings coincide for at most k contestants. Prove that

$$\frac{k}{a} \geq \frac{b-1}{2b}.$$

Problem 83 (USAMO 2012)**.** A circle is divided into congruent arcs by 432 points. The points are colored in four colors such that some 108 points are colored red, some 108 points are colored green, some 108 points are colored blue, and the remaining 108 points are colored yellow. Prove that one can choose three points of each color in such a way that the four triangles formed by the chosen points of the same color are congruent.

Problem 84 (IMO 2016)**.** Find all integers n for which each cell of $n \times n$ table can be filled with one of the letters I, M and O in such a way that:

- In each row and column, one third of the entries are I, one third are M and one third are O; and
- in any diagonal, if the number of entries on the diagonal is a multiple of three, then one third of the entries are I, one third are M and one third are O.

Note that an $n \times n$ table has $4n - 2$ diagonals.

Problem 85 (Online Math Open 2013)**.** Kevin has $2^n - 1$ cookies, each labeled with a unique nonempty subset of $\{1, 2, \ldots, n\}$. Each day, he chooses one cookie uniformly at random out of the cookies not yet eaten. Then, he eats that cookie, and all remaining cookies that are labeled with a subset of that cookie. Determine the expected value of the number of days that Kevin eats a cookie before all cookies are gone.

Problem 86 (IMO 2005)**.** In a mathematical competition 6 problems were posed to the contestants. Each pair of problems was solved by more than $\frac{2}{5}$ of the contestants. Nobody solved all 6 problems. Show that there were at least 2 contestants who each solved exactly 5 problems.

§7.6 Solutions

Solution 75. The minimum is clearly zero — consider a tournament where there are players of different skill levels, and no upsets.

For the maximum, count the number of non-cyclic triplets. In any non-cyclic triplet there is exactly one vertex dominating the other two. So the number of non-cyclic triplets is equal to

$$\sum_v \binom{\text{outdeg } v}{2}$$

which by Jensen is at least $(2k+1)\binom{k}{2}$. Hence the answer is $\binom{2k+1}{3} - (2k+1)\binom{k}{2}$.

Solution 76. Observe that from $\gcd(n, 6) = 1$ we find there are no equilateral triangles, and

- Every segment is the base of exactly one isosceles triangle.
- Every segment is the left leg of exactly one isosceles triangle.
- Every segment is the right leg of exactly one isosceles triangle.

Now, assume for contradiction there are no rainbow isosceles triangles. Let Y be the number of monochromatic isosceles triangles, and X the number of isosceles triangles with two vertices of one color and the last vertex of a different color.

Let a, b, c be the number of vertices of each color. On the one hand, we have

$$X + 3Y = 3\left[\binom{a}{2} + \binom{b}{2} + \binom{c}{2}\right]$$

just by double-counting the triangles: the conditions $\gcd(n, 6) = 1$ imply that exactly three isosceles triangles use any given edge. (To be precise, we are counting pairs $(\triangle, e)$, where $\triangle$ is isosceles and has edge e with matching colors. The left-hand side counts by $\triangle$ while the right-hand side counts by e.)

On the other hand, we have

$$X + Y = \binom{n}{2}$$

since an isosceles triangle is determined by its base (again since $\gcd(n, 6) = 1$). Therefore we have the following equality modulo 2:

$$\binom{n}{2} \equiv \binom{a}{2} + \binom{b}{2} + \binom{c}{2} \pmod 2.$$

Doubling and expanding we get

$$\begin{aligned} n^2 - n &\equiv (a^2 - a) + (b^2 - b) + (c^2 - c) \pmod 4 \\ &\equiv a^2 + b^2 + c^2 - n \pmod 4. \end{aligned}$$

But since n, a, b, c are odd this is impossible.

Remark. One can equally well work with $ab+bc+ca$ instead of $\binom{a}{2}+\binom{b}{2}+\binom{c}{2}$, which turns out be somewhat cleaner technically, although it feels less natural to me.

Remark. I think the conditions $\gcd(n, 6) = 1$ and a, b, c odd are huge giveaways that this will be "global obstructions modulo 2". I solved this during a Synco concert.

Remark. In fact one can drop the condition that $\gcd(n, 3) = 1$. Indeed, the only change is that fixing any two vertices A, B, either exactly one or exactly three isosceles triangles pass through them (instead of always exactly three). These are the same modulo two anyways.

Solution 77. The probability a given $1 \le n \le 2013$ appears is $1-(2012/2013)^{2013}$. Thus by linearity of expectation, the answer is

$$2013\left(1 - \left(\frac{2012}{2013}\right)^{2013}\right).$$

Solution 78. Call the players n strong players and 10 weak players. There are three categories of points given:

- In games between two strong players, a total of $A = \binom{n}{2}$ points are given out.
- In games between a strong player and a weak player, a total of $B = 10n$ points are given out.
- In games between two weak players, a total of $C = \binom{10}{2}$ points are given out.

Breaking down further, the points in category B are divided into two subcategories:

- Points given where strong players beat weak players. Each individual strong player earns the same number of points from this category as points in category A. So summing over all strong players, there are $A = \binom{n}{2}$ points in this sub-category.

- Points given where weak players beat strong players. Each individual weak player earns the same number of points from this category as points in category C. So summing over all weak players, there are $C = \binom{10}{2}$ points in this sub-category.

So, we conclude

$$A + C = B.$$

Solving gives $n^2 - 21n + 90 = 0$, so either $n = 6$ or $n = 15$. Finally, note that the strong players have an average score of $2A/n$ while the weak players have an average score of $2C/10$. As $2A/n > 2C/10$, we require $n > 10$. So $n = 15$ and $n + 10 = \boxed{25}$.

Solution 79. The answer is iff n a perfect square.

To see such n work, just take

$$\left(\frac{1}{2} \times \frac{2}{3} \times \ldots \frac{\sqrt{n}-1}{\sqrt{n}}\right) \times \left(\frac{\sqrt{n}+1}{\sqrt{n}} \times \cdots \times \frac{n}{n-1}\right) = \frac{1}{\sqrt{n}} \times \sqrt{n} = 1$$

as a construction.

Now we show n must be a square. If we divide

$$\frac{2}{1} \times \frac{3}{2} \times \cdots \times \frac{n}{n-1} = n$$

with a construction equal to 1, we will get

$$\left(\frac{2}{1}\right)^{1+\varepsilon_1} \times \left(\frac{3}{2}\right)^{1+\varepsilon_2} \times \cdots \times \left(\frac{n}{n-1}\right)^{1+\varepsilon_{n-1}} = n$$

where $\varepsilon_i \in \{-1, 1\}$ for each i. Since $1 + \varepsilon_i$ is even for each i, this means n is the square of a rational number. Thus n must be itself a perfect square.

Solution 80. Construct a $m \times n$ multiplication table: both sides are counting the number of terms $\le x$. In other words, both sides count the number of ordered pairs

$$(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, m\} \quad \text{such that} \quad ij \le x.$$

Alternatively, induction on x works fine.

Solution 81. Let a_1, ..., a_{1600} denote the number of committees the ith delegate is in. Thus $\sum a_i = 16000 \cdot 80$, and $\frac{1}{1600} \sum a_i = 800$.

Now sample a random pair of committees. The expected number of common members is

$$\begin{aligned}\frac{\sum \binom{a_i}{2}}{\binom{16000}{2}} &\overset{\text{Jensen}}{\geq} \frac{1600 \cdot \binom{800}{2}}{\binom{16000}{2}} \\ &= \frac{1600 \cdot 800 \cdot 799}{16000 \cdot 15999} \\ &= \frac{800 \cdot 799}{10 \cdot 15999} > 3\end{aligned}$$

(in fact quite close to 4).

Solution 82. This is a "routine" problem with global ideas. We count pairs of coinciding ratings, i.e. the number N of tuples

$$(\{J_1, J_2\}, C)$$

of two distinct judges and a contestant for which the judges gave the same rating.

On the one hand, if we count by the judges, we have

$$N \leq \binom{b}{2} k$$

by he problem condition.

On the other hand, if $b = 2m + 1$, then each contestant C contributes at least $\binom{m}{2} + \binom{m+1}{2} = m^2$ to N, and so

$$N \geq a \cdot \left(\frac{b-1}{2}\right)^2$$

Putting together the two estimates for N yields the conclusion.

Solution 83. First, consider the 431 possible non-identity rotations of the red points, and count overlaps with green points. If we select a rotation randomly, then each red point lies over a green point with probability $\frac{108}{431}$; hence the expected number of red-green incidences is

$$\frac{108}{431} \cdot 108 > 27$$

and so by pigeonhole, we can find a red 28-gon and a green 28-gon which are rotations of each other.

Now, look at the 430 rotations of this 28-gon (that do not give the all-red or all-green configuration) and compare it with the blue points. The same approach gives

$$\frac{108}{430} \cdot 28 > 7$$

incidences, so we can find red, green, blue 8-gons which are similar under rotation.

Finally, the 429 nontrivial rotations of this 8-gon expect

$$\frac{108}{429} \cdot 8 > 2$$

incidences with yellow. So finally we have four monochromatic 3-gons, one of each color, which are rotations of each other.

Solution 84. The answer is n divisible by 9.

First we construct $n = 9$ and by extension every multiple of 9.

I	I	I	M	M	M	O	O	O
M	M	M	O	O	O	I	I	I
O	O	O	I	I	I	M	M	M
I	I	I	M	M	M	O	O	O
M	M	M	O	O	O	I	I	I
O	O	O	I	I	I	M	M	M
I	I	I	M	M	M	O	O	O
M	M	M	O	O	O	I	I	I
O	O	O	I	I	I	M	M	M

We now prove $9 \mid n$ is necessary.

Let $n = 3k$, which divides the given grid into k^2 sub-boxes (of size 3×3 each). We say a multiset of squares S is *clean* if the letters distribute equally among them; note that unions of clean multisets are clean.

Consider the following clean sets (given to us by problem statement):

- All columns indexed $2 \pmod 3$,
- All rows indexed $2 \pmod 3$, and
- All $4k - 2$ diagonals mentioned in the problem.

Take their union. This covers the center of each box four times, and every other cell exactly once. We conclude the set of k^2 center squares are clean, hence $3 \mid k^2$ and so $9 \mid n$, as desired.

Shown below is the sums over all diagonals only, and of the entire union.

1		1	1		1	1		1
	2			2			2	
1		1	1		1	1		1
1		1	1		1	1		1
	2			2			2	
1		1	1		1	1		1
1		1	1		1	1		1
	2			2			2	
1		1	1		1	1		1

1	1	1	1	1	1	1	1	1
1	4	1	1	4	1	1	4	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	4	1	1	4	1	1	4	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	4	1	1	4	1	1	4	1
1	1	1	1	1	1	1	1	1

Solution 85. The key insight is that the number of days that elapse is exactly equal to the number of cookies that are *chosen*. Thus we can compute the probability each given cookie is chosen (which is easy, since if a cookie is alive so are all supersets), and sum using linearity of expectation.

Given a cookie labelled with S, the probability it is chosen at all is $1/2^{n-|S|}$, and the expected value of the number of days that pass is (by linearity) the sum of all these. By the Binomial Theorem, we obtain an answer of

$$\begin{aligned}\sum_{\text{cookie } C} \mathbb{P}(C \text{ chosen}) &= \sum_{C} \frac{1}{2^{n-|C|}} \\ &= \sum_{k=1}^{n} \frac{\binom{n}{k}}{2^{n-k}} \\ &= \left(\frac{3}{2}\right)^n - \left(\frac{1}{2}\right)^n.\end{aligned}$$

We obtain an answer $\frac{3^n-1}{2^n}$.

Note that omitting $\varnothing$ is red herring, since one can read off that it adds 2^{-n}. It was added just to make answer harder to guess.

Solution 86. Assume not and at most one contestant solved five problems. By adding in solves, we can assume WLOG that one contestant solved problems one through five, and every other contestant solved four of the six problems.

We split the remaining contestants based on whether they solved P6. Let a_i denote the number of contestants who solved $\{1, 2, \dots, 5\} \setminus \{i\}$ (and missed P6). Let b_{ij} denote the number of contestants who solved $\{1, 2, \dots, 5, 6\} \setminus \{i, j\}$, for $1 \le i < j \le 5$ (thus in particular they solved P6). Thus

$$n = 1 + \sum_{1 \le i \le 5} a_i + \sum_{1 \le i < j \le 5} b_{ij}$$

denotes the total number of contestants.

Considering contestants who solved P1/P6 we have

$$t_1 \coloneqq b_{23} + b_{24} + b_{25} + b_{34} + b_{35} + b_{45} \ge \frac{2}{5}n + \frac{1}{5}$$

and we similarly define t_2, t_3, t_4, t_5. (We have written $\frac{2}{5}n + \frac{1}{5}$ since we know the left-hand side is an integer strictly larger than $\frac{2}{5}n$.) Also, by considering contestants who solved P1/P2 we have

$$t_{12} = 1 + a_3 + a_4 + a_5 + b_{34} + b_{35} + b_{45} \ge \frac{2}{5}n + \frac{1}{5}$$

and we similarly define t_{ij} for $1 \le i < j \le 5$.

Claim. The number $\frac{2n+1}{5}$ is equal to some integer k, fourteen of the t's are equal to k, and the last one is equal to $k+1$.

Proof. First, summing all fifteen equations gives

$$\begin{aligned} 6n+4 = 10+6(n-1) &= 10 + \sum_{1\le i\le 5} 6a_i + \sum_{1\le i<j\le 5} 6b_{ij} \\ &= \sum_{1\le i\le 5} t_i + \sum_{1\le i<j\le 5} t_{ij}. \end{aligned}$$

Thus the sum of the 15 t's is $6n+4$. But since all the t's are integers at least $\frac{2n+1}{5} = \frac{6n+3}{15}$, the conclusion follows. □

However, we will also manipulate the equations to get the following.

Claim. We have

$$t_{45} \equiv 1 + t_1 + t_2 + t_3 + t_{12} + t_{23} + t_{31} \pmod 3.$$

Proof. This follows directly by computing the coefficient of the a's and b's. We will nonetheless write out a derivation of this equation, to motivate it, but the proof stands without it.

Let $B = \sum_{1\le i<j\le 5} b_{ij}$ be the sum of all b's. First, note that

$$\begin{aligned} t_1 + t_2 &= B + b_{34} + b_{45} + b_{35} - b_{12} \\ &= B + (t_{12} - 1 - a_3 - a_4 - a_5) - b_{12} \\ \implies b_{12} &= B - (t_1 + t_2) + t_{12} - 1 - (a_3 + a_4 + a_5). \end{aligned}$$

This means we have more or less solved for each b_{ij} in terms of only t and a variables. Now

$$\begin{aligned} t_{45} &= 1 + a_1 + a_2 + a_3 + b_{12} + b_{23} + b_{31} \\ &= 1 + a_1 + a_2 + a_3 \\ &\quad + [B - (t_1 + t_2) + t_{12} - 1 - (a_3 + a_4 + a_5)] \\ &\quad + [B - (t_2 + t_3) + t_{23} - 1 - (a_1 + a_4 + a_5)] \\ &\quad + [B - (t_3 + t_1) + t_{13} - 1 - (a_2 + a_4 + a_5)] \\ &\equiv 1 + t_1 + t_2 + t_3 + t_{12} + t_{23} + t_{31} \pmod 3 \end{aligned}$$

as desired. □

However, we now show the two claims are incompatible (and this is easy, many ways to do this). There are two cases.

- Say $t_5 = k+1$ and the others are k. Then the equation for t_{45} gives that $k \equiv 6k + 1 \pmod 3$. But now the equation for t_{12} give $k \equiv 6k \pmod 3$.

- Say $t_{45} = k+1$ and the others are k. Then the equation for t_{45} gives that $k+1 \equiv 6k \pmod 3$. But now the equation for t_{12} give $k \equiv 6k+1 \pmod 3$.

Remark. It is significantly easier to prove that there is at least one contestant who solved five problems. One can see it by dropping the +10 in the proof of the claim, and arrives at a contradiction. In this situation it is not even necessary to set up the many a and b variables; just note that the expected number of contestants solving any particular pair of problems is $\frac{\binom{4}{2}n}{\binom{6}{2}} = \frac{2}{5}n$.

The fact that $\frac{2n+1}{5}$ should be an integer also follows quickly, since if not one can improve the bound to $\frac{2n+2}{5}$ and quickly run into a contradiction. Again one can get here without setting up a and b.

The main difficulty seems to be the precision required in order to nail down the second 5-problem solve.

Remark. The second claim may look miraculous, but the proof shows that it is not too unnatural to consider $t_1 + t_2 - t_{12}$ to isolate b_{12} in terms of a's and t's. The main trick is: why mod 3?

The reason is that if one looks closely, for a fixed k we have a system of 15 equations in 15 variables. Unless the determinant D of that system happens to be zero, this means there will be a rational solution in a and b, whose denominators are bounded by D. However if $p \mid D$ then we may conceivably run into mod p issues.

This motivates the choice $p = 3$, since it is easy to see the determinant is divisible by 3, since constant shifts of $\vec{a}$ and $\vec{b}$ are also solutions mod 3. (The choice $p = 2$ is a possible guess as well for this reason, but the problem seems to have better 3-symmetry.)

---

8 Local

§8.1 Synopsis

Last chapter I talked about using a "global" argument, looking at the entire problem at once and e.g. summing, in order to get a problem. This lecture is about the opposite idea: **making small disturbances** in a concentrated area.

Unlike the last chapter, there is no theory I want to develop, and it will be better for you to jump straight into the walkthroughs after this short interruption.

The idea is to look at a *small part of the structure*, rather than the whole problem at once. The most common thing to do is to then *perturb* the problem a little by making a small disturbance. Then, **repeat until stuck**.

Examples of this idea:

- (Greedy) algorithms: for example, to find n objects satisfying so-and-so condition, one might try to grab them one at a time without causing any issues. Then look at the situation in which we can't add any more.
- Extremal principle: often used as a write-up mechanic, e.g. many greedy algorithms may be rephrased in terms of "look at a maximal set".
- Smoothing: inequalities technique involving moving variables. For example, if we find that we can replace two variables a and b with their average, the resulting inequality still must be true.

  As a simple example, suppose we want to prove $abc \le 1$ subject to $a, b, c > 0$ and $a + b + c = 3$. Note that if $a > b$, then one can replace (a, b) with $(a-\varepsilon, b+\varepsilon)$ for $\varepsilon < \frac{1}{2}(a-b)$, and this will increase the product. By doing this operation we can arrive at a situation in which $a = b = c$ at which point the inequality is obviously true.

In many cases the small changes made follow a heuristic, like in a greedy algorithm. In the best cases they are optimizations in the sense that the problem has to "remain true" after the operation.

§8.2 Walkthroughs

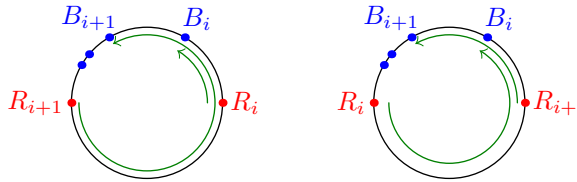
Problem 87 (USAMO 2017)**.** Let P_1, P_2, ..., P_{2n} be $2n$ distinct points on the unit circle $x^2 + y^2 = 1$, other than $(1, 0)$. Each point is colored either red or blue, with exactly n red points and n blue points. Let R_1, R_2, ..., R_n

be any ordering of the red points. Let B_1 be the nearest blue point to R_1 traveling counterclockwise around the circle starting from R_1. Then let B_2 be the nearest of the remaining blue points to R_2 travelling counterclockwise around the circle from R_2, and so on, until we have labeled all of the blue points B_1, ..., B_n. Show that the number of counterclockwise arcs of the form $R_i \to B_i$ that contain the point $(1, 0)$ is independent of the way we chose the ordering R_1, ..., R_n of the red points.

Walkthrough. There is actually a fairly nice characterization of the number of such arcs, which leads to a satisfying "rigid" characterization of the number. (You can try to find it yourself, or read the official solution.)

However, suppose one does not care about having a satisfying solution, and only cares about having *any* solution. Then it is much easier to simply mess with the permutation $(R_i)_i$ a small amount, and show that the number of arcs does not change. And in fact, this is much easier to do.

(a) Make sense of the following picture, where we swap R_i and R_{i+1}.



(b) Figure out what the possible other pictures might look like, if we try to swap R_i and R_{i+1}.

(c) Draw the other two cases, and verify that swapping R_i and R_{i+1} does not change the number of arcs containing any point on the circumference.

Then we "repeat until stuck"; fortunately in this case we don't get stuck at all, and simply solve the problem outright.

(d) Show that the operation of (c) is enough to imply the problem. (You may have seen this before.)

(e) Does the same proof work if you try to swap R_i and R_j for any $i < j$?

In some sense this solution should seem unsatisfying, since we still do not really understand why the result "should be true". However, it is strictly speaking a correct solution.

Problem 88. Suppose 4951 distinct points in the plane are given such that no four points are collinear. Show that it is possible to select 100 of the points for which no three points are collinear.

Walkthrough. Following RUST, keep grabbing points until we cannot take any more. Suppose at this point we have n points.

(a) Show that $4951 - n \le \binom{n}{2}$.

(b) Prove that $n \ge 100$.

So this is an example of a greedy algorithm of the most direct sort.

Problem 89 (Putnam 1979)**.** Given n red points and n blue points in the plane, no three collinear, prove that we can draw n segments, each joining a red point to a blue point, such that no segments intersect.

Walkthrough. Starting from an arbitrary configuration, we will use the algorithm "given a crossing, un-cross it". This is a very natural algorithm to come up with, and playing with some simple examples one finds that it always work. So we just have to prove that.

(a) Show that a step of this algorithm does *not* necessarily decrease the total number of intersections. (But this is the first thing we *should* try, given that our goal is to get zero intersections at the end.)

(b) Find a different monovariant M which *does* decrease at each step of the algorithm.

(c) Remark on the finiteness of the configuration space, and complete the problem using (b).

I want to say a few words about why I chose this example. This problem is touted in olympiad cultures as an example of "extremal principle", with "choose the minimal M" as the poster description. In my humble opinion, I think this is hogwash. The motivation should be the natural algorithm we used; the monovariant comes after the fact.

Indeed, the fact that the natural guess of the monovariant in (a) fails is what makes this problem a little interesting (and not completely standard). However, it doesn't change the fact that the algorithm comes before the monovariant in our thought process.

§8.3 Problems

Problem 90 (Princeton Competition 2013)**.** Let G be a graph and let k be a positive integer. A *k-star* is a set of k edges with a common endpoint and a *k-matching* is a set of k edges such that no two have a common endpoint. Prove that if G has more than $2(k-1)^2$ edges then it either has a k-star or a k-matching.

Problem 91 (IMO Shortlist 2013)**.** Let n be a positive integer. Find the smallest integer k (in terms of n) with the following property: given any finite multiset of real numbers in $[0, 1]$ whose sum is n, 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.

Problem 92. Let G be a finite simple graph. Show that one can partition the vertices into two groups such that for each vertex, at least half the neighbors are in the other group.

Problem 93 (IMO 2003)**.** Let A be a 101-element subset of $S = \{1, 2, \dots, 10^6\}$. Prove that there exist numbers t_1, $t_2, \dots, t_{100}$ in S such that the sets

$$A_j = \{x + t_j \mid x \in A\}, \qquad j = 1, 2, \dots, 100$$

are pairwise disjoint.

Problem 94. Let G be a finite simple graph with $m > 0$ edges and $n > 1$ vertices. Show that one can delete some number of vertices of G to obtain a graph with at least one vertex whose minimum degree is at least m/n.

Problem 95 (USA TST 2017)**.** In a sports league, each team uses a set of at most t signature colors. A set S of teams is *color-identifiable* if one can assign each team in S one of their signature colors, such that no team in S is assigned *any* signature color of a different team in S. For all positive integers n and t, determine the maximum integer $g(n, t)$ such that: In any sports league with exactly n distinct colors present over all teams, one can always find a color-identifiable set of size at least $g(n, t)$.

Problem 96 (IMO 2014)**.** A set of lines in the plane is in *general position* if no two are parallel and no three pass through the same point. A set of lines in general position cuts the plane into regions, some of which have finite area; we call these its *finite regions*. Prove that for all sufficiently large n, in any set of n lines in general position it is possible to colour at least $\sqrt{n}$ lines blue in such a way that none of its finite regions has a completely blue boundary.

Problem 97 (USA TST 2018)**.** At a university dinner, there are 2017 mathematicians who each order two distinct entrées, with no two mathematicians ordering the same pair of entrées. The cost of each entrée is equal to the number of mathematicians who ordered it, and the university pays for each mathematician's less expensive entrée (ties broken arbitrarily). Over all possible sets of orders, what is the maximum total amount the university could have paid?

§8.4 Solutions

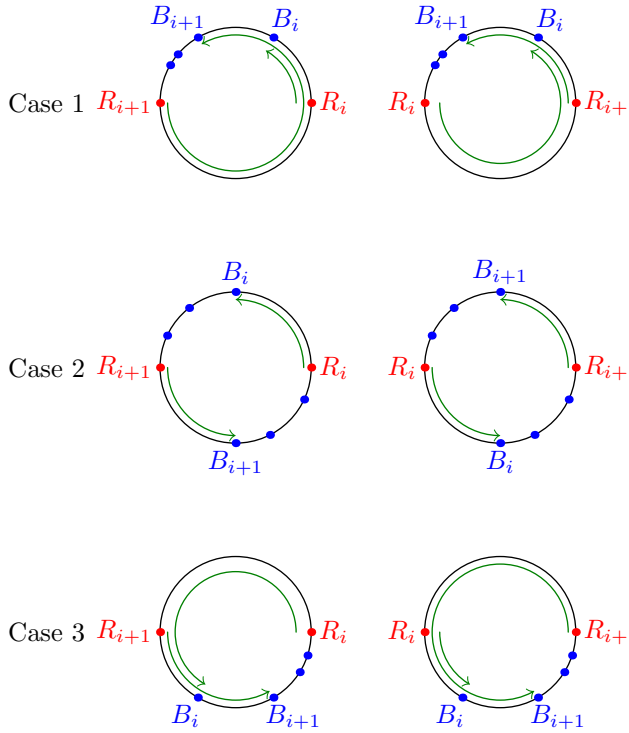
Solution 87. We present two solutions, one based on swapping and one based on an invariant.

First "local" solution by swapping two points Let $1 \le i < n$ be any index and consider the two red points R_i and R_{i+1}. There are two blue points B_i and B_{i+1} associated with them.

Claim. If we swap the locations of points R_i and R_{i+1} then the new arcs $R_i \to B_i$ and $R_{i+1} \to B_{i+1}$ will cover the same points.

Proof. Delete all the points R_1, ..., R_{i-1} and B_1, ..., B_{i-1}; instead focus on the positions of R_i and R_{i+1}.

The two blue points can then be located in three possible ways: either 0, 1, or 2 of them lie on the arc $R_i \to R_{i+1}$. For each of the cases below, we illustrate on the left the locations of B_i and B_{i+1} and the corresponding arcs in green; then on the right we show the modified picture where R_i and R_{i+1} have swapped. (Note that by hypothesis there are no other blue points in the green arcs).

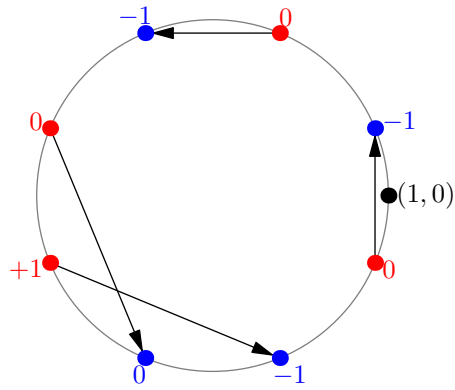


Observe that in all cases, the number of arcs covering any given point on the circumference is not changed. Consequently, this proves the claim. □

Finally, it is enough to recall that any permutation of the red points can be achieved by swapping consecutive points (put another way: $(i\ i+1)$ generates the permutation group S_n). This solves the problem.

Remark. This proof does *not* work if one tries to swap R_i and R_j if $|i-j| \neq 1$. For example if we swapped R_i and R_{i+2} then there are some issues caused by the possible presence of the blue point B_{i+1} in the green arc $R_{i+2} \to B_{i+2}$.

Second longer solution using an invariant Visually, if we draw all the segments $R_i \to B_i$ then we obtain a set of n chords. Say a chord is *inverted* if satisfies the problem condition, and *stable* otherwise. The problem contends that the number of stable/inverted chords depends only on the layout of the points and not on the choice of chords.



In fact we'll describe the number of inverted chords explicitly. Starting from $(1,0)$ we keep a running tally of $R - B$; in other words we start the counter at 0 and decrement by 1 at each blue point and increment by 1 at each red point. Let $x \leq 0$ be the lowest number ever recorded. Then:

Claim. The number of inverted chords is $-x$ (and hence independent of the choice of chords).

This is by induction on n. I think the easiest thing is to delete chord R_1B_1; note that the arc cut out by this chord contains no blue points. So if the chord was stable certainly no change to x. On the other hand, if the chord is inverted, then in particular the last point before $(1,0)$ was red, and so $x < 0$. In this situation one sees that deleting the chord changes x to $x+1$, as desired.

Solution 88. This is an example of a direct greedy algorithm: we will simply grab points until we are stuck.

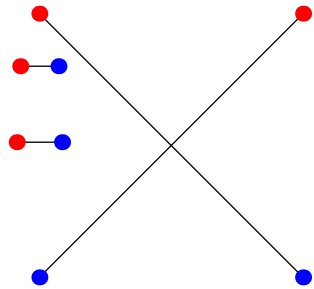
Consider a maximal set S of the points as described (meaning no more additional points can be added), and suppose $|S| = n$. Then the $4951 - n$ other points must each lie on a line determined by two points in S, meaning

$$4951 - n \le \binom{n}{2} \implies n + \binom{n}{2} \ge 4951.$$

This requires $n \ge 100$.

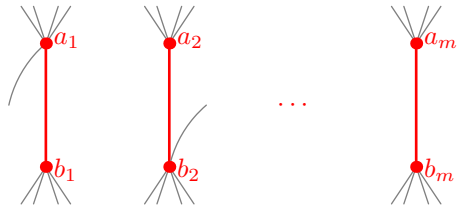
Solution 89. The idea is that given any two segments which cross, then we can un-cross them.

Unfortunately, this does not necessarily decrease the number of intersections, but it *does* decrease the sum of the Euclidean lengths. Hence this serves as a monovariant that shows the "uncross any intersection" algorithm works.



In other words, if we take the connection for which the sum of the lengths is minimal, then there will necessarily be no intersections.

Solution 90. Assume for contradiction there is neither a k-matching nor a k-star. Take a maximal matching of size $m \le k - 1$.



Every edge must touch an edge in the matching. But the degrees of all vertices are all bounded by $k - 1$, so the number of edges not in the matching is at most $2m(k-2)$. Hence the total number of edges in G is

$$2m(k-2) + m \le 2(k-1)(k-2) + (k-1) = (k-1)(2k-3) < 2(k-1)^2$$

contradicting the hypothesis.

Solution 91. Answer: $k = 2n - 1$. To see that at least $2n - 1$ groups may be necessary, take $2n - 1$ copies of the number $\frac{n}{2n-1}$.

To see $2n - 1$ groups is sufficient, consider a minimal partitions into groups with sums $g_1 \le g_2 \le \cdots \le g_m$. That implies that $g_i + g_j > 1$ for any distinct i, j (otherwise we could merge those two groups together). Moreover, $g_1 + \cdots + g_m = n$. Then

$$2n = (g_1 + g_2) + (g_2 + g_3) + \cdots + (g_m + g_1) > 1 + 1 + \cdots + 1 = m$$

as required.

Solution 92. Take a partition $A \cup B$ which maximizes the number of edges between A and B.

Then this partition works, because if a vertex $v \in A$ has the property that less than half its neighbors are in B, we could move v from A to B and increase the number of edges.

Remark. Equivalently (and more naturally): start with an arbitrary partition $A \cup B$, and move a vertex if more than half its neighbors are in the same set. This increases the number of edges, so eventually this process terminates, and it terminates at a desired partition.

Solution 93. A greedy algorithm works: suppose we have picked

$$T = \{t_1, \ldots, t_n\}$$

as large as possible, meaning it's impossible to add any more elements to T. That means, for each $t \in \{1, \ldots, 10^6\}$ either $t \in T$ already or there exists two distinct elements $a, b \in A$ and $t_i \in T$ such that

$$t = t_i + b - a \qquad (\star).$$

There are at most $|T| \cdot |A| \cdot (|A| - 1) = n \cdot 101 \cdot 100$ possible values for the right-hand side of $(\star)$. So we therefore must have

$$101 \cdot 100 \cdot n + n \ge 10^6$$

which implies $n > 99$, as desired.

Remark. It is possible to improve the bound significantly with a small optimization; rather than adding any t, we require that $t_1 < \cdots < t_n$ and that at each step we add the *least* $t \in S$ which is permitted. In that case, one finds we only need to consider $b > a$ in $(\bigstar)$, and so this will essentially save us a factor of $2 + o(1)$ as the main term $101 \cdot 100$ becomes $\binom{101}{2}$ instead.

Solution 94. For a graph Γ, we let $a(\Gamma)$ denote the average degree. A vertex of Γ is called *deficient* if its degree is less than $\frac{1}{2}a(\Gamma)$.

Claim. Let Γ be any graph, and suppose it has a deficient vertex v. Then $\Gamma - v$ has average degree at least that of Γ.

Proof. Deleting a deficient vertex does not decrease the average degree, since the new graph has average degree at least

$$a(\Gamma - v) \ge 2 \cdot \frac{\#E(\Gamma) - \#E(\Gamma)/\#V(\Gamma)}{\#V(\Gamma) - 1} = 2\frac{\#E(\Gamma)}{\#V(\Gamma)} = a(\Gamma). \qquad \square$$

Thus if we repeatedly delete a single deficient vertex from G, we get a sequence of graphs $G = G_0 \supset G_1 \supset \dots$ with

$$a(G_0) \le a(G_1) \le a(G_2) \le \dots.$$

This process terminates only when we have a graph $H = G_N$ with no deficient vertices. Since the graph with one vertex has average degree zero, H is nonempty, and every vertex of H has degree at least $\frac{1}{2}a(H) \ge \frac{1}{2}a(G)$.

Remark. Note that one *must* delete vertices with degree $\le m/n$ anyways, which motivates this solution.

Solution 95. Answer: $\lceil n/t \rceil$.

To see this is an upper bound, note that one can easily construct a sports league with that many teams anyways.

A quick warning:

Remark (Misreading the problem)**.** It is common to misread the problem by ignoring the word "any". Here is an illustration.

Suppose we have two teams, MIT and Harvard; the colors of MIT are red/grey/black, and the colors of Harvard are red/white. (Thus $n = 4$ and $t = 3$.) The assignment of MIT to grey and Harvard to red is *not* acceptable because red is a signature color of MIT, even though not the one assigned.

We present two proofs of the lower bound.

Approach by deleting teams (Gopal Goel) Initially, place all teams in a set S. Then we repeat the following algorithm:

> If there is a team all of whose signature colors are shared by some other team in S already, then we delete that team.

(If there is more than one such team, we pick arbitrarily.)

At the end of the process, all n colors are still present at least once, so at least $\lceil n/t \rceil$ teams remain. Moreover, since the algorithm is no longer possible, the remaining set S is already color-identifiable.

Remark (Gopal Goel)**.** It might seem counter-intuitive that we are *deleting* teams from the full set when the original problem is trying to get a large set S.

This is less strange when one thinks of it instead as "safely deleting useless teams". Basically, if one deletes such a team, the problem statement implies that the task must still be possible, since $g(n, t)$ does not depend on the number of teams: n is the number of colors present, and deleting a useless team does not change this. It turns out that this optimization is already enough to solve the problem.

Approach by adding colors For a constructive algorithmic approach, the idea is to greedy pick by color (rather than by team), taking at each step the least used color. Select the color C_1 with the *fewest* teams using it, and a team T_1 using it. Then delete all colors T_1 uses, and all teams which use C_1. Note that

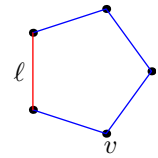
- By problem condition, this deletes at most t teams total.
- Any remaining color C still has at least one user. Indeed, if not, then C had the same set of teams as C_1 did (by minimality of C), but then it should have deleted as a color of T_1.

Now repeat this algorithm with C_2 and T_2, and so on. This operations uses at most t colors each time, so we select at least $\lceil n/t \rceil$ colors.

Remark. A greedy approach by team *does not work*. For example, suppose we try to "grab teams until no more can be added".

As before, assume our league has teams, MIT and Harvard; the colors of MIT are red/grey/black, and the colors of Harvard are red/white. (Thus $n = 4$ and $t = 3$.) If we start by selecting MIT and red, then it is impossible to select any more teams; but $g(n, t) = 2$.

Solution 96. Suppose we have colored k of the lines blue, and that it is not possible to color any additional lines. That means any of the $n - k$ non-blue lines is the side of some finite region with an otherwise entirely blue perimeter. For each such line ℓ, select one such region, and take the next counterclockwise vertex; this is the intersection of two blue lines v. We'll say ℓ is the *eyelid* of v.

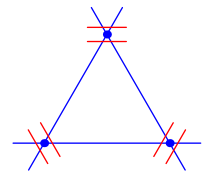


You can prove without too much difficulty that every intersection of two blue lines has at most two eyelids. Since there are $\binom{k}{2}$ such intersections, we see that

$$n - k \le 2\binom{k}{2} = k^2 - k$$

so $n \le k^2$, as required.

Remark. In fact, $k = \sqrt{n}$ is "sharp for greedy algorithms", as illustrated below for $k = 3$:



Solution 97. In graph theoretic terms: we wish to determine the maximum possible value of

$$S(G) \coloneqq \sum_{e=vw} \min(\deg v, \deg w)$$

across all graphs G with 2017 edges. We claim the answer is $63 \cdot \binom{64}{2} + 1 = 127009$.

First solution (combinatorial, Evan Chen) First define L_k to consist of a clique on k vertices, plus a single vertex connected to exactly one vertex of the clique. Hence L_k has $k+1$ vertices, $\binom{k}{2}+1$ edges, and $S(L_k) = (k-1)\binom{k}{2}+1$. In particular, L_{64} achieves the claimed maximum, so it suffices to prove the upper bound.

Lemma. *Let G be a graph such that either*

- *G has $\binom{k}{2}$ edges for some $k \ge 3$ or*
- *G has $\binom{k}{2}+1$ edges for some $k \ge 4$.*

Then there exists a graph G^* with the same number of edges such that $S(G^*) \ge S(G)$, and moreover G^* has a universal vertex (i.e. a vertex adjacent to every other vertex).

Proof. Fix k and the number m of edges. We prove the result by induction on the number n of vertices in G. Since the lemma has two parts, we will need two different base cases:

1. Suppose $n = k$ and $m = \binom{k}{2}$. Then G must be a clique so pick $G^* = G$.

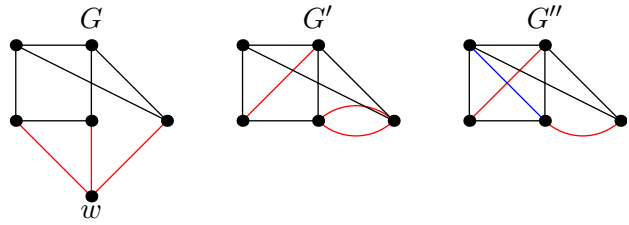
2. Suppose $n = k+1$ and $m = \binom{k}{2} + 1$. If G has no universal vertex, we claim we may take $G^* = L_k$. Indeed each vertex of G has degree at most $k-1$, and the average degree is

$$\frac{2m}{n} = \frac{k^2 - k + 1}{k+1} < k - 1$$

using here $k \geq 4$. Thus there exists a vertex w of degree $1 \leq d \leq k-2$. The edges touching w will have label at most d and hence

$$\begin{aligned} S(G) &\leq (k-1)(m-d) + d^2 = (k-1)m - d(k-1-d) \\ &\leq (k-1)m - (k-2) = (k-1)\binom{k}{2} + 1 = S(G^*). \end{aligned}$$

Now we settle the inductive step. Let w be a vertex with minimal degree $0 \leq d < k-1$, with neighbors $w_1, \ldots, w_d$. By our assumption, for each w_i there exists a vertex v_i for which $v_i w_i \notin E$. Now, we may delete all edges ww_i and in their place put $v_i w_i$, and then delete the vertex w. This gives a graph G', possibly with multiple edges (if $v_i = w_j$ and $w_j = v_i$), and with one fewer vertex.



We then construct a graph G'' by taking any pair of double edges, deleting one of them, and adding any missing edge of G'' in its place. (This is always possible, since when $m = \binom{k}{2}$ we have $n - 1 \geq k$ and when $m = \binom{k}{2} + 1$ we have $n - 1 \geq k + 1$.)

Thus we have arrived at a simple graph G'' with one fewer vertex. We also observe that we have $S(G'') \geq S(G)$; after all every vertex in G'' has degree at least as large as it did in G, and the d edges we deleted have been replaced with new edges which will have labels at least d. Hence we may apply the inductive hypothesis to the graph G'' to obtain G^* with $S(G^*) \geq S(G'') \geq S(G)$. □

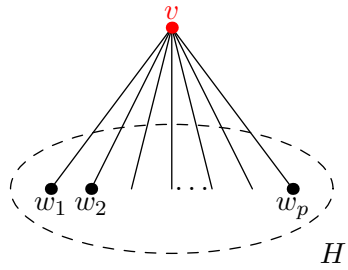
The problem then is completed once we prove the following:

Claim. For any graph G,

- If G has $\binom{k}{2}$ edges for $k \geq 3$, then $S(G) \leq \binom{k}{2} \cdot (k-1)$.
- If G has $\binom{k}{2}+1$ edges for $k \geq 4$, then $S(G) \leq \binom{k}{2} \cdot (k-1) + 1$.

Proof. We prove both parts at once by induction on k, with the base case $k = 3$ being plain (there is nothing to prove in the second part for $k = 3$). Thus assume $k \geq 4$. By the earlier lemma, we may assume G has a universal vertex v. For notational convenience, we say G has $\binom{k}{2}+\varepsilon$ edges for $\varepsilon \in \{0,1\}$, and G has $p+1$ vertices, where $p \geq k-1+\varepsilon$.

Let H be the subgraph obtained when v is deleted. Then $m = \binom{k}{2} + \varepsilon - p$ is the number of edges in H; from $p \geq k-1+\varepsilon$ we have $m \leq \binom{k-1}{2}$ and so we may apply the inductive hypothesis to H to deduce $S(H) \leq \binom{k-1}{2} \cdot (k-2)$.



Now the labels of edges vw_i have sum

$$\sum_{i=1}^{p} \min\left(\deg_G v, \deg_G w_i\right) = \sum_{i=1}^{p} \deg_G w_i = \sum_{i=1}^{p} (\deg_H w_i + 1) = 2m + p.$$

For each of the edges contained in H, the label on that edge has increased by exactly 1, so those edges contribute $S(H) + m$. In total,

$$\begin{aligned} S(G) &= 2m + p + (S(H) + m) = (m+p) + 2m + S(H) \\ &\leq \binom{k}{2} + \varepsilon + 2\binom{k-1}{2} + \binom{k-1}{2}(k-2) = \binom{k}{2}(k-1) + \varepsilon. \end{aligned}$$

□

Second solution (algebraic, submitted by contestant James Lin) We give a different proof of $S(G) \leq 127009$. The proof proceeds using the following two claims, which will show that $S(G) \leq 127010$ for all graphs G. Then a careful analysis of the equality cases will show that this bound is not achieved for any graph G. Since the example L_{64} earlier has $S(L_{64}) = 127009$, this will solve the problem.

Lemma (Combinatorial bound)**.** *Let G be a graph with 2017 edges and let $d_1 \geq d_2 \geq \cdots \geq d_n$ be the degree sequence of the graph (thus $n \geq 65$). Then*

$$S(G) \leq d_2 + 2d_3 + 3d_4 + \cdots + 63d_{64} + d_{65}.$$

Proof. Let v_1, ..., v_n be the corresponding vertices. For any edge $e = \{v_i, v_j\}$ with $i < j$, we consider associating each edge e with v_j, and computing the sum $S(G)$ indexing over associated vertices. To be precise, if we let a_i denote the number of edges associated to v_i, we now have $a_i \le i - 1$, $\sum a_i = 2017$, and

$$S(G) = \sum_{i=1}^{n} a_i d_i.$$

The inequality $\sum a_i d_i \le d_2 + 2d_3 + 3d_4 + \cdots + 63d_{64} + d_{65}$ then follows for smoothing reasons (by "smoothing" the a_i), since the d_i are monotone. This proves the given inequality. □

Once we have this property, we handle the bounding completely algebraically.

Lemma (Algebraic bound)**.** *Let $x_1 \ge x_2 \ge \cdots \ge x_{65}$ be any nonnegative integers such that $\sum_{i=1}^{65} x_i \le 4034$. Then*

$$x_2 + 2x_3 + \cdots + 63x_{64} + x_{65} \le 127010.$$

Moreover, equality occurs if and only if $x_1 = x_2 = x_3 = \cdots = x_{64} = 63$ and $x_{65} = 2$.

Proof. Let A denote the left-hand side of the inequality. We begin with a smoothing argument.

- Suppose there are indices $1 \le i < j \le 64$ such that $x_i > x_{i+1} \ge x_{j-1} > x_j$. Then replacing (x_i, x_j) by $(x_i - 1, x_j + 1)$ strictly increases A preserving all conditions. Thus we may assume all numbers in $\{x_1, \dots, x_{64}\}$ differ by at most 1.

- Suppose $x_{65} \ge 4$. Then we can replace $(x_1, x_2, x_3, x_4, x_{65})$ by $(x_1+1, x_2+1, x_3+1, x_4+1, x_{65}-4)$ and strictly increase A. Hence we may assume $x_{65} \le 3$.

We will also tacitly assume $\sum x_i = 4034$, since otherwise we can increase x_1. These two properties leave only four sequences to examine:

- $x_1 = x_2 = x_3 = \cdots = x_{63} = 63$, $x_{64} = 62$, and $x_{65} = 3$, which gives $A = 126948$.

- $x_1 = x_2 = x_3 = \cdots = x_{63} = x_{64} = 63$ and $x_{65} = 2$, which gives $A = 127010$.

- $x_1 = 64$, $x_2 = x_3 = \cdots = x_{63} = x_{64} = 63$ and $x_{65} = 1$, which gives $A = 127009$.

- $x_1 = x_2 = 64$, $x_3 = \cdots = x_{63} = x_{64} = 63$ and $x_{65} = 0$, which gives $A = 127009$.

This proves that $A \le 127010$. To see that equality occurs only in the second case above, note that all the smoothing operations other than incrementing x_1 were strict, and that x_1 could not have been incremented in this way as $x_1 = x_2 = 63$. □

This shows that $S(G) \le 127010$ for all graphs G, so it remains to show equality never occurs. Retain the notation d_i and a_i of the combinatorial bound now; we would need to have $d_1 = \cdots = d_{64} = 63$ and $d_{65} = 2$ (in particular, deleting isolated vertices from G, we may assume $n = 65$). In that case, we have $a_i \le i - 1$ but also $a_{65} = 2$ by definition (the last vertex gets all edges associated to it). Finally,

$$\begin{aligned} S(G) &= \sum_{i=1}^{n} a_i d_i = 63(a_1 + \cdots + a_{64}) + a_{65} \\ &= 63(2017 - a_{65}) + a_{65} \le 63 \cdot 2015 + 2 = 126947 \end{aligned}$$

completing the proof.

Remark. Another way to finish once $S(G) \le 127010$ is note there is a unique graph (up to isomorphism and deletion of universal vertices) with degree sequence $(d_1, \dots, d_{65}) = (63, \dots, 63, 2)$. Indeed, the complement of the graph has degree sequence $(1, \dots, 1, 63)$, and so it must be a 63-star plus a single edge. One can then compute $S(G)$ explicitly for this graph.

Some further remarks

Remark. Interestingly, the graph C_4 has $\binom{3}{2} + 1 = 4$ edges and $S(C_4) = 8$, while $S(L_3) = 7$. This boundary case is visible in the combinatorial solution in the base case of the first claim. It also explains why we end up with the bound $S(G) \le 127010$ in the second algebraic solution, and why it is necessary to analyze the equality cases so carefully; observe in $k = 3$ the situation $d_1 = d_2 = d_3 = d_4 = 2$.

Remark. Some comments about further context for this problem:

- The obvious generalization of 2017 to any constant was resolved in September 2018 by Mehtaab Sawhney and Ashwin Sah. The relevant paper is *On the discrepancy between two Zagreb indices*, published in Discrete Mathematics, Volume 341, Issue 9, pages 2575-2589. The arXiv link is `https://arxiv.org/pdf/1801.02532.pdf`.
- The quantity
$$S(G) = \sum_{e=vw} \min(\deg v, \deg w)$$
in the problem has an interpretation: it can be used to provide a bound on the number of triangles in a graph G. To be precise, $\#E(G) \le \frac{1}{3} S(G)$, since an edge $e = vw$ is part of at most $\min(\deg v, \deg w)$ triangles.

- For *planar* graphs it is known $S(G) \leq 18n - 36$ and it is conjectured that for n large enough, $S(G) \leq 18n - 72$. See `https://mathoverflow.net/a/273694/70654`.

---

9 Rigid

§9.1 Synopsis

By "rigid" problems, I mean a class of problems which focus on a specific *concrete* structure. This can't be defined formally, but here are some characteristics:

- Often the problem has very few degrees of freedom.
- The structure is often pretty complex, and understanding it well is the entire point of the problem.
- The particular task you're asked to prove can feel very superficial, almost like an answer extraction (like on the AIME). For example, you might be asked to count the number of objects satisfying P, but in fact you simply characterize all the objects satisfying P and then do the counting as a little step at the end.
- One feels that one is *discovering* mathematics, rather than inventing it — the properties which you prove are forced upon you, rather than your design.
- Often there is only one solution to the problem, up to isomorphism.

You will hopefully begin to see what I mean from the two examples below.

§9.2 Walkthroughs

Problem 98 (TSTST 2016)**.** Prove that if n and k are positive integers satisfying $\varphi^k(n) = 1$, then $n \le 3^k$. (Here φ^k denotes k applications of the Euler phi function.)

Walkthrough. Let a, b, c, ..., denote positive integers.

(a) For positive integers a, b, show that $n = 2^a \cdot 3^b$ takes $a + b$ steps.

(b) How many steps does each of $n = 2^a5^b$, $n = 2^a17^b$, $2^a3^b7^c$, 2^a11^b take?

(c) Show that 2^a2017^b takes $a + 9b$ steps.

(d) Define the function $w \colon \mathbb{N} \to \mathbb{Z}_{\ge 0}$ by $w(ab) = w(a) + w(b)$, $w(2) = 1$, and $w(p) = w(p-1)$ for odd primes p. Figure out the connection between the values of $w(p)$ and the answer to your answer in (b).

(e) By looking at ν_2 prove the conjecture in (d).

(f) Show that $w(n)$ is the number of steps required for n, if n is even. What if n is odd?

(g) Show that $w(n) \geq \log_3 n$ by induction on $n \geq 1$. (The case where n is composite is immediate, so the only work is when n is prime.)

(h) In fact, prove that the stronger estimate $n \leq 2 \cdot 3^{k-1}$ holds (and is best possible).

As a rigid problem, this is a chief example: the point of the problem is to determine the function w, and the "extraction" of comparing to $\log_3$ occurs at the end. It's important to realize that w is "God-given"; we were not permitted any decisions in deriving it.

It might be tempting to try and prove $\varphi(n) \geq n/3$ or similar statements, but this is false, and in any case not representative of small cases. However, I think trying the "small cases": which in this situation are those n with relatively few prime factors — suggests that this is the wrong approach.

Problem 99 (IMO Shortlist 2015)**.** Suppose that $a_0, a_1, \ldots$ and $b_0, b_1, \ldots$ are two sequences of positive integers satisfying $a_0, b_0 \geq 2$ and

$$a_{n+1} = \gcd(a_n, b_n) + 1, \qquad b_{n+1} = \operatorname{lcm}(a_n, b_n) - 1$$

for all $n \geq 0$. Prove that the sequence (a_n) is eventually periodic.

Walkthrough. The rigid philosophy is great here because you can get a lot of concrete data by just picking your two favorite choices of (a_0, b_0). We start by doing that:

(a) Work through the case $(a_0, b_0) = (2, 4)$.

(b) Work through the case $(a_0, b_0) = (2, 10)$.

(c) Work through the case $(a_0, b_0) = (2, 16)$.

(d) Work through the case $(a_0, b_0) = (2, 58)$.

With these tables in front of you, you should be able to start seeing some patterns, or find some counterexamples to hopeful conjectures.

(e) A term a_i in the sequence is said to be a *peak* if $a_{i+1} \leq a_i$. Show that if a_i is not a peak then $a_{i+1} = a_i + 1$.

(f) Make a conjecture about the terms in the sub-sequence of peaks which would imply that the peaks are eventually constant.

(g) Prove your conjecture in (f).

(h) Finish the problem with the claim.

§9.3 Problems

Problem 100 (IMO 2017)**.** For each integer $a_0 > 1$, define the sequence a_0, a_1, a_2, ..., by

$$a_{n+1} = \begin{cases} \sqrt{a_n} & \text{if } \sqrt{a_n} \text{ is an integer,} \\ a_n + 3 & \text{otherwise} \end{cases}$$

for each $n \geq 0$. Determine all values of a_0 for which there is a number A such that $a_n = A$ for infinitely many values of n.

Problem 101 (USAJMO 2013)**.** Each cell of an $m \times n$ board is filled with some nonnegative integer. Two numbers in the filling are said to be *adjacent* if their cells share a common side. The filling is called a *garden* if it satisfies the following two conditions:

(i) The difference between any two adjacent numbers is either 0 or 1.

(ii) If a number is less than or equal to all of its adjacent numbers, then it is equal to 0.

Determine the number of distinct gardens in terms of m and n.

Problem 102 (IMO Shortlist 1995)**.** For an integer $x \geq 1$, let $p(x)$ be the least prime that does not divide x, and define $q(x)$ to be the product of all primes less than $p(x)$. In particular, $p(1) = 2$. For x having $p(x) = 2$, define $q(x) = 1$. Consider the sequence $x_0, x_1, x_2, \ldots$ defined by $x_0 = 1$ and

$$x_{n+1} = \frac{x_n p(x_n)}{q(x_n)}$$

for $n \geq 0$. Find all n such that $x_n = 1995$.

Problem 103 (EGMO 2014)**.** Let n be a positive integer. We have n boxes where each box contains a non-negative number of pebbles. In each move we are allowed to take two pebbles from a box we choose, throw away one of the pebbles and put the other pebble in another box we choose. An initial configuration of pebbles is called *solvable* if it is possible to reach a configuration with no empty box, in a finite (possibly zero) number of moves. Determine all initial configurations of pebbles which are not solvable, but become solvable when an additional pebble is added to a box, no matter which box is chosen.

Problem 104 (TSTST 2014)**.** Let $\leftarrow$ denote the left arrow key on a standard keyboard. If one opens a text editor and types the keys "ab$\leftarrow$ cd $\leftarrow\leftarrow$ e $\leftarrow\leftarrow$ f", the result is "faecdb". We say that a string B is *reachable* from a string A if it is possible to insert some amount of $\leftarrow$'s in A, such that typing the resulting characters produces B. So, our example shows that "faecdb" is reachable from "abcdef".

Prove that for any two strings A and B, A is reachable from B if and only if B is reachable from A.

Problem 105 (IMO 2005)**.** Let a_1, a_2, ...be a sequence of integers with infinitely many positive and negative terms. Suppose that for every positive integer n the numbers a_1, a_2, ..., a_n leave n different remainders upon division by n. Prove that every integer occurs exactly once in the sequence.

Problem 106 (USAMO 2010)**.** There are n students standing in a circle, one behind the other. The students have heights $h_1 < h_2 < \cdots < h_n$. If a student with height h_k is standing directly behind a student with height h_{k-2} or less, the two students are permitted to switch places. Prove that it is not possible to make more than $\binom{n}{3}$ such switches before reaching a position in which no further switches are possible.

Problem 107 (IMO Shortlist 2017)**.** Let $f\colon \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \to \{0,1\}$ be a function such that $f(2,1) = f(1,1) = 0$. Assume that for any relatively prime positive integers (a,b) not both equal to 1, we have

$$f(a,b) = 1 - f(b,a) = f(a+b,b).$$

Let p be an odd prime. Prove that

$$\sum_{n=1}^{p-1} f(n^2,p) \geq \sqrt{2p} - 2.$$

§9.4 Solutions

Solution 98. The main observation is that the exponent of 2 decreases by at most 1 with each application of φ. This will give us the desired estimate.

Define the *weight* function w on positive integers as follows: it satisfies $w(ab) = w(a) + w(b)$, $w(2) = 1$ and $w(p) = w(p-1)$ for any prime $p > 2$. By induction, we see that $w(n)$ counts the powers of 2 that are produced as φ is repeatedly applied to n. In particular, $k \geq w(n)$.

From $w(2) = 1$, it suffices to prove that $w(p) \geq \log_3 p$ for every $p > 2$. We use strong induction and note that

$$w(p) = w(2) + w\left(\frac{p-1}{2}\right) \geq 1 + \log_3(p-1) - \log_3 2 \geq \log_3 p$$

for any $p > 2$. This solves the problem.

Remark. You can even motivate this solution through small cases $2^x 3^y$ like $2^x 17^w$, $2^x 3^y 7^z$, $2^x 11^t$.

Moreover, the stronger bound $n \leq 2 \cdot 3^{k-1}$ is true and best possible.

Solution 99. Here are, first, several examples:

2	4
3	3
4	2
3	3
4	2
3	3
⋮	⋮

2	10
3	9
4	8
5	7
2	34
3	33
4	32
5	31
2	154
3	153
⋮	⋮

2	16
3	15
4	14
3	27
4	26
3	51
4	50
3	99
⋮	⋮

2	58
3	57
4	56
5	55
6	54
7	53
2	370
3	369
4	368
5	367
2	1834
3	1833
4	1832
5	1831
2	9064
⋮	⋮

Define an index i to be a *peak* if $a_{i+1} \leq a_i$. Note that if i is not a peak then $(a_{i+1}, b_{i+1}) = (a_i + 1, b_i - 1)$.

Claim. The sub-sequences of peaks is non-increasing.

Proof. Equivalently, if $i < k$ are adjacent peaks, we show $a_i \geq a_k$. Assume not. Set $(a_i, b_i) = (dx, dy)$ where $\gcd(x, y) = 1$. Then $(a_{i+1}, b_{i+1}) = (d+1, dxy-1)$. By assumption, we then arrive at the pair $(a_j, b_j) = (dx, dxy + d - dx)$ later, where $j < k$. At this point the GCD of these two numbers is $d < dx$, so j must be a peak, contradiction. □

Thus, the peaks are monotonic, so (a_n) bounded. Let $M = (\max a_i)!$. Now, note that $(a_i, b_i \mod M)$ determines $(a_{i+1}, b_{i+1} \mod M)$ since $\gcd(a_i, b_i)$ is determined by $b_i \mod M$, and also

$$b_{i+1} = \underbrace{\frac{a_i}{\gcd(a_i, b_i)}}_{\in \mathbb{Z}} b_i + 1.$$

Since the number of such pairs is finite, a_i is eventually periodic. Alternatively, using a direct calculation one can show that after the peaks stabilize, the sequence becomes periodic.

Solution 100. The answer is $a_0 \equiv 0 \pmod 3$ only.

First solution We first compute the minimal term of any sequence, periodic or not.

Lemma. *Let c be the smallest term in a_n. Then either $c \equiv 2 \pmod 3$ or $c = 3$.*

Proof. Clearly $c \neq 1, 4$. Assume $c \not\equiv 2 \pmod 3$. As c is not itself a square, the next perfect square after c in the sequence is one of $(\lfloor \sqrt{c} \rfloor + 1)^2$, $(\lfloor \sqrt{c} \rfloor + 2)^2$, or $(\lfloor \sqrt{c} \rfloor + 3)^2$. So by minimality we require

$$c \leq \lfloor \sqrt{c} \rfloor + 3 \leq \sqrt{c} + 3$$

which requires $c \leq 5$. Since $c \neq 1, 2, 4, 5$ we conclude $c = 3$. □

Now we split the problem into two cases:

- If $a_0 \equiv 0 \pmod 3$, then all terms of the sequence are $0 \pmod 3$. The smallest term of the sequence is thus 3 by the lemma and we have

  $$3 \to 6 \to 9 \to 3$$

  so $A = 3$ works fine.
- If $a_0 \not\equiv 0 \pmod 3$, then no term of the sequence is $0 \pmod 3$, and so in particular 3 does not appear in the sequence. So the smallest term of the sequence is $2 \pmod 3$ by lemma. But since no squares are $2 \pmod 3$, the sequence a_k grows without bound forever after, so no such A can exist.

Hence the answer is $a_0 \equiv 0 \pmod 3$ only.

Second solution We clean up the argument by proving the following lemma.

Lemma. *If a_n is constant modulo 3 and not 2 (mod 3), then a_n must eventually cycle in the form $(m, m+3, m+6, \dots, m^2)$, with no squares inside the cycle except m^2.*

Proof. Observe that a_n must eventually hit a square, say $a_k = c^2$; the next term is $a_{k+1} = c$. Then it is forever impossible to exceed c^2 again, by what is essentially discrete intermediate value theorem. Indeed, suppose $a_\ell > c^2$ and take $\ell > k$ minimal (in particular $a_\ell \neq \sqrt{a_{\ell-1}}$). Thus $a_{\ell-1} \in \{c^2-2, c^2-1, c^2\}$ and thus for modulo 3 reasons we have $a_{\ell-1} = c^2$. But that should imply $a_\ell = c < c^2$, contradiction.

We therefore conclude $\sup\{a_n, a_{n+1}, \dots\}$ is a decreasing integer sequence in n. It must eventually stabilize, say at m^2. Now we can't hit a square between m and m^2, and so we are done. □

Now, we contend that all $a_0 \equiv 0 \pmod 3$ work. Indeed, for such a_0 we have $a_n \equiv 0 \pmod 3$ for all n, so the lemma implies that the problem statement is valid.

Next, we observe that if $a_i \equiv 2 \pmod 3$, then the sequence grows without bound afterwards since no squares are 2 (mod 3). In particular, if $a_0 \equiv 2 \pmod 3$ the answer is no.

Finally, we claim that if $a_0 \equiv 1 \pmod 3$, then eventually some term is 2 (mod 3). Assume for contradiction this is not so; then $a_n \equiv 1 \pmod 3$ must hold forever, and the lemma applies to give us a cycle of the form $(m, m+3, \dots, m^2)$ where $m \equiv 1 \pmod 3$. In particular $m \geq 4$ and

$$m \leq (m-2)^2 < m^2$$

but $(m-2)^2 \equiv 1 \pmod 3$ which is a contradiction.

Solution 101. The numerical answer is $2^{mn} - 1$. But we claim much more, by giving an explicit description of all gardens:

> Let S be any nonempty subset of the mn cells. Suppose we fill each cell θ with the minimum (taxicab) distance from θ to some cell in S (in particular, we write 0 if $\theta \in S$). Then
>
> - This gives a garden, and
> - All gardens are of this form.

Since there are $2^{mn}-1$ such nonempty subsets S, this would finish the problem. An example of a garden with $|S| = 3$ is shown below.

$$\begin{bmatrix} 2 & 1 & 2 & 1 & \mathbf{0} & 1 \\ 1 & \mathbf{0} & 1 & 2 & 1 & 2 \\ 1 & 1 & 2 & 3 & 2 & 3 \\ \mathbf{0} & 1 & 2 & 3 & 3 & 4 \end{bmatrix}$$

It is actually fairly easy to see that this procedure always gives a garden; so we focus our attention on showing that every garden is of this form.

Given a garden, note first that it has at least one cell with a zero in it — by considering the minimum number across the entire garden. Now let S be the (thus nonempty) set of cells with a zero written in them. We contend that this works, i.e. the following sentence holds:

Claim. If a cell θ is labeled d, then the minimum distance from that cell to a cell in S is d.

Proof. The proof is by induction on d, with $d = 0$ being by definition. Now, consider any cell θ labeled $d \geq 1$. Every neighbor of θ has label at least $d-1$, so any path will necessarily take $d-1$ steps after leaving θ. Conversely, there is *some* $d-1$ adjacent to θ by (ii). Stepping on this cell and using the minimal path (by induction hypothesis) gives us a path to a cell in S with length *exactly* d. So the shortest path does indeed have distance d, as desired. □

Solution 102. We write out a few terms:

$$\begin{aligned}
x_1 &= 2^1 \\
x_2 &= 2^0 \times 3^1 \\
x_3 &= 2^1 \times 3^1 \\
x_4 &= 2^0 \times 3^0 \times 5^1 \\
x_5 &= 2^1 \times 3^0 \times 5^1 \\
x_6 &= 2^0 \times 3^1 \times 5^1 \\
x_7 &= 2^1 \times 3^1 \times 5^1 \\
x_8 &= 2^0 \times 3^0 \times 5^0 \times 7^1 \\
x_9 &= 2^1 \times 3^0 \times 5^0 \times 7^1 \\
x_{10} &= 2^0 \times 3^1 \times 5^0 \times 7^1 \\
&\vdots
\end{aligned}$$

We observe that the exponents are 0 and 1, and moreover encode n in binary. More precisely, if $n = \overline{\ldots b_2 b_1}_{(2)}$ in binary, then we have the explicit form

$$x_n = 2^{b_1} \times 3^{b_2} \times 5^{b_3} \times \ldots.$$

This is more or less tautological by induction.

In particular, $1995 = 3 \times 5 \times 7 \times 19$, so $x_n = 1995$ exactly when $n = \overline{10001110}_{(2)} = 142$.

Solution 103. The point of the problem is to characterize all the solvable configurations. We claim that it is given by the following:

Claim. A configuration $(a_1, \dots, a_n)$ is solvable if and only if

$$\sum_1^n \left\lceil \frac{a_i}{2} \right\rceil \geq n.$$

Proof. The proof is by induction on the number of stones. If there are fewer than n stones there is nothing to prove. Now assume there are at least n stones, and let $S = \sum \lceil a_i/2 \rceil$. Then:

- If $S < n$, this remains true after any operation, so by induction the configuration is not solvable.
- Suppose $S \geq n$, and also that there is an empty box (else we are already done). Then there must be some box with at least two stones. In that case, using those two stones to fill the empty box does not change the value of S, but decreases the total number of stones by one, as desired.

□

From here we may then extract the answer to the original problem: we require all a_i to be even and $\sum a_i = 2n - 2$.

Remark. It should be unsurprising that a criteria of this form exists, since (1) intuitively, one loses nothing by filling empty boxes as soon as possible, and then ignoring boxes with one pebble in them, (2) the set of configurations is a graded partially ordered set, so one can inductively look at small cases.

Solution 104. Obviously A and B should have the same multiset of characters, and we focus only on that situation.

Claim. If $A = 123\dots n$ and $B = \sigma(1)\sigma(2)\dots\sigma(n)$ is a permutation of A, then B is reachable if and only if it is **213-avoiding**, i.e. there are no indices $i < j < k$ such that $\sigma(j) < \sigma(i) < \sigma(k)$.

Proof. This is clearly necessary. To see its sufficient, one can just type B inductively: after typing k, the only way to get stuck is if $k+1$ is to the right of k and there is some character in the way; this gives a 213 pattern. □

Claim. A permutation σ on $\{1, \dots, n\}$ is 213-avoiding if and only if the inverse σ^{-1} is.

Proof. Suppose $i < j < k$ and $\sigma(j) < \sigma(i) < \sigma(k)$. Let $i' = \sigma(j)$, $j' = \sigma(i)$, $k' = \sigma(k)$; then $i' < j' < k'$ and $\sigma^{-1}(j') < \sigma^{-1}(i') < \sigma^{-1}(k')$. □

This essentially finishes the problem. Suppose B is reachable from A. By using the typing pattern, we get some permutation $\sigma\colon \{1,\dots,n\}$ such that the ith character of A is the $\sigma(i)$th character of B, and which is 213-avoiding by the claim. (The permutation is unique if A has all distinct characters, but there could be multiple if A has repeated ones.) Then σ^{-1} is 213-avoiding too and gives us a way to change B into A.

Solution 105. Obviously every integer appears at most once (otherwise take n much larger). So we will prove every integer appears at least once.

Claim. For any $i < j$ we have $|a_i - a_j| < j$.

Proof. Otherwise, let $n = |a_i - a_j| \neq 0$. Then $i, j \in [1, n]$ and $a_i \equiv a_j \pmod{n}$, contradiction. □

Claim. For any n, the set $\{a_1, \dots, a_n\}$ is of the form $\{k+1, \dots, k+n\}$ for some integer k.

Proof. By induction, with the base case $n = 1$ being vacuous. For the inductive step, suppose $\{a_1, \dots, a_n\} = \{k+1, \dots, k+n\}$ are determined. Then

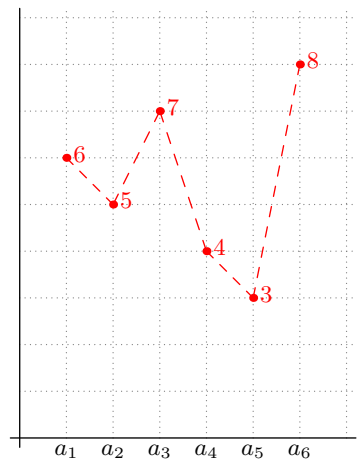
$$a_{n+1} \equiv k \pmod{n+1}.$$

Moreover by the earlier claim we have

$$|a_{n+1} - a_1| < n+1.$$

From this we deduce $a_{n+1} \in \{k, k+n+1\}$ as desired. □

This gives us actually a complete description of all possible sequences satisfying the hypothesis: choose any value of a_1 to start. Then, for the nth term, the set $S = \{a_1, \dots, a_{n-1}\}$ is (in some order) a set of $n-1$ consecutive integers. We then let $a_n = \max S + 1$ or $a_n = \min S - 1$. A picture of six possible starting terms is shown below.



Finally, we observe that the condition that the sequence has infinitely many positive and negative terms (which we have not used until now) implies it is unbounded above and below. Thus it must contain every integer.

Solution 106. The main claim is the following observation, which is most motivated in the situation $j - i = 2$.

Claim. The students with heights h_i and h_j switch at most $|j-i|-1$ times.

Proof. By induction on $d = |j-i|$, assuming $j > i$. For $d = 1$ there is nothing to prove.

For $d \geq 2$, look at only students h_j, h_{i+1} and h_i ignoring all other students. After h_j and h_i switch the first time, the relative ordering of the students must be $h_i \to h_j \to h_{i+1}$. Thereafter h_j must always switch with h_{i+1} before switching with h_i, so the inductive hypothesis applies to give the bound $1 + j - (i+1) - 1 = j - i - 1$. □

Hence, the number of switches is at most

$$\sum_{1 \leq i < j \leq n} (|j-i|-1) = \binom{n}{3}.$$

Solution 107. Incredibly, we have the following description of f.

Lemma. *For any relatively prime $(a,b) \neq (1,1)$,*

$$f(a,b) = \begin{cases} 1 & (a^{-1} \bmod b) \leq b/2 \\ 0 & (a^{-1} \bmod b) > b/2. \end{cases}$$

We give the short self-contained induction proof for now; see the remarks for a more reasonable and motivated proof.

Inductive proof by Ankan Bhattacharya. It is enough to show that if $a, b > 1$ are relatively prime then $a^{-1} \pmod b \leq b/2$ iff $b^{-1} \pmod a > a/2$. Let (x,y) be a minimal positive integer pair with $ax - by = 1$. Then $x \leq b-1$, $y \leq a-1$, and

$$\begin{aligned} a^{-1} &\equiv x \pmod b \\ b^{-1} &\equiv a - y \pmod a. \end{aligned}$$

Thus $a^{-1} \pmod b = x$, $b^{-1} \pmod a = a - y$. Finally

$$x \leq b/2 \iff ax \leq ab/2 \iff by < ab/2 \iff y < a/2 \iff a - y > a/2.$$

□

In particular, for any n such that $n \equiv \pm 1/k \pmod p$ with $k \in \{1, \dots, \left\lfloor \sqrt{p/2} \right\rfloor\}$, we have $f(n^2, p) = 1$, so this implies the result.

Remark. In general, we have

$$f(a,p) = 1 - f(p,a) = 1 - f(p-a,a) = f(a,p-a) = f(p,p-a) = 1 - f(p-a,p)$$

and so $f(a,p) + f(p-a,p) = 1$.

Note that if $p \equiv 1 \pmod 4$, this already solves the problem. If r is any quadratic residue, so is $-r$, and accordingly $f(-r,p) + f(r,p) = 1$; so we have actually

$$\sum_n f(n^2,p) = \frac{1}{2}(p-1) \qquad \forall p \equiv 1 \pmod 4.$$

Remark. In fact, for $p \equiv 3 \pmod 4$ it turns out the number of quadratic residues in $[1, p/2]$ is more than the number in $[p/2, p-1]$, and hence the $\frac{1}{2}(p-1)$ is actually sharp.

Indeed, if one defines the Dirichlet L-function

$$L(s) = \sum_n \left(\frac{n}{p}\right) n^{-s}$$

then it is known that

$$L(1) = \frac{\pi}{\left(2 - \left(\frac{2}{p}\right)\right)\sqrt{p}} \sum_{n=1}^{\frac{p-1}{2}} \left(\frac{n}{p}\right) > 0$$

which is the result we wanted. It seems no elementary proof is known, though.

Remark (Yang Liu)**.** The key lemma in the problem seems to come out of nowhere. Here is one way you can come up with it.

Denote by $\mathrm{GL}_2(\mathbb{Z})$ the set of 2×2 integer matrices with determinant ± 1. Suppose we consider only coprime pairs (a,b) with $a \geq b \geq 0$.

Consider first running the Euclidean algorithm backwards; starting from $(1,0)$ and trying to reach a given pair. An any point we can go from $(a,b) \to (a+b,b)$ or $(a,b) \to (a+b,a)$; the latter operation involves a *switch* and we're trying to count the parity of switches. (We don't count $(1,1) \to (2,1)$ as a switch.) If we interpret our pair as a column vector $\begin{bmatrix} a \\ b \end{bmatrix}$, then this means we are multiplying by either multiplying by $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ or $S = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ (for "switch"), one after another, several times. (For experts, I think T and S generate $\mathrm{GL}_2(\mathbb{Z})$.) As an example, to reach $(18,7)$ from $(1,0)$ we do

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \xrightarrow{\times S} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \xrightarrow{\times T} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \xrightarrow{\times T} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \xrightarrow{\times S} \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$
$$\xrightarrow{\times S} \begin{bmatrix} 7 \\ 4 \end{bmatrix} \xrightarrow{\times S} \begin{bmatrix} 11 \\ 7 \end{bmatrix} \xrightarrow{\times T} \begin{bmatrix} 18 \\ 7 \end{bmatrix}.$$

The punch line is that the overall matrix M we have is one whose first column is $\begin{bmatrix} a \\ b \end{bmatrix}$, and we want to count the number of times we used the matrix S. But $\det T = +1$ and $\det S = -1$, so this is given by the sign of $\det M \in \{\pm 1\}$, as we wanted!

Going forwards again, the idea is that given $\begin{bmatrix} a \\ b \end{bmatrix}$ that we are processing with the Euclidean algorithm, we can annotate by completing it to a 2×2 matrix in $\mathrm{GL}_2(\mathbb{Z})$ with nonnegative entries, such that the first row exceeds the second row. As an example, for $(a, b) = (18, 7)$ the process goes $(18, 7) \to (7, 4) \to (4, 3) \to (3, 1) \to (1, 0)$, and the set of accompanying annotated matrices is

$$\begin{bmatrix} 18 & 5 \\ 7 & 2 \end{bmatrix} \to \begin{bmatrix} 7 & 2 \\ 4 & 1 \end{bmatrix} \to \begin{bmatrix} 4 & 1 \\ 3 & 1 \end{bmatrix} \to \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Each steps corresponds to doing row reductions and then swapping rows; the determinant flips sign at every switch. The left column contains the actual (a, b) that are being processed while the right column contains the suitable inverses.

Thus the sign of the determinant of the initial matrix, when populated with nonnegative entries, determines the eventual parity. Essentially, there is a unique nonnegative pair of integers (x, y) for which $ay - bx = \pm 1$, $x \geq y$ and $x \leq a/2$, $y \leq b/2$. (You can prove this, but it's annoying.)

Note here $y \equiv \pm a^{-1} \pmod{b}$, with the choice of sign determining the sign of the determinant and hence the fate. So this implies the key lemma falls out.

---

10 Free

§10.1 Synopsis

This lecture discusses the other extreme to the rigid one: the problem gives you lots of freedom. It's impossible in these cases to try and understand the structure completely, but the problem will ask you only to do a small part of it (e.g. an existence proof). Often you have to impose the structure yourself so that you have something to work with.

Many of these problems end up being constructions. I think you can often think about these problems in two directions.

- **Experimenting**: in which you go forwards and write down an example, and hope that it works (or see why it doesn't work and adjust).
- **Restricting**: in which you add constraints on the thing you're constructing.

  The best case is if you can prove that the constraints you've added must be true for any example. For example if you're trying to construct $n \in \mathbb{Z}$ with a certain property, and you can prove that n even all fail, then of course you only look at odd n henceforth.

  However, if you're willing to take a gamble (which you should be willing to do sometimes), you can also narrow your search even more, even if it's not true (or you can't prove) that the condition is necessary. For example you might focus your attention to the case n is prime, because it helps simplify your experimentation, and you think that a prime n should exist anyways. This can turn an otherwise free problem into a more rigid one.

Most problems will involve a bit of both (and if you're working on a yes/no problem, it's super important to do both).

§10.2 Walkthroughs

Problem 108 (USAMO 2011)**.** Consider the assertion that for each positive integer $n \geq 2$, the remainder upon dividing 2^{2^n} by $2^n - 1$ is a power of 4. Either prove the assertion or find (with proof) a counterexample.

Walkthrough. This is a quick problem showing that you can (and should) often do constructions using both directions: parts (b) and (c) are restrictive, part (d) is experimental.

(a) Show that the problem is equivalent to whether there exists n such that the remainder $2^n \bmod n$ is odd.

(b) Prove that any working n must be odd.

(c) Prove that any working n is composite.

(d) Guess values of n until you find one that works.

In (b) and (c) we were even able to prove n must be odd composite in order to have a chance of working. In other problems you might not be so lucky that you can prove your restrictions are necessary, but it's often correct to take the restriction anyways.

Problem 109 (USAMO 2010)**.** The 2010 positive real numbers a_1, a_2, ..., a_{2010} satisfy the inequality $a_i a_j \le i + j$ for all $1 \le i < j \le 2010$. Determine, with proof, the largest possible value of the product $a_1 a_2 \dots a_{2010}$.

Walkthrough. While it's possible to write down a formula that achieves the maximum, I want to push the point of view that this is a problem which should be done almost entirely by restriction.

First, let's get a sense of what we expect the optimum to be.

(a) We can get an upper bound by multiplying 1005 disjoint inequalities together. For example, you can show 2011^{1005} is an upper bound by using $a_1 a_{2010} \le 2011$, $a_2 a_{2009} \le 2011$, and so on.

However, this is far from optimal: find the best possible bound you can get by multiplying 1005 disjoint inequalities. (You can even prove your answer is the best possible.)

(b) Convince yourself you are on the right track by showing the analogous upper bound where 2010 is replaced by 4 actually has an existing (a_1, a_2, a_3, a_4) with $a_1 a_2 a_3 a_4$ achieving your conjectured maximum.

(c) For any sequence achieving the bound found in (a), write down 1005 equalities which must be true.

(d) Find constants c, λ such that one expects $a_n \approx (cn)^\lambda$ in any valid construction. (Possible hint: AM-GM.)

Now we can start thinking about how to construct the optimum.

(e) Based on your answer to (d), which of the $\binom{2010}{2}$ inequalities are the sharpest?

(f) Try to come up with a construction by choosing an additional 1004 inequalities to set to equalities.

(g) If you made the right choice in (f), you can already prove that $a_i a_j \le i+j$ when i and j are different parity, using only the inequalities you chose in (c) and (f). Do so. The proof will be along the same lines as in (a).

(h) Show that with the additional condition $a_{2008}a_{2010} = 4018$, we are done. Why might we pick this one and not $a_1a_3 = 4$?

Thus we solve the problem almost entirely by adding restrictions. As mentioned, there does exist a solution which explicitly defines a_n but this one makes no attempt to do so whatsoever. Rather, one comes up with a heuristic reason to believe that certain inequalities are more important than others, and then sets those "sharpest" inequalities to equalities, and watches the problem solve itself.

After part (g), the following program might be fun to play with, to get a sense of why the answer to (h) is what it is. (Note: contains spoilers for the earlier parts).

```
def check(x):
        a = [x]
        for t in range(3,4021,2):
                a.append(t/a[-1])
        for i in range(1,2011):
                for j in range(i+1,2011):
                        if a[i-1] * a[j-1] > i+j+1e-10:
                                return (i,j)
        return None
```

§10.3 Problems

Problem 110 (USAMO 1985)**.** Determine whether or not there are any positive integral solutions of the simultaneous equations

$$x_1^2 + x_2^2 + \cdots + x_{1985}^2 = y^3$$
$$x_1^3 + x_2^3 + \cdots + x_{1985}^3 = z^2$$

with distinct integers x_1, x_2, ..., x_{1985}.

Problem 111 (RMM 2015)**.** Does there exist an infinite sequence of integers a_1, a_2, ...such that $\gcd(a_m, a_n) = 1$ if and only if $|m - n| = 1$?

Problem 112 (IMO 2015)**.** We say that a finite set $\mathcal{S}$ of points in the plane is *balanced* if, for any two different points A and B in $\mathcal{S}$, there is a point C in $\mathcal{S}$ such that $AC = BC$. We say that $\mathcal{S}$ is *centre-free* if for any three different points A, B and C in $\mathcal{S}$, there are no points P in $\mathcal{S}$ such that $PA = PB = PC$.

(a) Show that for all integers $n \geq 3$, there exists a balanced set consisting of n points.

(b) Determine all integers $n \geq 3$ for which there exists a balanced centre-free set consisting of n points.

Problem 113 (IMO 2014)**.** Let $n \geq 2$ be an integer. Consider an $n \times n$ chessboard consisting of n^2 unit squares. A configuration of n rooks on this board is *peaceful* if every row and every column contains exactly one rook. Find the greatest positive integer k such that, for each peaceful configuration of n rooks, there is a $k \times k$ square which does not contain a rook on any of its k^2 unit squares.

Problem 114 (IMO 2016)**.** The equation

$$(x-1)(x-2)\cdots(x-2016) = (x-1)(x-2)\cdots(x-2016)$$

is written on the board, with 2016 linear factors on each side. What is the least possible value of k for which it is possible to erase exactly k of these 4032 linear factors so that at least one factor remains on each side and the resulting equation has no real solutions?

Problem 115 (IMO Shortlist 2011)**.** Prove that for every positive integer n, the set $\{2, 3, 4, \ldots, 3n+1\}$ can be partitioned into n triples in such a way that the numbers from each triple are the lengths of the sides of some obtuse triangle.

Problem 116 (USAMO 2014)**.** Prove that there exists an infinite set of points

$$\ldots,\ P_{-3},\ P_{-2},\ P_{-1},\ P_0,\ P_1,\ P_2,\ P_3,\ \ldots$$

in the plane with the following property: For any three distinct integers a, b, and c, points P_a, P_b, and P_c are collinear if and only if $a + b + c = 2014$.

Problem 117 (TSTST 2015)**.** A *Nim-style game* is defined as follows. Two positive integers k and n are specified, along with a finite set S of k-tuples of integers (not necessarily positive). At the start of the game, the k-tuple $(n, 0, 0, \ldots, 0)$ is written on the blackboard.

A legal move consists of erasing the tuple $(a_1, a_2, \ldots, a_k)$ which is written on the blackboard and replacing it with $(a_1 + b_1, a_2 + b_2, \ldots, a_k + b_k)$, where $(b_1, b_2, \ldots, b_k)$ is an element of the set S. Two players take turns making legal moves, and the first to write a negative integer loses. In the event that neither player is ever forced to write a negative integer, the game is a draw.

Prove that there is a choice of k and S with the following property: the first player has a winning strategy if n is a power of 2, and otherwise the second player has a winning strategy.

§10.4 Solutions

Solution 108. We claim $n = 25$ is a counterexample. Since $2^{25} \equiv 2^0 \pmod{2^{25}-1}$, we have

$$2^{2^{25}} \equiv 2^{2^{25} \bmod 25} \equiv 2^7 \bmod 2^{25}-1$$

and the right-hand side is actually the remainder, since $0 < 2^7 < 2^{25}$. But 2^7 is not a power of 4.

Remark. Really, the problem is just equivalent for asking 2^n to have odd remainder when divided by n.

Solution 109. The answer is $3 \times 7 \times 11 \times \cdots \times 4019$, which is clearly an upper bound (and it's not too hard to show this is the lowest number we may obtain by multiplying 1005 equalities together; this is essentially the rearrangement inequality). The tricky part is the construction. Intuitively we want $a_i \approx \sqrt{2i}$, but the details require significant care.

Note that if this is achievable, we will require $a_n a_{n+1} = 2n+1$ for all odd n. Here are two constructions:

- One can take the sequence such that $a_{2008}a_{2010} = 4028$ and $a_n a_{n+1} = 2n+1$ for all $n = 1, 2, \ldots, 2009$. This can be shown to work by some calculation. As an illustrative example,
$$a_1 a_4 = \frac{a_1 a_2 \cdot a_3 a_4}{a_2 a_3} = \frac{3 \cdot 7}{5} < 5.$$
- In fact one can also take $a_n = \sqrt{2n}$ for all even n (and hence $a_{n-1} = \sqrt{2n} - \frac{1}{\sqrt{2n}}$ for such even n).

Remark. This is a chief example of an "abstract" restriction-based approach. One can motivate it in three steps:

- The bound $3 \cdot 7 \cdot \cdots \cdot 4019$ is provably best possible upper bound by pairing the inequalities; also the situation with 2010 replaced by 4 is constructible with bound 21.
- We have $a_n \approx \sqrt{2n}$ heuristically; in fact $a_n = \sqrt{2n}$ satisfies inequalities by AM-GM.
- So we are most worried about $a_i a_j \le i + j$ when $|i-j|$ is small, like $|i-j| = 1$.

I then proceeded to spend five hours on various constructions, but it turns out that the right thing to do was just require $a_k a_{k+1} = 2k+1$, to make sure these pass: and the problem almost solves itself.

Remark. When 2010 is replaced by 4 it is not too hard to manually write an explicit example: say $a_1 = \frac{\sqrt{3}}{1.1}$, $a_2 = 1.1\sqrt{3}$, $a_3 = \frac{\sqrt{7}}{1.1}$ and $a_4 = 1.1\sqrt{7}$. So this is a reason one might guess that $3 \times 7 \times \cdots \times 4019$ can actually be achieved in the large case.

Remark. Victor Wang says: I believe we can actually prove that WLOG (!) assume $a_i a_{i+1} = 2i + 1$ for all i (but there are other ways to motivate that as well, like linear programming after taking logs), which makes things a bit simpler to think about.

Solution 110. Yes, take

$$x_k = k \cdot (1^2 + 2^2 + \cdots + 1985^2)^4$$

which works (noting $1^3 + \cdots + 1985^3 = (1 + \cdots + 1985)^2$ is a square).

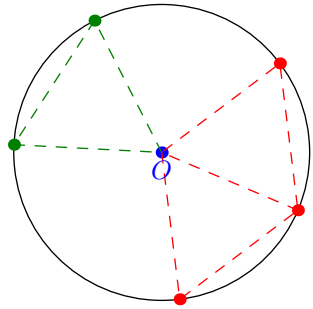
Solution 111. The answer is yes. Let $p_1 = 2$, $q_1 = 3$, $p_2 = 5$, $q_2 = 7$, $p_3 = 11$, …denote the sequence of prime numbers. Define

$$a_n = p_n q_n \cdot \begin{cases} \prod_{k=1}^{n-2} p_k & n \text{ even} \\ \prod_{k=1}^{n-2} q_k & n \text{ odd.} \end{cases}$$

This works by construction.

Remark. Here is an idea of how to come up with this. The idea is that you just take every pair $i < j$ you want not to be relatively prime (meaning $|i - j| \geq 2$) and throw in a prime. You can't do this by using a different prime for every pair (since each a_i must be finite) and you can't use the same prime for a fixed i, so you do the next best thing and alternate using even and odd and you're done.

Solution 112. For part (a), take a circle centered at a point O, and add $n - 1$ additional points by adding pairs of points separated by an arc of $60°$ or similar triples. An example for $n = 6$ is shown below.



For part (b), the answer is odd n, achieved by taking a regular n-gon. To show even n fail, note that some point is on the perpendicular bisector of

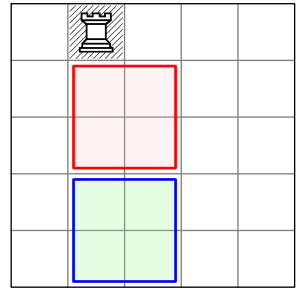
$$\left\lceil \frac{1}{n}\binom{n}{2} \right\rceil = \frac{n}{2}$$

pairs of points, which is enough. (This is a standard double-counting argument.)

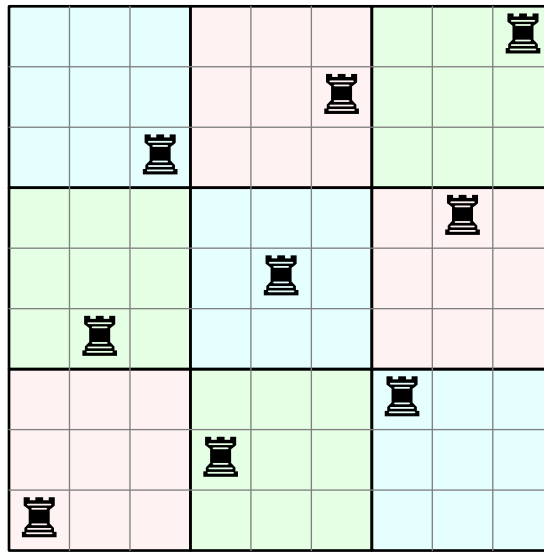
As an aside, there is a funny joke about this problem. There are two types of people in the world: those who solve (b) quickly and then take forever to solve (a), and those who solve (a) quickly and then can't solve (b) at all. (Empirically true when the Taiwan IMO 2014 team was working on it.)

Solution 113. The answer is $k = \lfloor \sqrt{n-1} \rfloor$, sir.

First, assume $n > k^2$ for some k. We will prove we can find an empty $k \times k$ square. Indeed, let R be a rook in the uppermost column, and draw k squares of size $k \times k$ directly below it, aligned. There are at most $k-1$ rooks among these squares, as desired.



Now for the construction for $n = k^2$. We draw the example for $k = 3$ (with the generalization being obvious);



To show that this works, consider for each rook drawing an $k \times k$ square of X's whose bottom-right hand corner is the rook (these may go off the board). These indicate positions where one cannot place the upper-left hand corner of any square. It's easy to see that these cover the entire board, except parts of the last $k-1$ columns, which don't matter anyways.

It remains to check that $n \le k^2$ also all work (omitting this step is a common mistake). For this, we can delete rows and column to get an $n \times n$ board, and then fill in any gaps where we accidentally deleted a rook.

Solution 114. The answer is 2016. Obviously this is necessary in order to delete duplicated factors. We now prove it suffices to deleted 2 (mod 4) and 3 (mod 4) guys from the left-hand side, and 0 (mod 4), 1 (mod 4) from the right-hand side.

Consider the 1008 inequalities

$$\begin{aligned}(x-1)(x-4) &< (x-2)(x-3)\\ (x-5)(x-8) &< (x-6)(x-7)\\ (x-9)(x-12) &< (x-10)(x-11)\\ &\vdots\\ (x-2013)(x-2016) &< (x-2014)(x-2015).\end{aligned}$$

Notice that in all these inequalities, at most one of them has non-positive numbers in it, and we never have both zero. If there is exactly one negative term among the $1008 \cdot 2 = 2016$ sides, it is on the left and we can multiply all together. Thus the only case that remains is if $x \in (4m-2, 4m-1)$ for some

m, say the mth inequality. In that case, the two sides of that inequality differ by a factor of at least 9.

Claim. We have
$$\prod_{k\geq 0} \frac{(4k+2)(4k+3)}{(4k+1)(4k+4)} < e.$$

Proof of claim using logarithms. To see this, note that it's equivalent to prove
$$\sum_{k\geq 0} \log\left(1 + \frac{2}{(4k+1)(4k+4)}\right) < 1.$$
To this end, we use the deep fact that $\log(1+t) \leq t$, and thus it follows from $\sum_{k\geq 0} \frac{1}{(4k+1)(4k+4)} < \frac{1}{2}$, which one can obtain for example by noticing it's less than $\frac{1}{4}\frac{\pi^2}{6}$. □

Elementary proof of claim, given by Espen Slettnes. For each $N \geq 0$, the partial product is bounded by
$$\begin{aligned}\prod_{k=0}^{N} \frac{(4k+2)(4k+3)}{(4k+1)(4k+4)} &= \frac{2}{1}\cdot\left(\frac{3}{4}\cdot\frac{6}{5}\right)\cdot\left(\frac{7}{8}\cdot\frac{10}{9}\right)\cdot\dots\cdot\frac{4N+3}{4N+4} \\ &< 2\cdot 1\cdot 1\cdot\dots\cdot\frac{4N+3}{4N+4} < 2 < e.\end{aligned}$$
□

This solves the problem, because then the factors being multiplied on by the positive inequalities before the mth one are both less than e, and $e^2 < 9$. In symbols, for $4m-2 < x < 4m-1$ we should have
$$\frac{(x-(4m-6))(x-(4m-5))}{(x-(4m-7))(x-(4m-4))} \times \dots \times \frac{(x-2)(x-3)}{(x-1)(x-4)} < e$$
and
$$\frac{(x-(4m+2))(x-(4m+3))}{(x-(4m+1))(x-(4m+4))} \times \dots \times \frac{(x-2014)(x-2015)}{(x-2013)(x-2016)} < e$$
because the $(k+1)$st term of each left-hand side is at most $\frac{(4k+2)(4k+3)}{(4k+1)(4k+4)}$, for $k \geq 0$. As $e^2 < 9$, we're okay.

Solution 115. Here is one of many possible constructions. We will prove one can form such a partition such that $\{2, 3, \dots, n+1\}$ are in different triples; let $P(n)$ denote this statement.

We make the following observation:

Fact. If $a < b < c$ is an obtuse triple, then so is $(a, b+x, c+x)$ for any $x > 0$.

Observe $P(1)$ is obviously true.

Claim. We have $P(n) \implies P(2n)$ for all $n \geq 1$.

Proof. Take the partition for $P(n)$ and use the observation to get a construction for $\{2, \dots, n+1\} \sqcup \{2n+2, \dots, 4n+1\}$. Now consider the following table:

$$\left[\begin{array}{cccc|cccc} 2 & 3 & \dots & n+1 & n+2 & n+3 & \dots & 2n+1 \\ \hline \multicolumn{4}{c|}{\text{Induct hypth}} & 4n+2 & 4n+3 & \dots & 5n+1 \\ \multicolumn{4}{c|}{+n} & 5n+2 & 5n+3 & \dots & 6n+1 \end{array}\right]$$

We claim all the column are obtuse. Indeed, they are obviously the sides of a triangle; now let $2 \leq k \leq n+1$ and note that

$$k^2 < 8n^2 \implies (n+k)^2 + (4n+k)^2 < (5n+k)^2$$

as desired. □

Claim. We have $P(n) \implies P(2n-1)$ for all $n \geq 2$.

Proof. Take the partition for $P(n)$ and use the observation to get a construction for $\{2, \dots, n+1\} \sqcup \{2n+1, \dots, 4n+1\}$. Now consider the following table:

$$\left[\begin{array}{cccc|cccc} 2 & 3 & \dots & n+1 & n+2 & n+3 & \dots & 2n \\ \hline \multicolumn{4}{c|}{\text{Induct hypth}} & 4n+1 & 4n+2 & \dots & 5n-1 \\ \multicolumn{4}{c|}{+n} & 5n & 5n+1 & \dots & 6n-2 \end{array}\right]$$

We claim all the columns are obtuse again. Indeed, they are obviously the sides of a triangle; now let $1 \leq k \leq n-1$ and note that

$$(k-2)^2 < 8n^2 - 12n + 4 \implies (n+1+k)^2 + (4n+k)^2 < (5n+k-1)^2$$

as desired. □

Together with the base case $P(1)$, we obtain $P(n)$ for all n.

Solution 116. The construction

$$P_n = \left(n - \frac{2014}{3}, \left(n - \frac{2014}{3} \right)^3 \right)$$

works fine, and follows from the following claim:

Claim. If x, y, z are distinct real numbers then the points (x, x^3), (y, y^3), (z, z^3) are collinear if and only if $x + y + z = 0$.

Proof. Note that by the "shoelace formula", the collinearity is equivalent to

$$0 = \det \begin{bmatrix} x & x^3 & 1 \\ y & y^3 & 1 \\ z & z^3 & 1 \end{bmatrix}$$

But the determinant equals

$$\sum_{\text{cyc}} x(y^3 - z^3) = (x-y)(y-z)(z-x)(x+y+z). \qquad \square$$

Solution 117. Here we present a solution with 14 registers and 22 moves. Initially $X = n$ and all other variables are zero.

	X	Y	Go	S_X^0	S_X	S_X'	S_Y^0	S_Y	S_Y'	Cl	A	B	Die	Die'
Init	-1		1									1	1	1
Begin	1		-1	1							-1	1		
Sleep											1	-1		
StartX				-1	1						-1	1		
WorkX	-1				-1	1					-1	1		
WorkX'	-1	1			1	-1					-1	1		
DoneX					-1		1				-1	1		
WrongX	-1						-1					-1		
StartY							-1	1			-1	1		
WorkY		-1						-1	1		-1	1		
WorkY'	1	-1						1	-1		-1	1		
DoneY				1				-1			-1	1		
WrongY		-1		-1								-1		
ClaimX	-1			-1						1	-1	1		
ClaimY		-1					-1			1	-1	1		
FakeX	-1									-1		-1		
FakeY		-1								-1		-1		
Win										-1	-1			
PunA												-2		
PunB											-1	-1		
Kill												-1	-2	1
Kill'												-1	1	-2

Now, the "game" is played as follows. The mechanics are controlled by the *turn counters* A and B.

Observe the game starts with Alice playing Init. Thereafter, we say that the game is

- In the *main part* if $A + B = 1$, and no one has played Init a second time.
- In the *death part* otherwise.

Observe that in the main state, on Alice's turn we always have $(A, B) = (1, 0)$ and on Bob's turn we always have $(A, B) = (0, 1)$.

Claim. A player who plays Init a second time must lose. In particular, a player who makes a move when $A = B = 0$ must lose.

Proof. Situations with $A + B \geq 2$ cannot occur during main part, so there are only a few possibilities.

- Suppose the offending player is in a situation where $A = B = 0$. Then he/she must play Init. At this point, the opposing player can respond by playing Kill. Then the offending player must play Init again. The opposing player now responds with Kill'. This iteration continues until X reaches a negative number and the offending player loses.
- Suppose Alice has $(A, B) = (1, 0)$ but plays Init again anyways. Then Bob responds with PunA to punish her; he then wins as in the first case.
- Suppose Bob has $(A, B) = (0, 1)$ but plays Init again anyways. Alice responds with PunB in the same way. □

Thus we may assume that players avoid the death part at all costs. Hence the second moves consist of Bob playing Sleep, and then Alice playing Begin (thus restoring the value of n in X), then Bob playing Sleep.

Now we return to analysis of the main part. We say the game is in *state* S for $S \in \{S_X^0, S_X, S_X', S_Y^0, S_Y, S_Y', \mathrm{Cl}\}$ if $S = 1$ and all other variables are zero. By construction, this is always the case. From then on the main part is divided into several phases:

- An *X-phase*: this begins with Alice at S_X^0, and ends when the game is in a state other than S_X and S_X'. (She can never return to S_X^0 during an X-phase.)
- A *Y-phase*: this begins with Alice at S_Y^0, and ends when the game is in a state other than S_Y and S_Y'. (She can never return to S_Y^0 during a Y-phase.)

Claim. Consider an X-phase in which $(X, Y) = (x, 0)$, $x > 1$. Then Alice can complete the phase without losing if and only if x is even; if so she begins a Y-phase with $(X, Y) = (0, x/2)$.

Proof. As $x > 1$, Alice cannot play ClaimX since Bob will respond with FakeX and win. Now by alternating between WorkX and WorkX', Alice can repeatedly deduct 2 from X and add 1 to Y, leading to $(x - 2, y + 1)$, $(x - 4, y + 2)$, and so on. (During this time, Bob can only play Sleep.) Eventually, she must stop this process by playing DoneX, which begins a Y-phase.

Now note that unless $X = 0$, Bob now has a winning move WrongX. Conversely he may only play Sleep if $X = 0$. □

We have an analogous claim for Y-phases. Thus if n is not a power of 2, we see that Alice eventually loses.

Now suppose $n = 2^k$; then Alice reaches $(X, Y) = (0, 2^{k-1})$, $(2^{k-2}, 0)$, ... until either reaching $(1, 0)$ or $(0, 1)$. At this point she can play ClaimX or ClaimY, respectively; the game is now in state Cl. Bob cannot play either FakeX or FakeY, so he must play Sleep, and then Alice wins by playing Win. Thus Alice has a winning strategy when $n = 2^k$.

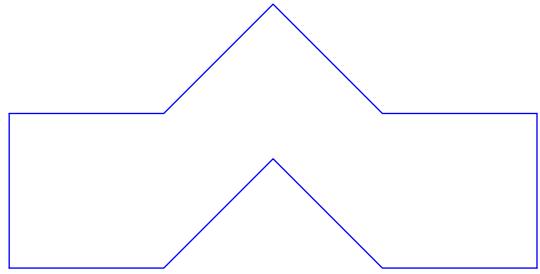
11 Anti-Problems

This is a silly chapter dedicated to problems whose solutions make you groan. No particular theory, other than advice: keep things simple.

§11.1 Walkthroughs

Problem 118. Let $ABCDEZYXWV$ be an equilateral decagon with interior angles $\angle A = \angle V = \angle E = \angle Z = \angle C = 90°$, $\angle W = \angle Y = 135°$, $\angle B = \angle D = 225°$, and $\angle X = 270°$. Determine whether or not one can dissect $ABCDEZYXWV$ into four congruent polygons.

Walkthrough. Stare at the shape until you either give up or figure out what's going on.



Problem 119 (Math Prize for Girls 2017)**.** Define a *lattice line* as a line containing at least 2 lattice points. Is it possible to color every lattice point red or blue such that every lattice line contains exactly 2017 red points?

Walkthrough. Let L_1, L_2, ...denote the countably many lattice lines, in some order. It is not hard to do the "finite" step:

(a) Show that for every integer n, we can construct a set T_n of lattice points such that each line L_1, ..., L_n passes through exactly 2017 points in T_n.

(b) Make sure your solution to (a) works. Is it possible to get stuck because you accidentally colored 2018 points on L_N already for some N in the future?

The issue is that we need a set T_∞ that works for all lines at once: there is a difference between "unbounded" and "infinite"! Put another way, we have proven the statement $P(n)$ that "there exists a set T_n as in (a)" for every $n = 1, 2, \ldots$, by induction of the usual shape $P(n) \implies P(n+1)$ but we really need the statement $P(\infty)$, which we cannot reach by using a normal induction. Thus, we need to do a little more work.

(c) Modify your approach to (a) such that we have the additional property $T_1 \subseteq T_2 \subseteq \dots$. (For some people, no additional modification is needed.)

(d) Prove that

$$T_\infty = \bigcup_{n \geq 1} T_n$$

fits the bill.

(e) Why was part (c) necessary? (In other words, what goes wrong if you try to fix over-red lines retroactively?)

As an aside, this is sort of a simple case of a "transfinite induction": the last step breaks the realm of normal induction and brings us into the world of statements $P(\alpha)$ for infinite ordinals α. In set theory, transfinite induction proves a statement $P(-)$ for any *ordinal* α, and this proof typically involves both a successor case $P(\alpha) \implies P(\alpha+1)$, as well as a limit case similar to the above.

§11.2 Problems

Problem 120 (Russian Olympiad 2015)**.** We define a *chessboard polygon* to be a polygon in the xy-plane whose edges are situated along lines of the form $x = a$ and $y = b$, where a and b are integers. These lines divide the interior into unit squares, which we call cells.

Let n and k be positive integers. Assume that a square can be partitioned into n congruent chessboard polygons of k cells each. Prove that this square may also be partitioned into k congruent chessboard polygons of n cells each.

Problem 121 (IMO Shortlist 2016)**.** Find all positive integers n for which it is possible to arrange all positive divisors of n (including 1 and n) in a rectangular grid of some size (with all cells filled) such that

- each divisor appears exactly once,
- all columns have equal sum,
- all rows have equal sum.

Problem 122. Show that one can find a set S of 210 distinct points in $\mathbb{R}^{20}$ and two positive real numbers a and b, such that for any two distinct points in S, the distance between them is either a or b.

Problem 123 (Putnam 2018 B6)**.** Prove that the number of length 2018-tuples whose entries are in $\{1, 2, 3, 4, 5, 6, 10\}$ and sum to 3860, is at most

$$2^{3860} \cdot \left(\frac{2018}{2048}\right)^{2018}.$$

Problem 124 (USAMO 2002)**.** Prove that any monic polynomial (a polynomial with leading coefficient 1) of degree n with real coefficients is the average of two monic polynomials of degree n with n real roots.

Problem 125 (China TST 2016)**.** In the coordinate plane the points with both coordinates being rational numbers are called rational points. For any positive integer n, is there a way to use n colours to colour all rational points, every point is coloured one colour, such that any line segment with both endpoints being rational points contains the rational points of every colour?

Problem 126 (Ankan Bhattacharya)**.** A *diamond* is a rhombus with side length 1 whose interior angles are 60° and 120° (hence with area $\sqrt{3}/2$). A regular hexagon $\mathcal{H}$ of side length 10 is dissected into diamonds. In a move, if three pairwise adjacent diamonds form a regular hexagon of side length 1, one may rotate all three 60° about the center of that hexagon.

Find the smallest positive integer N such that any tiling of $\mathcal{H}$ can be transformed into any other in at most N moves, or show that no such N exists.

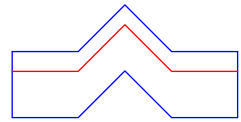
Problem 127 (USA TST 2013)**.** In a table with n rows and $2n$ columns where n is a fixed positive integer, we write either zero or one into each cell so that each row has n zeros and n ones. For $1 \le k \le n$ and $1 \le i \le n$, we define $a_{k,i}$ so that the i^{th} zero in the k^{th} row is the ${a_{k,i}}^{\text{th}}$ column. Let $\mathcal{F}$ be the set of such tables with $a_{1,i} \ge a_{2,i} \ge \cdots \ge a_{n,i}$ for every i with $1 \le i \le n$. We associate another $n \times 2n$ table $f(C)$ for each $C \in \mathcal{F}$ as follows: for the k^{th} row of $f(C)$, we write n ones in the columns $a_{n,k} - k + 1, a_{n-1,k} - k + 2, \ldots, a_{1,k} - k + n$ (and we write zeros in the other cells in the row).

(a) Show that $f(C) \in \mathcal{F}$.

(b) Show that $f(f(f(f(f(f(C)))))) = C$ for any $C \in \mathcal{F}$.

§11.3 Solutions

Solution 118. Make a stack of four copies of the same polygon except with $AV = EZ = 1/4$.



Solution 119. Let L_1, L_2, ...denote the countably many lattice lines, in some order. We construct by induction a set T_n of lattice points (for each $n \geq 1$) such that each line L_1, ..., L_n passes through exactly 2017 points in T_n.

To do this, at the nth step, we take T_{n-1} and add in between 2015 and 2017 red points on L_n such that

- no red point we add is on any of L_1, ..., L_{n-1}, and
- no red point we add is collinear with any two red points in T_{n-1}. (This ensures that at future steps of the algorithm, each line passes through at most two red points already).

Finally, note that our construction has the property that $T_1 \subseteq T_2 \subseteq \dots$; thus the union

$$T_\infty = \bigcup_{n \geq 1} T_n$$

satisfies the construction.

Remark. One incorrect approach is to try and edit the choice of red points retroactively if the line L_n is already full. This makes it impossible to take the union at the last step.

Solution 120. Let the side length of the square be s. Because of the partition, we have

$$nk = s^2.$$

By the so-called factor lemma, there exist positive integers a, b, c, d with $n = ab$, $k = cd$, and $s = ac = bd$.

Therefore, we can tile this $s \times s = ac \times bd$ square with $a \times b$ rectangles! Ha, ha, ha...

Solution 121. This is a somewhat silly problem — it's impossible for size reasons, except in the trivial situation $n = 1$. (One can gain this intuition very quickly from small cases. I solved this problem during the Synco concert.)

Suppose the grid has dimensions a rows and b columns, $a \geq b > 1$ (the $b = 1$ situation gives $n = 1$).

Clearly the common sum is more than n. On the other hand, there are at most $b - 1$ divisors exceeding $\frac{n}{b}$. Since there are $a > b - 1$ rows, some row has all entries at most n/b. So that row has sum at most $b \cdot n/b = n$, impossible.

Solution 122. Take the $\binom{21}{2} = 210$ points on the hyperplane

$$x_0 + x_1 + \cdots + x_{20} = 2$$

which have two coordinates equal to one, and the others zero. The hyperplane is a 20-dimensional space, and using $a = \sqrt{2}$, $b = 2$ works fine,

---

Solution 123. Let a_n be the number of n-tuples if 3860 is replaced by n. Consider the usual generating function

$$F(X) = (X^1 + X^2 + X^3 + X^4 + X^5 + X^6 + X^{10})^{2018} = \sum_n a_n X^n.$$

Observe that

$$F\left(\frac{1}{2}\right) = \left(\frac{1009}{1024}\right)^{2018} = \sum_n a_n \left(\frac{1}{2}\right)^n$$

and hence $a_n \leq 2^n \left(\frac{1009}{1024}\right)^{2018}$ for any integer n.

Remark. Alexendar Givental notes that the bound $X^{-3860}F(X)$ is minimized when $X = \frac{1}{2}$ (which one can check by computing the derivative), i.e. we don't get a better bound by replacing X. Therefore, this gives a reason why 3860 might have been chosen.

Solution 124. First,

Lemma. *If p is a monic polynomial of degree n, and $p(1)p(2) < 0$, $p(2)p(3) < 0$, ..., $p(n-1)p(n) < 0$ then p has n real roots.*

Proof. The intermediate value theorem already guarantees the existence of $n - 1$ real roots.

The last root is obtained by considering cases on $n \pmod 2$. If n is even, then $p(1)$ and $p(n)$ have opposite sign, while we must have either

$$\lim_{x \to -\infty} p(x) = \lim_{x \to \infty} p(x) = \pm\infty$$

so we get one more root. The n odd case is similar, with $p(1)$ and $p(n)$ now having the same sign, but $\lim_{x\to-\infty} p(x) = -\lim_{x\to\infty} p(x)$ instead. □

Let $f(n)$ be the monic polynomial and let $M > 1000 \max_{t=1,\dots,n} |f(t)|+1000$. Then we may select reals $a_1, \dots, a_n$ and $b_1, \dots, b_n$ such that for each $k = 1, \dots, n$, we have

$$\begin{aligned} a_k + b_k &= 2f(k) \\ (-1)^k a_k &> M \\ (-1)^{k+1} b_k &> M. \end{aligned}$$

We may interpolate monic polynomials g and h through the a_k and b_k (if the a_k, b_k are selected "generically" from each other). Then one can easily check $f = \frac{1}{2}(g+h)$ works.

Remark. This is like Cape Town all over again...

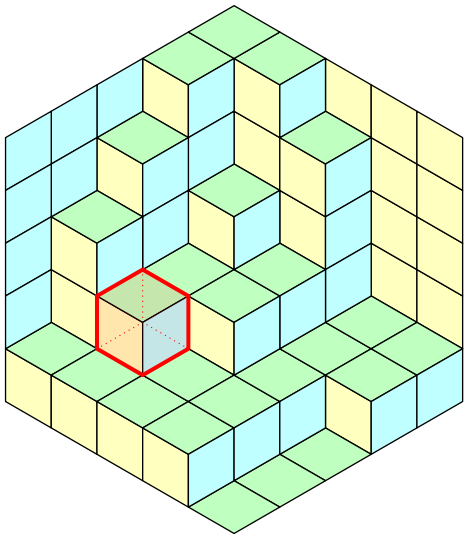
Solution 125. Always possible.

The number of rational points is countable, and so is the number of line segments with rational endpoints. Let us list these segments as $s_1, s_2, \dots$. First pick n arbitrary rational points on s_1 and give them distinct colors. Then do the same for s_2, except that we need to avoid choosing points that have already been colored. But that is possible because s_2 contains infinitely many rational points. Keep doing this for each s_k, and the desired conclusion follows. To be complete, give the uncolored points any color at the end.

Remark. CeuAzul in `https://aops.com/community/q3h1299532p6916022` recounts a story:

The day before the TST teacher QuZhenHua noticed this and said it will be obvious using countable, but other teachers don't think many students would use this, so they put on a bet. Out of 60 students, the other teachers bet the number of people who use this way is strictly less than 5 while Qu said more than 5. The result is, only 4 students during the contest used this.

Solution 126. As in the AOPS logo, view the picture in 3-D. Then one imagines a cube of side length 10, which is filled with some unit cubes under gravity. The operation consists of either adding or removing a visible unit cube.



To finish, there are $10^3 = 1000$ unit cubes and we claim the answer is just $N = 1000$.

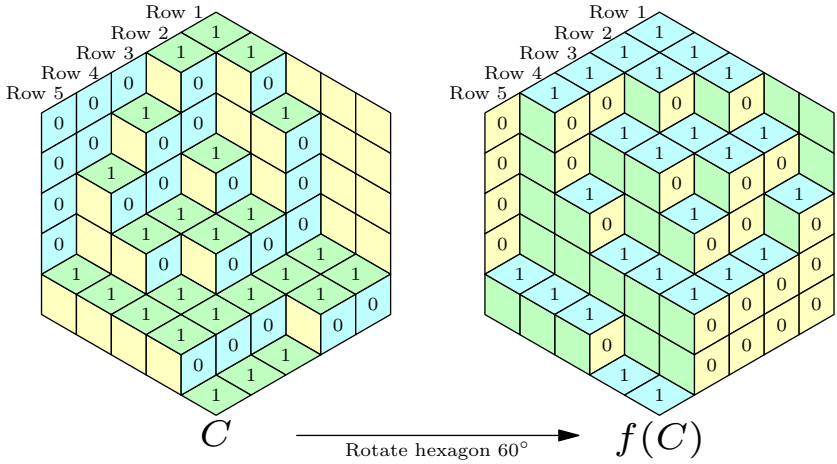
Suppose there are a unit cubes in our starting configuration and b unit cubes in our ending one. If we remove all a unit cubes and rebuild to get the desired configuration of b unit cubes in the end, this takes $a + b$ steps. On the other hand, if we add unit cubes until we have all 1000, and then delete down to our desired configuration, this takes $(1000 - a) + (1000 - b)$ steps. Now, $\min((a + b), 2000 - (a + b)) \le 1000$, proving the bound. To see the bound cannot be improved, consider the configuration with $a = 0$ and $b = 1000$; it always takes at least 1000 steps to fill.

Remark. One other nice corollary of the 3D perspective is that the number of diamonds of each orientation is always equal.

Solution 127. Part (a) is easy and (b) is typically just a very long calculation.

The official solution to (b) is quite nice, but it is essentially completely unmotivated. Nonetheless, since I don't want to type the long calculation (you can find plenty of those on AOPS), here is the "nice" solution.

We give an interpretation of C in terms of the AOPS logo. Consider subsets cubes as shown below, supported by gravity in all three directions. Write 1 on the top of every cube, 0 on the right face. Then we can read off the rows of a $2n \times n$ table in the obvious way.



In that case, $f(C)$ corresponds to rotating the hexagon 60°. So $f(C) \in \mathcal{F}$ and $f^6(C) = \mathrm{id}$.

---

12 Selected Combinatorics from USA TST

§12.1 Problems

Problem 128 (USA TST 2015)**.** Fix a positive integer n. A tournament on n vertices has all its edges colored by χ colors, so that any two directed edges $u \to v$ and $v \to w$ have different colors. Over all possible tournaments on n vertices, determine the minimum possible value of χ.

Problem 129 (TSTST 2018)**.** In the nation of Onewaynia, certain pairs of cities are connected by one-way roads. Every road connects exactly two cities (roads are allowed to cross each other, e.g., via bridges), and each pair of cities has at most one road between them. Moreover, every city has exactly two roads leaving it and exactly two roads entering it.

We wish to close half the roads of Onewaynia in such a way that every city has exactly one road leaving it and exactly one road entering it. Show that the number of ways to do so is a power of 2 greater than 1 (i.e. of the form 2^n for some integer $n \geq 1$).

Problem 130 (TSTST 2018)**.** Let n be a positive integer. A frog starts on the number line at 0. Suppose it makes a finite sequence of hops, subject to two conditions:

- The frog visits only points in $\{1, 2, \dots, 2^n - 1\}$, each at most once.
- The length of each hop is in $\{2^0, 2^1, 2^2, \dots\}$. (The hops may be either direction, left or right.)

Let S be the sum of the (positive) lengths of all hops in the sequence. What is the maximum possible value of S?

Problem 131 (TSTST 2016)**.** In the coordinate plane are finitely many *walls*, which are disjoint line segments, none of which are parallel to either axis. A bulldozer starts at an arbitrary point and moves in the $+x$ direction. Every time it hits a wall, it turns at a right angle to its path, away from the wall, and continues moving. (Thus the bulldozer always moves parallel to the axes.)

Prove that it is impossible for the bulldozer to hit both sides of every wall.

Problem 132 (USA TST 2017)**.** You are cheating at a trivia contest. For each question, you can peek at each of the $n > 1$ other contestant's guesses before writing your own. For each question, after all guesses are submitted,

the emcee announces the correct answer. A correct guess is worth 0 points. An incorrect guess is worth -2 points for other contestants, but only -1 point for you, because you hacked the scoring system. After announcing the correct answer, the emcee proceeds to read out the next question. Show that if you are leading by 2^{n-1} points at any time, then you can surely win first place.

Problem 133 (TSTST 2018)**.** Show that there is an absolute constant $c < 1$ with the following property: whenever $\mathcal{P}$ is a polygon with area 1 in the plane, one can translate it by a distance of $\frac{1}{100}$ in some direction to obtain a polygon $\mathcal{Q}$, for which the intersection of the interiors of $\mathcal{P}$ and $\mathcal{Q}$ has total area at most c.

Problem 134 (USAMO 2017)**.** Find all real numbers $c > 0$ such that there exists a labeling of the lattice points in $\mathbb{Z}^2$ with positive integers for which:

- only finitely many distinct labels occur, and
- for each label i, the distance between any two points labeled i is at least c^i.

Problem 135 (USA TST 2019)**.** A *snake of length* k is an animal which occupies an ordered k-tuple $(s_1, \dots, s_k)$ of cells in an $n \times n$ grid of square unit cells. These cells must be pairwise distinct, and s_i and s_{i+1} must share a side for $i = 1, \dots, k-1$. If the snake is currently occupying $(s_1, \dots, s_k)$ and s is an unoccupied cell sharing a side with s_1, the snake can *move* to occupy $(s, s_1, \dots, s_{k-1})$ instead. The snake has *turned around* if it occupied $(s_1, s_2, \dots, s_k)$ at the beginning, but after a finite number of moves occupies $(s_k, s_{k-1}, \dots, s_1)$ instead.

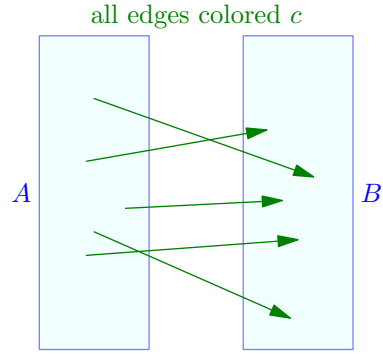
Determine whether there exists an integer $n > 1$ such that one can place some snake of length at least $0.9n^2$ in an $n \times n$ grid which can turn around.

§12.2 Solutions

Solution 128. The answer is

$$\chi = \lceil \log_2 n \rceil .$$

First, we prove by induction on n that $\chi \geq \log_2 n$ for any coloring and any tournament. The base case $n = 1$ is obvious. Now given any tournament, consider any used color c. Then it should be possible to divide the tournament into two subsets A and B such that all c-colored edges point from A to B (for example by letting A be all vertices which are the starting point of a c-edge).



One of A and B has size at least $n/2$, say A. Since A has no c edges, and uses at least $\log_2 |A|$ colors other than c, we get

$$\chi \geq 1 + \log_2(n/2) = \log_2 n$$

completing the induction.

One can read the construction off from the argument above, but here is a concrete description. For each integer n, consider the tournament whose vertices are the binary representations of $S = \{0, \dots, n-1\}$. Instantiate colors c_1, c_2, ... Then for $v, w \in S$, we look at the smallest order bit for which they differ; say the kth one. If v has a zero in the kth bit, and w has a one in the kth bit, we draw $v \to w$. Moreover we color the edge with color c_k. This works and uses at most $\lceil \log_2 n \rceil$ colors.

Remark (Motivation)**.** The philosophy "combinatorial optimization" applies here. The idea is given any color c, we can find sets A and B such that all c-edges point A to B. Once you realize this, the next insight is to realize that you may as well color *all* the edges from A to B by c; after all, this doesn't hurt the condition and makes your life easier. Hence, if f is the answer, we have already a proof that $f(n) = 1 + \max(f(|A|), f(|B|))$ and we choose $|A| \approx |B|$. This optimization also gives the inductive construction.

Solution 129. In the language of graph theory, we have a simple digraph G which is 2-regular and we seek the number of sub-digraphs which are 1-regular. We now present two solution paths.

First solution, combinatorial We construct a simple undirected bipartite graph Γ as follows:

- the vertex set consists of two copies of $V(G)$, say V_{out} and V_{in}; and
- for $v \in V_{\text{out}}$ and $w \in V_{\text{in}}$ we have an undirected edge $vw \in E(\Gamma)$ if and only if the directed edge $v \to w$ is in G.

Moreover, the desired sub-digraphs of H correspond exactly to perfect matchings of Γ.

However the graph Γ is 2-regular and hence consists of several disjoint (simple) cycles of even length. If there are n such cycles, the number of perfect matchings is 2^n, as desired.

Remark. The construction of Γ is not as magical as it may first seem.

Suppose we pick a road $v_1 \to v_2$ to use. Then, the other road $v_3 \to v_2$ is certainly *not* used; hence some other road $v_3 \to v_4$ must be used, etc. We thus get a cycle of forced decisions until we eventually return to the vertex v_1.

These cycles in the original graph G (where the arrows alternate directions) correspond to the cycles we found in Γ. It's merely that phrasing the solution in terms of Γ makes it cleaner in a linguistic sense, but not really in a mathematical sense.

Second solution by linear algebra over $\mathbb{F}_2$ (Brian Lawrence) This is actually not that different from the first solution. For each edge e, we create an indicator variable x_e. We then require for each vertex v that:

- If e_1 and e_2 are the two edges leaving v, then we require $x_{e_1} + x_{e_2} \equiv 1 \pmod 2$.
- If e_3 and e_4 are the two edges entering v, then we require $x_{e_3} + x_{e_4} \equiv 1 \pmod 2$.

We thus get a large system of equations. Moreover, the solutions come in natural pairs $\vec{x}$ and $\vec{x} + \vec{1}$ and therefore the number of solutions is either zero, or a power of two. So we just have to prove there is at least one solution.

For linear algebra reasons, there can only be zero solutions if some nontrivial linear combination of the equations gives the sum $0 \equiv 1$. So suppose we added up some subset S of the equations for which every variable appeared on the left-hand side an even number of times. Then every variable that did appear appeared exactly twice; and accordingly we see that the edges corresponding

to these variables form one or more even cycles as in the previous solution. Of course, this means $|S|$ is even, so we really have $0 \equiv 0 \pmod 2$ as needed.

Remark. The author's original proposal contained a second part asking to show that it was not always possible for the resulting H to be connected, even if G was strongly connected. This problem is related to IMO Shortlist 2002 C6, which gives an example of a strongly connected graph which does have a full directed Hamiltonian cycle.

Solution 130. We claim the answer is $\frac{4^n-1}{3}$.

We first prove the bound. First notice that the hop sizes are in $\{2^0, 2^1, \dots, 2^{n-1}\}$, since the frog must stay within bounds the whole time. Let a_i be the number of hops of size 2^i the frog makes, for $0 \le i \le n-1$.

Claim. For any $k = 1, \dots, n$ we have

$$a_{n-1} + \cdots + a_{n-k} \le 2^n - 2^{n-k}.$$

Proof. Let $m = n - k$ and look modulo 2^m. Call a jump *small* if its length is at most 2^{m-1}, and *large* if it is at least 2^m; the former changes the residue class of the frog modulo 2^m while the latter does not.

Within each fixed residue modulo 2^m, the frog can make at most $\frac{2^n}{2^m} - 1$ large jumps. So the total number of large jumps is at most $2^m \left(\frac{2^n}{2^m} - 1\right) = 2^n - 2^m$. □

(As an example, when $n = 3$ this means there are at most four hops of length 4, at most six hops of length 2 or 4, and at most seven hops total. Of course, if we want to max the length of the hops, we see that we want $a_2 = 4$, $a_1 = 2$, $a_0 = 1$, and in general equality is achieved when $a_m = 2^m$ for any m.)

Now, the total distance the frog travels is

$$S = a_0 + 2a_1 + 4a_2 + \cdots + 2^{n-1}a_{n-1}.$$

We rewrite using the so-called "summation by parts":

$$\begin{aligned}
S = a_0 + a_1 + \ & a_2 + \ a_3 + \ldots + \ a_{n-1} \\
+ \quad a_1 + \ & a_2 + \ a_3 + \ldots + \ a_{n-1} \\
+ \quad 2a_2 + \ & 2a_3 + \ldots + \ 2a_{n-1} \\
+ \quad 4a_3 + \ & \ldots + \ 4a_{n-1} \\
\vdots \quad \ddots \quad & \vdots \\
+ \quad 2^{n-2}a_{n-1}&.
\end{aligned}$$

Hence

$$\begin{aligned} S &\le (2^n - 2^0) + (2^n - 2^1) + 2(2^n - 2^2) + \cdots + 2^{n-2}(2^n - 2^{n-1}) \\ &= \frac{4^n - 1}{3}. \end{aligned}$$

It remains to show that equality can hold. There are many such constructions but most are inductive. Here is one approach. We will construct two family of paths such that there are 2^k hops of size 2^k, for every $0 \le k \le n-1$, and we visit each of $\{0, \dots, 2^n - 1\}$ once, starting on 0 and ending on x, for the two values $x \in \{1, 2^n - 1\}$.

The base case $n = 1$ is clear. To take a path from 0 to $2^{n+1} - 1$.

- Take a path on $\{0, 2, 4, \dots, 2^{n+1} - 2\}$ starting from 0 and ending on 2 (by inductive hypothesis).
- Take a path on $\{1, 3, 5, \dots, 2^{n+1} - 1\}$ starting from 1 and ending on $2^{n+1} - 1$ (by inductive hypothesis).
- Link them together by adding a single jump $2 \to 1$.

The other case is similar, but we route $0 \to (2^{n+1} - 2) \to (2^{n+1} - 1) \to 1$ instead. (This can also be visualized as hopping along a hypercube of binary strings; each inductive step takes two copies of the hypercube and links them together by a single edge.)

Remark (Ashwin Sah)**.** The problem can also be altered to ask for the minimum value of the sum of the reciprocals of the hop sizes, where further we stipulate that the frog must hit every point precisely once (to avoid triviality). With a nearly identical proof that also exploits the added condition $a_0 + \cdots + a_{n-1} = 2^n - 1$, the answer is n. This yields a nicer form for the generalization. The natural generalization changes the above problem by replacing 2^k with a_k where $a_k \mid a_{k+1}$, so that the interval covered by hops is of size a_n and the hop sizes are restricted to the a_i, where $a_0 = 1$. In this case, similar bounding yields

$$\sum_{i=1}^{2^n - 1} \frac{1}{b_k} \ge \sum_{i=0}^{n-1} \left(\frac{a_{k+1}}{a_k} - 1 \right).$$

Bounds for the total distance traveled happen in the same way as the solution above, and equality for both can be constructed in an analogous fashion.

Solution 131. We say a wall v is *above* another wall w if some point on v is directly above a point on w. (This relation is anti-symmetric, as walls do not intersect).

The critical claim is as follows:

Claim. There exists a lowest wall, i.e. a wall not above any other walls.

Proof. Assume not. Then we get a directed cycle of some length $n \geq 3$: it's possible to construct a series of points P_i, Q_i, for $i = 1, \dots, n$ (indices modulo n), such that the point Q_i is directly above P_{i+1} for each i, the segment $\overline{Q_iP_{i+1}}$ does not intersect any wall in its interior, and finally each segment $\overline{P_iQ_i}$ is contained inside a wall. This gives us a broken line on $2n$ vertices which is not self-intersecting.

Now consider the leftmost vertical segment $\overline{Q_iP_{i+1}}$ and the rightmost vertical segment $\overline{Q_jP_{j+1}}$. The broken line gives a path from P_{i+1} to Q_j, as well as a path from P_{j+1} to Q_i. These clearly must intersect, contradiction. □

Remark. This claim is Iran TST 2010.

Thus if the bulldozer eventually moves upwards indefinitely, it may never hit the bottom side of the lowest wall. Similarly, if the bulldozer eventually moves downwards indefinitely, it may never hit the upper side of the highest wall.

Solution 132. We will prove the result with 2^{n-1} replaced even by $2^{n-2}+1$.

We first make the following reductions. First, change the weights to be $+1$, -1, 0 respectively (rather than 0, -2, -1); this clearly has no effect. Also, WLOG that all contestants except you initially have score zero (and that your score exceeds 2^{n-2}). WLOG ignore rounds in which all answers are the same. Finally, ignore rounds in which you get the correct answer, since that leaves you at least as well off as before — in other words, we'll assume your score is always fixed, but you can pick any group of people with the same answers and ensure they lose 1 point, while some other group gains 1 point.

The key observation is the following. Consider two rounds R_1 and R_2 such that:

- In round R_1, some set S of contestants gains a point.
- In round R_2, the set S of contestants all have the same answer.

Then, if we copy the answers of contestants in S during R_2, then the sum of the scorings in R_1 and R_2 cancel each other out. In other words we can then ignore R_1 and R_2 forever.

We thus consider the following strategy. We keep a list $\mathcal{L}$ of subsets of $\{1, \dots, n\}$, initially empty. Now do the following strategy:

- On a round, suppose there exists a set S of people with the same answer such that $S \in \mathcal{L}$. (If multiple exist, choose one arbitrarily.) Then, copy the answer of S, causing them to lose a point. Delete S from $\mathcal{L}$. (Importantly, we do not add any new sets to $\mathcal{L}$.)

- Otherwise, copy any set T of contestants, selecting $|T| \geq n/2$ if possible. Let S be the set of contestants who answer correctly (if any), and add S to the list $\mathcal{L}$. Note that $|S| \leq n/2$, since S is disjoint from T.

By construction, $\mathcal{L}$ has no duplicate sets. So the score of any contestant c is bounded above by the number of times that c appears among sets in $\mathcal{L}$. The number of such sets is clearly at most $\frac{1}{2} \cdot 2^{n-1}$. So, if you lead by $2^{n-2}+1$ then you ensure victory. This completes the proof!

Remark. Several remarks are in order. First, we comment on the bound $2^{n-2}+1$ itself. The most natural solution using only the list idea gives an upper bound of $(2^n-2)+1$, which is the number of nonempty proper subsets of $\{1,\dots,n\}$. Then, there are two optimizations one can observe:

- In fact we can improve to the number of times any particular contestant c appears in some set, rather than the total number of sets.
- When adding new sets S to $\mathcal{L}$, one can ensure $|S| \leq n/2$.

Either observation alone improves the bound from 2^n-1 to 2^{n-1}, but both together give the bound $2^{n-2}+1$. Additionally, when n is odd the calculation of subsets actually gives $2^{n-2} - \frac{1}{2}\binom{n-1}{\frac{n-1}{2}} + 1$. This gives the best possible value at both $n=2$ and $n=3$. It seems likely some further improvements are possible, and the true bound is suspected to be polynomial in n.

Secondly, the solution is highly motivated by considering a true/false contest in which only two distinct answers are given per question. However, a very natural mistake (which graders assessed as a two-point deduction) is to try and prove that in fact one can "WLOG" we are in the two-question case. The proof of this requires substantially more care than expected. For instance, set $n=3$. If $\mathcal{L} = \{\{1\},\{2\},\{3\}\}$ then it becomes impossible to prevent a duplicate set from appearing in $\mathcal{L}$ if all contestants give distinct answers. One might attempt to fix this by instead adding to $\mathcal{L}$ the *complement* of the set T described above. The example $\mathcal{L} = \{\{1,2\},\{2,3\},\{3,1\}\}$ (followed again by a round with all distinct answers) shows that this proposed fix does not work either. This issue affects all variations of the above approach.

Remark. Here are some motivations for the solution:

1. The exponential bound 2^n suggests looking at subsets.
2. The $n=2$ case suggests the idea of "repeated rounds". (I think this $n=2$ case is actually really good.)
3. The "two distinct answers" case suggests looking at rounds as partitions (even though the WLOG does not work, at least not without further thought).
4. There's something weird about this problem: it's a finite bound over unbounded time. This is a hint to *not* worry excessively about the actual scores, which turn out to be almost irrelevant.

Solution 133. The following solution is due to Brian Lawrence. We will prove the result with the generality of any measurable set $\mathcal{P}$ (rather than a polygon). For a vector v in the plane, write $\mathcal{P}+v$ for the translate of $\mathcal{P}$ by v.

Suppose $\mathcal{P}$ is a polygon of area 1, and $\varepsilon > 0$ is a constant, such that for any translate $\mathcal{Q} = \mathcal{P} + v$, where v has length exactly $\frac{1}{100}$, the intersection of $\mathcal{P}$ and $\mathcal{Q}$ has area at least $1 - \varepsilon$. The problem asks us to prove a lower bound on ε.

Lemma. *Fix a sequence of n vectors v_1, v_2, ..., v_n, each of length $\frac{1}{100}$. A grasshopper starts at a random point x of $\mathcal{P}$, and makes n jumps to $x + v_1 + \cdots + v_n$. Then it remains in $\mathcal{P}$ with probability at least $1 - n\varepsilon$.*

Proof. In order for the grasshopper to leave $\mathcal{P}$ at step i, the grasshopper's position before step i must be inside the difference set $\mathcal{P} \backslash (\mathcal{P} - v_i)$. Since this difference set has area at most ε, the probability the grasshopper leaves $\mathcal{P}$ at step i is at most ε. Summing over the n steps, the probability that the grasshopper ever manages to leave $\mathcal{P}$ is at most $n\varepsilon$. □

Corollary. *Fix a vector w of length at most 8. A grasshopper starts at a random point x of $\mathcal{P}$, and jumps to $x + w$. Then it remains in $\mathcal{P}$ with probability at least $1 - 800\varepsilon$.*

Proof. Apply the previous lemma with 800 jumps. Any vector w of length at most 8 can be written as $w = v_1 + v_2 + \cdots + v_{800}$, where each v_i has length exactly $\frac{1}{100}$. □

Now consider the process where we select a random starting point $x \in \mathcal{P}$ for our grasshopper, *and* a random vector w of length at most 8 (sampled uniformly from the closed disk of radius 8). Let q denote the probability of staying inside $\mathcal{P}$ we will bound q from above and below.

- On the one hand, suppose we pick w first. By the previous corollary, $q \geq 1 - 800\varepsilon$ (irrespective of the chosen w).
- On the other hand, suppose we pick x first. Then the possible landing points $x + w$ are uniformly distributed over a closed disk of radius 8, which has area 64π. The probability of landing in $\mathcal{P}$ is certainly at most $\frac{[\mathcal{P}]}{64\pi}$.

Consequently, we deduce

$$1 - 800\varepsilon \leq q \leq \frac{[\mathcal{P}]}{64\pi} \implies \varepsilon > \frac{1 - \frac{[\mathcal{P}]}{64\pi}}{800} > 0.001$$

as desired.

Remark. The choice of 800 jumps is only for concreteness; any constant n for which $\pi(n/100)^2 > 1$ works. I think $n = 98$ gives the best bound following this approach.

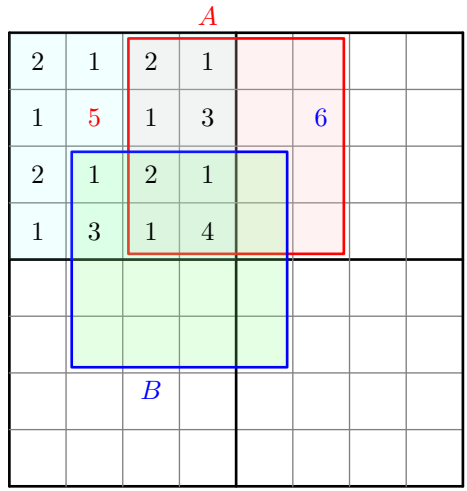
Solution 134. The answer is $c < \sqrt{2}$. Here is a solution with Calvin Deng.

The construction for any $c < \sqrt{2}$ can be done as follows. Checkerboard color the lattice points and label the black ones with 1. The white points then form a copy of $\mathbb{Z}^2$ again scaled up by $\sqrt{2}$ so we can repeat the procedure with 2 on half the resulting points. Continue this dyadic construction until a large N for which $c^N < 2^{\frac{1}{2}(N-1)}$, at which point we can just label all the points with N.

I'll now prove that $c = \sqrt{2}$ (and hence $c \geq \sqrt{2}$) can't be done.

Claim. It is impossible to fill a $2^n \times 2^n$ square with labels not exceeding $2n$.

The case $n = 1$ is clear. So now assume it's true up to $n - 1$; and assume for contradiction a $2^n \times 2^n$ square S only contains labels up to $2n$. (Of course every $2^{n-1} \times 2^{n-1}$ square contains an instance of a label at least $2n - 1$.)



Now, we contend there are fewer than four copies of $2n$:

Lemma. *In a unit square, among any four points, two of these points have distance ≤ 1 apart.*

Proof. Look at the four rays emanating from the origin and note that two of them have included angle $\leq 90°$. □

So WLOG the northwest quadrant has no $2n$'s. Take a $2n - 1$ in the northwest and draw a square of size $2^{n-1} \times 2^{n-1}$ directly right of it (with its top

edge coinciding with the top of S). Then A can't contain $2n-1$, so it must contain a $2n$ label; that $2n$ label must be in the northeast quadrant.

Then we define a square B of size $2^{n-1} \times 2^{n-1}$ as follows. If $2n-1$ is at least as high $2n$, let B be a $2^{n-1} \times 2^{n-1}$ square which touches $2n-1$ north and is bounded east by $2n$. Otherwise let B be the square that touches $2n-1$ west and is bounded north by $2n$. We then observe B can neither have $2n-1$ nor $2n$ in it, contradiction.

Remark. To my knowledge, essentially all density arguments fail because of hexagonal lattice packing.

Solution 135. The answer is yes (and 0.9 is arbitrary).

First grid-based solution The following solution is due to Brian Lawrence. For illustration reasons, we give below a figure of a snake of length 89 turning around in an 11×11 square (which generalizes readily to odd n). We will see that a snake of length $(n-1)(n-2)-1$ can turn around in an $n \times n$ square, so this certainly implies the problem.

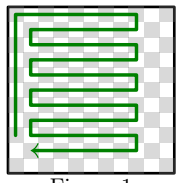


Figure 1

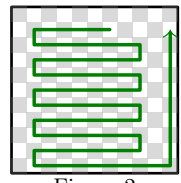


Figure 2

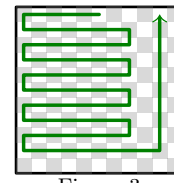


Figure 3

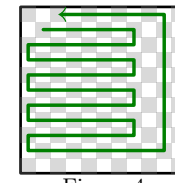


Figure 4

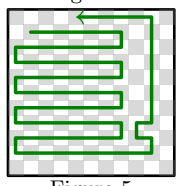


Figure 5

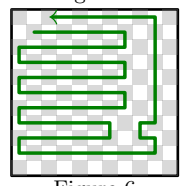


Figure 6

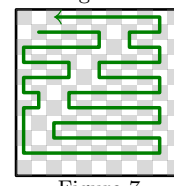


Figure 7

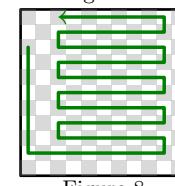


Figure 8

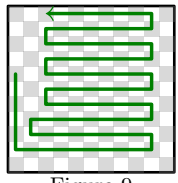


Figure 9

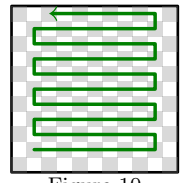


Figure 10

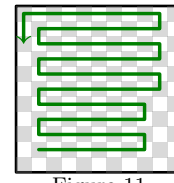


Figure 11

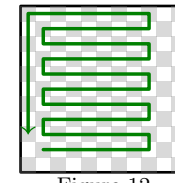


Figure 12

Use the obvious coordinate system with $(1,1)$ in the bottom left. Start with the snake as shown in Figure 1, then have it move to $(2,1)$, $(2,n)$, $(n,n-1)$ as in Figure 2. Then, have the snake shift to the position in Figure 3; this is possible since the snake can just walk to (n,n), then start walking to the left and then follow the route; by the time it reaches the ith row from the top its

tail will have vacated by then. Once it achieves Figure 3, move the head of the snake to $(3, n)$ to achieve Figure 4.

In Figure 5 and 6, the snake begins to "deform" its loop continuously. In general, this deformation by two squares is possible in the following way. The snake walks first to $(1, n)$ then retraces the steps left by its tail, except when it reaches $(n-1, 3)$ it makes a brief detour to $(n-2, 3)$, $(n-2, 4)$, $(n-1, 4)$ and continues along its way; this gives the position in Figure 5. Then it retraces the entire loop again, except that when it reaches $(n-4, 4)$ it turns directly down, and continues retracing its path; thus at the end of this second revolution, we arrive at Figure 6.

By repeatedly doing perturbations of two cells, we can move move all the "bumps" in the path gradually to protrude from the right; Figure 7 shows a partial application of the procedure, with the final state as shown in Figure 8.

In Figure 9, we stretch the bottom-most bump by two more cells; this shortens the "tail" by two units, which is fine. Doing this for all $(n-3)/2$ bumps arrives at the situation in Figure 10, with the snake's head at $(3, n)$. We then begin deforming the turns on the bottom-right by two steps each as in Figure 11, which visually will increase the length of the head. Doing this arrives finally at the situation in Figure 12. Thus the snake has turned around.

Second solution phrased using graph theory (Nikolai Beluhov) Let G be any undirected graph. Consider a snake of length k lying within G, with each segment of the snake occupying one vertex, consecutive segments occupying adjacent vertices, and no two segments occupying the same vertex. One move of the snake consists of the snake's head advancing to an adjacent empty vertex and segment i advancing to the vertex of segment $i-1$ for $i = 2, 3, \ldots, k$.

The solution proceeds in two stages. First we construct a planar graph G such that it is possible for a snake that occupies nearly all of G to turn around inside G. Then we construct a subgraph H of a grid adjacency graph such that H is isomorphic to G and H occupies nearly all of the grid.

For the first stage of the solution, we construct G as follows.

Let r and ℓ be positive integers. Start with r disjoint *main* paths $p_1, p_2, \ldots, p_r$, each of length at least ℓ, with p_i leading from A_i to B_i for $i = 1, 2, \ldots, r$. Add to those r *linking* paths, one leading from B_i to A_{i+1} for each $i = 1, 2, \ldots, r-1$, and one leading from B_r to A_1. Finally, add to those two families of *transit* paths, with one family containing one transit path joining A_1 to each of $A_2, A_3, \ldots, A_r$ and the other containing one path joining B_r to each of $B_1, B_2, \ldots, B_{r-1}$. We require that all paths specified in the construction have no interior vertices in common, with the exception of transit paths in the same family.

We claim that a snake of length $(r-1)\ell$ can turn around inside G.

To this end, let the concatenation $A_1B_1A_2B_2 \ldots A_rB_r$ of all main and linking paths be the *great cycle*. We refer to $A_1B_1A_2B_2 \ldots A_rB_r$ as the counterclockwise orientation of the great cycle, and to $B_rA_rB_{r-1}A_{r-1} \ldots B_1A_1$ as its

clockwise orientation.

Place the snake so that its tail is at A_1 and its body extends counterclockwise along the great cycle. Then let the snake manoeuvre as follows. (We track only the snake's head, as its movement uniquely determines the movement of the complete body of the snake.)

At phase 1, advance counterclockwise along the great cycle to B_{r-1}, take a detour along a transit path to B_r, and advance clockwise along the great cycle to A_r.

For $i = 2, 3, \ldots, r-1$, at phase i, take a detour along a transit path to A_1, advance counterclockwise along the great cycle to B_{r-i}, take a detour along a transit path to B_r, and advance clockwise along the great cycle to A_{r-i+1}.

At phase r, simply advance clockwise along the great cycle to A_1.

For the second stage of the solution, let n be a sufficiently large positive integer. Consider an $n \times n$ grid S. Number the columns of S from 1 to n from left to right, and its rows from 1 to n from bottom to top.

Let $a_1, a_2, \ldots, a_{r+1}$ be cells of S such that all of $a_1, a_2, \ldots, a_{r+1}$ lie in column 2, a_1 lies in row 2, a_{r+1} lies in row $n-1$, and $a_1, a_2, \ldots, a_{r+1}$ are approximately equally spaced. Let $b_1, b_2, \ldots, b_r$ be cells of S such that all of $b_1, b_2, \ldots, b_r$ lie in column $n-2$ and b_i lies in the row of a_{i+1} for $i = 1, 2, \ldots, r$.

Construct H as follows. For $i = 1, 2, \ldots, r$, let the main path from a_i to b_i fill up the rectangle bounded by the rows and columns of a_i and b_i nearly completely. Then every main path is of length approximately $\frac{1}{r}n^2$.

For $i = 1, 2, \ldots, r-1$, let the linking path that leads from b_i to a_{i+1} lie inside the row of b_i and a_{i+1} and let the linking path that leads from b_r to a_1 lie inside row n, column n, and row 1.

Lastly, let the union of the first family of transit paths be column 1 and let the union of the second family of transit paths be column $n-1$, with the exception of their bottommost and topmost squares.

As in the first stage of the solution, by this construction a snake of length k approximately equal to $\frac{r-1}{r}n^2$ can turn around inside an $n \times n$ grid S. When r is fixed and n tends to infinity, $\frac{k}{n^2}$ tends to $\frac{r-1}{r}$. Furthermore, when r tends to infinity, $\frac{r-1}{r}$ tends to 1. This gives the answer.

III

Number Theory

13 Orders

We will assume general comfort with modular arithmetic; readers without this background should first consult texts such as [Ste16].

§13.1 Definition and examples of order

Let p be a prime and take $a \not\equiv 0 \pmod{p}$.

Definition 13.1. The **order** of $a \pmod{p}$ is defined to be the smallest positive integer e such that

$$a^e \equiv 1 \pmod{p}.$$

This order is clearly finite because *Fermat's Little Theorem* tells us

$$a^{p-1} \equiv 1 \pmod{p}.$$

As an example, listed below are the orders of each residue $a \pmod{11}$ and $a \pmod{13}$.

a	mod 11	mod 13
1	1	1
2	10	12
3	5	3
4	5	6
5	5	4
6	10	12

a	mod 11	mod 13
7	10	12
8	10	4
9	5	3
10	2	6
11		12
12		2

The most important result about orders is that, due to minimality, they actually divide all other exponents: for example, they will all divide $p-1$, as the above examples suggest.

Theorem 13.2 (Fundamental theorem of orders)**.** *Suppose $a^N \equiv 1 \pmod{p}$. Then the order of $a \pmod{p}$ divides N.*

Proof. Let e be the order, and let $N = de + r$ by division algorithm (where $0 \le r < e$). Since $a^N \equiv 1 \pmod{p}$ and $a^e \equiv 1 \pmod{p}$, it follows $1 \equiv a^N (a^e)^{-d} \equiv a^r \pmod{p}$. This can only happen if $r = 0$ since e was minimal, which is the same as saying $e \mid N$. □

Although the prime case is the most frequently used, we note here that the order of $a \pmod{m}$ can be defined as long as $\gcd(a, m) = 1$; in that case it is the smallest exponent $e > 0$ such that $a^e \equiv 1 \pmod{m}$. In this situation, we have $e \mid \varphi(m)$.

§13.2 Application: Fermat's Christmas theorem

The upshot is that existence of elements with a certain order modulo p can usually tell you something about the prime p. Here is a classical application, for which orders are not part of the theorem statement but play an essential role.

Theorem 13.3. *Let a and b be positive integers, and let $p \equiv 3 \pmod 4$ be a prime. Suppose p divides $n = a^2 + b^2$. Then p divides both a and b (and in particular n is divisible by p^2).*

Proof. We proceed by contradiction: suppose p is an odd prime, and a, b are both nonzero modulo p, with

$$a^2 + b^2 \equiv 0 \pmod p$$

We will show that $p \equiv 1 \pmod 4$, which will be the desired contradiction.

Indeed, since b is nonzero we may invert it, and let $x = a/b \pmod p$. Then

$$\begin{aligned} x^2 + 1 &\equiv 0 \pmod p \\ x^2 &\equiv -1 \pmod p \\ x^4 &\equiv 1 \pmod p. \end{aligned}$$

Then, x has order four! Indeed the order should divide four, but $x^2 \equiv -1 \not\equiv 1 \pmod p$ since $p \neq 2$, so it must actually equal four.

In particular, $4 \mid p-1$, as desired. □

Remark 13.4. It turns out that a strong converse is true: if $p \equiv 1 \pmod 4$, then p can be written as a sum of squares. The most natural proofs of this result are not elementary so we will not prove it here, but it is good to know.

§13.3 Primitive roots

We now state an existence result, which we will not prove. The first is about orders modulo a prime being tight:

Theorem 13.5. *Let p be a prime. Then there exists an element $g \pmod p$ of order $p-1$.*

Definition 13.6. Such an element is known as a **primitive root** modulo p.

Here is a concrete example. It turns out that $g = 2$ is a primitive root

modulo both 11 and 13; let's write out what this means.

2^n	mod 11	mod 13
2^1	2	2
2^2	4	4
2^3	8	8
2^4	5	3
2^5	$\boxed{10}$	6
2^6	9	$\boxed{12}$
2^7	7	11
2^8	3	9
2^9	6	5
2^{10}	1	10
2^{11}		7
2^{12}		1

I've boxed the two "half-way" points: $2^5 \equiv 10 \equiv -1 \pmod{11}$ and $2^6 \equiv 12 \equiv -1 \pmod{13}$.

Consider $p = 11$. We already know that -1 cannot be a square modulo p from the proof of our earlier result, and you can intuitively see this come through: since $\frac{p-1}{2} = 5$ is odd, it's not possible to cut $g^5 \equiv -1$ into a perfect square.

On the other hand, if $p = 13$ then $p \equiv 1 \pmod 4$, and you can see intuitively why $g^6 \equiv -1$ is a perfect square: just write $g^6 = (g^3)^2$ and we're home free!

Based on the following discussion, we can for instance prove:

Example 136. Prove that if $p \equiv 1 \pmod 4$ is prime, then there exists an element x with $x^2 \equiv -1 \pmod p$.

Solution 136. All we need to do is generate an element of order 4. It is enough to pick $x = g^{\frac{1}{4}(p-1)}$ where g is some primitive root. ■

§13.4 Walkthroughs

Problem 137. Find all integers $n \geq 1$ such that n divides $2^n - 1$.

Walkthrough. The answer is $n = 1$ only. Assume for contradiction $n > 1$ works; consider a p dividing n.

(a) Show that $p \neq 2$.

(b) Show that if p is a prime dividing n, then the order of $2 \pmod p$ divides $\gcd(n, p-1)$.

(c) Prove that for any positive integer $n > 1$, there exists a prime $p \mid n$ with $\gcd(n, p-1) = 1$. (Hint: try several examples of n.)

(d) Conclude that the order of 2 (mod p) is 1, which produces the required contradiction.

Problem 138 (Ali Gurel)**.** Solve $a^{11} + 11b^{11} + 111c^{11} = 0$ over $\mathbb{Z}$.

Walkthrough. This is sort of the standard example showing how you're supposed to pick a modulus when given a generic large power.

(a) Prove that if $p \equiv 1 \pmod{11}$, then there are $\frac{p-1}{11} + 1$ possible eleventh powers modulo p, but otherwise, every number is an 11th power modulo p.

(b) Try taking modulo the smallest prime p that satisfies the hypothesis of (a).

(c) Take modulo that prime p. What can we conclude about a, b, c in that case?

(d) Show that $a = b = c = 0$ is the only solution.

Problem 139 (Online Math Open 2013)**.** Find the sum of all integers m with $1 \le m \le 300$ such that for any integer n with $n \ge 2$, if $2013m$ divides $n^n - 1$ then $2013m$ also divides $n - 1$.

Walkthrough. This is a really good test of how well you understand the notion of orders modulo a prime. It's one of my favorite instructional problems for this reason.

We'll say an integer $M > 1$ is *good* if whenever $n^n \equiv 1 \pmod{M}$ we also have $n \equiv 1 \pmod{M}$, and bad otherwise. The goal of this walkthrough will be to characterize all good integers.

(a) Which of $M \in \{2, 3, 4, 5\}$ are good?

(b) Prove that all odd M are bad.

(c) Prove that $M = 6$ and $M = 8$ are good.

(d) Show that $M = 10$ is good. (General idea: assume $n^n \equiv 1 \pmod{10}$. Prove that n is odd, $n^n \equiv 1 \pmod{5}$ and use this to deduce $n \equiv 1 \pmod{5}$).

So far it looks like even M are good. This luck won't hold:

(e) Check that $M = 12$ is good.

(f) Find an example of n such that $n^n \equiv 1 \pmod{14}$ but $n \not\equiv 1 \pmod{14}$. Thus $M = 14$ is *bad*.

(g) Prove that $M = 42$ is good, nonetheless.

(h) Show that $M = 30$ is good.

(i) Show that $M = 22$ is bad.

(j) Formulate a general conjecture about when an integer M is good. This doesn't require a new idea, just being able to piece together the general pattern from the specific cases you did earlier.

It remains to tackle the answer extraction.

(k) Show that if $2013m$ is good then $10 \mid m$.

(l) Among $m \in \{10, 20, \dots, 300\}$, only one of these leads to $2013m$ being bad. Which m is it?

§13.5 Problems

Problem 140. Let p be a prime. How many nonzero elements modulo p have order $p-1$ (i.e. are primitive roots)?

Problem 141 (HMMT February 2016)**.** For positive integers n, let c_n be the smallest positive integer for which $n^{c_n} - 1$ is divisible by 210, if such a positive integer exists, and $c_n = 0$ otherwise. What is $c_1 + c_2 + \cdots + c_{210}$?

Problem 142. Let p be a prime and n a positive integer. Determine the remainder when $1^n + 2^n + \cdots + (p-1)^n$ is divided by p, as a function of n and p.

Problem 143 (IMO Shortlist 2000)**.** Determine all positive integers $n \geq 2$ that satisfy the following condition: for a and b relatively prime to n we have $a \equiv b \pmod{n}$ if and only if $ab \equiv 1 \pmod{n}$.

Problem 144 (HMMT November 2014)**.** Determine all positive integers $1 \leq m \leq 50$ for which there exists an integer n for which m divides $n^{n+1} + 1$.

Problem 145 (China TST 2006)**.** Find all positive integers a and n for which n divides $(a+1)^n - a^n$.

Problem 146 (Don Zagier)**.** Let S denote the integers $n \geq 2$ with the property that for any positive integer a we have

$$a^{n+1} \equiv a \pmod{n}.$$

Show that S is finite and determine its elements.

Problem 147. Find all integers $n \geq 1$ such that n divides $2^{n-1} + 1$.

Problem 148 (Math Prize for Girls 2017)**.** Determine the value of the sum

$$\sum_{k=1}^{11} \frac{\sin\left(2^{k+4}\frac{\pi}{89}\right)}{\sin\left(2^k\frac{\pi}{89}\right)}.$$

§13.6 Solutions

Solution 137. The answer is $n = 1$ only, which obviously works.

If $n > 1$, consider the smallest prime p such that $p \mid n$. Note $p > 2$. Then we have the order of 2 modulo p divides $\gcd(p-1, n) = 1$. This is a contradiction.

Solution 138. The only solution is $(a, b, c) = (0, 0, 0)$, which obviously works.

Working modulo 23, we observe that $x^{11} \pmod{23}$ is either 0, 1, -1 for each x. One can check that the only combination for which $a^{11} + 11b^{11} + 111c^{11} \equiv 0 \pmod{23}$ is when $a^{11} \equiv b^{11} \equiv c^{11} \equiv 0 \pmod{23}$, so $23 \mid a, b, c$.

Consequently, whenever (a, b, c) is an integer solution, the triple $\left(\frac{a}{23}, \frac{b}{23}, \frac{c}{23}\right)$ is an integer solution too. Thus by infinite descent, the only solution is $(0, 0, 0)$.

Solution 139. Call an integer M *stable* if $n^n \equiv 1 \pmod{M}$ implies $n \equiv 1 \pmod{M}$.

Claim. The number M is stable if and only if for every prime $p \mid M$, all prime factors of $p - 1$ divide M.

Proof. We prove this only in the forwards direction (you can get the idea for the backwards direction by looking at the proof that 290 is stable). Suppose $n^n \equiv 1 \pmod{M}$. It suffices to show that $n \equiv 1 \pmod{p^k}$ for each $p^k \mid M$. Let $u \mid \varphi(p^k)$ be the order of n modulo p^k. Because $n^n \equiv 1 \pmod{M} \implies (n, M) = 1$, we find that $(n, \varphi(p^k)) = 1$. But $u \mid n$ as well. This forces $u = 1$, which is the desired. □

Let $M = 2013m$. First, we claim that M must be even. Otherwise, take $n = M - 1$. Then, we claim that 5 must divide M. Otherwise, take $n \equiv 0 \pmod{5}$, $n \equiv 3 \pmod{11}$, and $n \equiv 1$ modulo any other primes powers dividing M.

Now for $m = 10, 20, \ldots, 300$, it is easy to check by the condition that M is stable by our condition above, except for $m = 290$. It turns out that $m = 290$ is not stable; simply select $n \equiv 0 \pmod{7}$, $n \equiv 2^4 \pmod{29}$, and $n \equiv 1 \pmod{10 \cdot 2013}$. It is not hard to check that $n^n - 1 \equiv 1 \pmod{29 \cdot 10 \cdot 2013}$ and yet $n \not\equiv 1 \pmod{29}$, as desired.

So, the answer is $10 + 20 + 30 + \cdots + 280 + 300 = 4360$.

In fact, the converse to the stability lemma is true as well. We can generate the necessary counterexamples using primitive roots.

Solution 140. The answer is $\varphi(p-1)$. Indeed, let g be *one* particular primitive root. Then the nonzero elements modulo p are $1 = g^0$, g^1, …, g^{p-2}.

In general, the element g^k is the smallest exponent e such that $(g^k)^e \equiv 1 \pmod{p}$. Since g was primitive, this is the smallest integer e such that $ke \equiv 0 \pmod{p-1}$. That integer is $\frac{p-1}{\gcd(k,p-1)}$.

So in particular g^k is a primitive root if and only if $\gcd(k, p-1) = 1$ which occurs for $\varphi(p-1)$ values of k.

Solution 141. In order for $c_n \neq 0$, we must have $\gcd(n, 210) = 1$, so we need only consider such n. The number $n^{c_n} - 1$ is divisible by 210 iff it is divisible by each of 2, 3, 5, and 7, and we can consider the order of n modulo each modulus separately; c_n will simply be the LCM of these orders. We can ignore the modulus 2 because order is always 1. For the other moduli, the sets of orders are

$$\begin{aligned} a &\in \{1, 2\} && \bmod 3 \\ b &\in \{1, 2, 4, 4\} && \bmod 5 \\ c &\in \{1, 2, 3, 3, 6, 6\} && \bmod 7. \end{aligned}$$

By the Chinese Remainder Theorem, each triplet of choices from these three multisets occurs for exactly one n in the range $\{1, 2, \ldots, 210\}$, so the answer we seek is the sum of $\operatorname{lcm}(a, b, c)$ over a, b, c in the Cartesian product of these multisets. For $a = 1$ this table of LCMs is as follows:

	1	2	3	3	6	6
1	1	2	3	3	6	6
2	2	2	6	6	6	6
4	4	4	12	12	12	12
4	4	4	12	12	12	12

which has a sum of $21 + 56 + 28 + 56 = 161$. The table for $a = 2$ is identical except for the top row, where $1, 3, 3$ are replaced by $2, 6, 6$, and thus has a total sum of 7 more, or 168. So our answer is $161 + 168 = \boxed{\mathbf{329}}$.

This can also be computed by counting how many times each LCM occurs:

- 12 appears 16 times when $b = 4$ and $c \in \{3, 6\}$, for a contribution of $12 \times 16 = 192$;
- 6 appears 14 times, 8 times when $c = 6$ and $b \leq 2$ and 6 times when $c = 3$ and $(a, b) \in \{(1, 2), (2, 1), (2, 2)\}$, for a contribution of $6 \times 14 = 84$;
- 4 appears 8 times when $b = 4$ and $a, c \in \{1, 2\}$, for a contribution of $4 \times 8 = 32$;
- 3 appears 2 times when $c = 3$ and $a = b = 1$, for a contribution of $3 \times 2 = 6$;

- 2 appears 7 times when $a, b, c \in \{1, 2\}$ and $(a, b, c) \neq (1, 1, 1)$, for a contribution of $2 \times 7 = 14$;
- 1 appears 1 time when $a = b = c = 1$, for a contribution of $1 \times 1 = 1$.

The result is again $192 + 84 + 32 + 6 + 14 + 1 = 329$.

Solution 142. The answer is

$$1^n + \cdots + (p-1)^n \equiv \begin{cases} -1 & \text{if } p-1 \mid n \\ 0 & \text{otherwise.} \end{cases}$$

Indeed, the first case follows by Fermat's little theorem, since in that case every term is $1 \pmod p$ by Fermat's little theorem.

Now suppose $n \nmid p-1$, and let g be a primitive root modulo p. Then the above sum rewrites as

$$\begin{aligned} S = 1^n + \cdots + (p-1)^n &= (g^0)^n + (g^1)^n + \cdots + (g^{p-2})^n \\ &= g^0 + g^n + \cdots + g^{(p-2)n}. \end{aligned}$$

Using the "geometric series" formula, we get

$$\begin{aligned} (1-g^n)S &= (1-g^n)(g^0 + g^n + \cdots + g^{(p-2)n}) \\ &= 1 - g^{(p-1)n} = 1 - (g^{p-1})^n \\ &\equiv 1 - 1^n = 1 - 1 = 0 \pmod p. \end{aligned}$$

Since $g^n \not\equiv 1 \pmod p$ (as $p-1 \nmid n$) dividing gives $S \equiv 0 \pmod p$.

Solution 143. The answers are $n = 2^s \cdot 3^t$ for $s \in \{0, 1, 2, 3\}$, $t \in \{0, 1\}$ (and $n \neq 1$).

The problem is equivalent to $x^2 \equiv 1 \pmod n$ for all $\gcd(x, n) = 1$. By the Chinese remainder theorem, this is equivalent to the following property:

> For any prime power $q \mid n$, the maximal order modulo q is 2.

Now we note that:

- If $\gcd(q, 2) = 1$ we need $2^2 \equiv 1 \pmod q$ or $q \mid 3$.
- If $\gcd(q, 3) = 1$ we need $3^2 \equiv 1 \pmod q$ or $q \mid 8$.

Therefore, the only possible q are $q \in \{2, 3, 4, 8\}$ and one may check all of them work. This gives the answer above.

Remark. This is essentially asking when the `https://en.wikipedia.org/wiki/Carmichael_function` is equal to 2. The Carmichael function can be computed in this same way.

Solution 144. The answers are odd m, as well as

$$m \in \{2, 2 \cdot 5, 2 \cdot 13, 2 \cdot 17, 2 \cdot 5^2\} = \{2, 10, 26, 34, 50\}.$$

First, we show all m must be as above. Assume m is even; then n is odd. In that case $n^{n+1} + 1 \equiv 2 \pmod 4$ and moreover is a sum of squares. Thus by Fermat's Christmas theorem m must be equal to $2p_1 \dots p_k$ where $p_i \equiv 1 \pmod 4$ are not necessarily distinct primes.

It remains to do the construction. For odd m, we simply select $n = m - 1$. For the five special even values m, one can use the Chinese remainder theorem to generate constructions: just require $n \equiv 1 \pmod 4$ and $n^2 + 1 \equiv 0 \pmod{m/2}$. We give them explicitly here for concreteness:

- $m = 2$: use $n = 1$.
- $m = 10$: use any $n \equiv 1 \pmod 4$, $n \equiv 2 \pmod 5$.
- $m = 26$: use any $n \equiv 1 \pmod 4$, $n \equiv 5 \pmod{13}$.
- $m = 34$: use any $n \equiv 1 \pmod 4$, $n \equiv 4 \pmod{17}$.
- $m = 50$: use any $n \equiv 1 \pmod 4$, $n \equiv 7 \pmod{50}$.

Solution 145. The answer is that (a, n) works if and only if $n = 1$. When $n = 1$ there is nothing to prove.

Assume for contradiction $n > 1$. Look at the smallest prime p dividing n; note that it does not divide a. Then

$$\left(1 + \frac{1}{a}\right)^n \equiv 1 \pmod p.$$

Let $x = 1 + 1/a$ and let e be the *order* of x modulo p (meaning e is the smallest positive integer with $x^e \equiv 1 \pmod p$). Then $e \mid n$, but also $e \mid p - 1$ by Fermat's little theorem. So $e \mid \gcd(n, p-1)$. Since p was the smallest prime dividing n, we have $e = 1$. Thus $x \equiv 1 \pmod p$, which is absurd.

Solution 146. The answer is $n = 2$, $n = 6$, $n = 42$, $n = 1806$.

Clearly n must be squarefree (else take a a prime dividing n more than once). So let us focus our attention on n of the form

$$n = p_1 p_2 \dots p_k \qquad p_1 < p_2 < \dots < p_k$$

with p_i prime.

By the Chinese remainder theorem, n works if and only if for every i, the statement

$$a^n \equiv 1 \pmod{p_i} \qquad \forall \gcd(a, p_i) = 1$$

holds true. In particular, if we take a to be a primitive root modulo p_i then this is equivalent to

$$p_i - 1 \mid n = p_1 \dots p_k$$

for every i.

This last relation is enough to solve the problem now, since it means

$$p_1 - 1 \mid 1, \quad p_2 - 1 \mid p_1, \quad p_3 - 1 \mid p_1 p_2, \quad p_4 - 1 \mid p_1 p_2 p_3$$

and so on. A computation now gives that $p_1 = 2$, $p_2 = 3$ if $k \ge 2$, $p_3 = 7$ if $k \ge 3$, $p_4 = 43$ if $k \ge 4$, but no possibilities for $k \ge 5$.

Solution 147. The answer is $n = 1$ only which obviously works.

Clearly n is odd. Then:

- For every prime $p \mid 2^{n-1} + 1$, we must have $p \equiv 1 \pmod 4$ since the right-hand side is a sum of squares (equivalently, $2^{\frac{n-1}{2}}$ has order 4, so $4 \mid p - 1$).

  In particular, $n \equiv 1 \pmod 4$ as well, being the product of 1 (mod 4) primes.

- But now we claim every prime dividing $2^{n-1} + 1$ is 1 (mod 8), by the same reason! Indeed, the right-hand side is a sum of fourth powers; equivalently, $2^{\frac{n-1}{4}}$ has order 8, forcing $8 \mid p - 1$.

  In particular, $n \equiv 1 \pmod 8$ as well, being the product of 1 (mod 8) primes.

- But then repeating the same argument shows that all prime divisors of $2^{n-1} + 1$ are $1 \pmod{16}$, and so on…

Repeating the same logic we find $n \equiv 1 \pmod{2^k}$ for any positive integer k. Thus $n = 1$.

Solution 148. The solution is divided into three parts.

Part I: complex numbers. We begin (as I always do) by rewriting all the trigonometry in terms of roots of unity. For brevity, let

$$\zeta = \exp\left(\frac{2\pi i}{89}\right) \qquad \text{and} \qquad z_k = \zeta^{2^k}$$

Then $\sin\left(2^k \frac{\pi}{89}\right) = \frac{1}{2i}(z_{k-1} - \overline{z_{k-1}})$, and hence the sum rewrites (shifting indices) as

$$\begin{aligned} S &= \sum_{k=0}^{10} \frac{z_k^{16} - z_k^{-16}}{z_k - z_k^{-1}} \\ &= \sum_{k=0}^{10} \underbrace{\left(z_k^{15} + z_k^{13} + z_k^{11} + \cdots + z_k^{-15}\right)}_{16 \text{ terms}}. \end{aligned}$$

Part II: guessing the answer. At this point, we can take all the exponents in the expression S modulo 89. None of the exponents are zero.

To summarize: we have written S as a linear combination of $16 \cdot 11 = 176$ powers of ζ, each in the set $\{\zeta^1, \zeta^2, \ldots, \zeta^{88}\}$.

However: the minimal polynomial P of ζ is the 89th cyclotomic polynomial:

$$P(\zeta) = 1 + \zeta + \zeta^2 + \cdots + \zeta^{88} = 0.$$

Now P has 89 terms and degree 88. If we view S as a polynomial in ζ and then take the remainder modulo P, we'll get a polynomial of degree at most 87, which (due to the minimality of P) That suggests $\boxed{S = -2}$, as $S + 2$ has $11 \cdot 16 + 2 = 178$ terms, which is exactly the right number of terms in two copies of P. In other words, our claim is that $S(\zeta) + 2 = 2P(\zeta)$ as polynomials in ζ.

Let's see what it would take to check this is correct. The exponents of ζ that appear in the sum S are exactly the numbers of the form

$$\pm 2^k m \pmod{89} \qquad \text{where } 0 \le k \le 10 \text{ and } m \in \{1, 3, 5, \ldots 15\}.$$

This is a total of 176 numbers, and we aim to show

Claim. Every nonzero residue modulo 89 appears exactly twice among numbers in the above form.

Part III: the grind. We now turn our attention to proving the claim at the end of Part II. Note that 2 (mod 89) has order 11, since $2^{11} - 1 = 2047 = 89 \cdot 23$. (Actually, that part is guessable from the fact the index k in the original definition of S runs to 11.)

Let $G = (\mathbb{Z}/89)^\times$ denote the nonzero residues modulo 89. Let $H \subset G$ denote the 11th powers modulo 89, hence $|H| = 8$. Then the map

$$\psi : G \twoheadrightarrow H \qquad \text{by} \qquad x \mapsto x^{11}$$

is surjective, and vanishes exactly at powers of 2. Thus to prove the claim, it suffices to show the pre-image of every element of H contains exactly two elements from the set $\{\pm 1, \pm 3, \pm 5, \ldots, \pm 15\}$.

So, we compute a bunch of 11th powers (whee):

$$\begin{aligned}(\pm1)^{11} &= \pm1\\ (\pm3)^{11} &= \pm37\\ (\pm5)^{11} &= \pm55\\ (\pm7)^{11} &= \pm37\\ (\pm9)^{11} &= \pm34\\ (\pm11)^{11} &= \pm88\\ (\pm13)^{11} &= \pm77\\ (\pm15)^{11} &= \pm77.\end{aligned}$$

From here we see $H = \{1, 22, 34, 37, 52, 55, 77, 88\}$ and indeed that each element appears exactly twice.

Remark. The paragraph with G and H can be rewritten in more elementary terms. Fix a primitive root g for which $g^8 = 2$. Then for each element $m = g^k$ we only care about $k \pmod 8$. We can read this off from $m^{11} = g^{11k} \pmod{89}$.

---

14 Look at the exponent

§14.1 Definition

For a prime p and nonzero integer n, we let $\nu_p(n)$ denote the largest integer e with p^e dividing n. We can extend this valuation to rational numbers by $\nu_p(a/b) = \nu_p(a) - \nu_p(b)$. Thus we get a function

$$\nu_p \colon \mathbb{Q}_{\neq 0} \to \mathbb{Z}.$$

By convention we then set $\nu_p(0) = \infty$.

Then:

- $\nu_p(xy) = \nu_p(x) + \nu_p(y)$.
- $\nu_p(x+y) \geq \min\{\nu_p(x), \nu_p(y)\}$, and equality holds if $\nu_p(x) \neq \nu_p(y)$.

Note also that for integers x and y we have

$$\nu_p(x-y) \geq e \iff x \equiv y \pmod{p^e}$$

and we will use these perspectives interchangeably.

The idea of this lecture is to see just how far we can go using just this relation. For example, you might try to do Problem 149 now, to see an application using "bare hands".

§14.2 Exponent lifting

I am about to state the exponent lifting lemma, which is a useful tool for evaluating $\nu_p(a^n \pm b^n)$. However, if you have never seen this lemma before, you are required to do Problem 150 before continuing, to make sure you actually internalize the statement of the lemma.

Okay, here is the lemma.

Theorem 14.1 (Exponent lifting lemma)**.** *Let a and b be integers and p a prime. Assume that*

(i) *$a - b$ is divisible by p,*

(ii) *$p \nmid a, b$, and*

(iii) *$p \neq 2$.*

Then

$$\nu_p(a^n - b^n) = \underbrace{\nu_p(a-b)}_{>0} + \nu_p(n).$$

I want to stress that if you choose to apply this lemma, then you really must *remember to verify all three conditions*! It is worse to apply the lemma blindly and forget a hypothesis (rather than just pretend you did not know it and work out the special case that you need), because this will lead to plausible-looking but wrong results.

I highlighted the fact that $\nu_p(a-b) > 0$ in the conclusion: if you apply the lemma and find that term is zero, you messed up (probably forgot the first condition).

§14.3 Walkthroughs

Problem 149. Let a and b be positive integers such that $a \mid b^2$, $b^2 \mid a^3$, $a^3 \mid b^4$, $b^4 \mid a^5$, $a^5 \mid b^6$, and so on. Show that $a = b$.

Walkthrough. This is actually a quite easy problem, but it showcases the idea well.

(a) Show that
$$(2n-1)\nu_p(a) \le 2n\nu_p(b)$$
for any positive integer n and prime p.

(b) What happens as we take n large?

(c) Conclude $\nu_p(a) \le \nu_p(b)$ for every prime p.

(d) Show $b \mid a$ similarly.

(e) Conclude.

Problem 150. For each positive integer n, compute
$$\nu_3\left(2^{3^n}+1\right).$$

Walkthrough. This is an induction, using the fact that each step we get a sum of cubes. First, here are some base cases:

$$\begin{aligned}
\nu_3(2^1+1) &= \nu_3(3) = 1 \\
\nu_3(2^3+1) &= \nu_3(9) = \nu_3(3^2) = 2 \\
\nu_3(2^9+1) &= \nu_3(513) = \nu_3(3^3 \cdot 19) = 3 \\
\nu_3(2^{27}+1) &= \nu_3(134217729) = \nu_3(3^4 \cdot 19 \cdot 87211) = 4 \\
&\vdots
\end{aligned}$$

So you can probably guess the answer!

(a) Show that
$$\frac{2^{3^{n+1}}+1}{2^{3^n}+1}$$
is divisible by 3.

(b) Show that it is not divisible by 9.

(c) Use induction to figure out what the answer is now.

You can think of this without induction by writing
$$2^{3^n}+1 = (2+1)(2^2-2+1)(8^2-8+1)(512^2-512+1)\dots$$
and then showing that $\nu_3(A^2-A+1)=1$.

Problem 151 (AIME 2018)**.** Find the smallest positive integer n such that 3^n ends with 01 when written in base 143.

Walkthrough. The idea is that to use the order of 3 (mod p) to get the order of 3 (mod p^e) for $e \geq 1$. Naturally, this is exponent lifting lemma.

(a) Find all n for which $3^n \equiv 1 \pmod{11}$.

(b) Show that for all n in (a), we also have $3^n \equiv 1 \pmod{11^2}$.

(c) Show that $3^n \equiv 1 \pmod{13}$ iff $3 \mid n$.

(d) Prove that $3^n \equiv 1 \pmod{13^2}$ iff $3 \cdot 13 \mid n$. You'll probably want to use exponent lifting.

(e) Why do parts (b) and (d) behave differently? Describe what would happen for a general prime $p \neq 3$ instead of 11 or 13.

(f) Extract the final answer by combining (b) and (d) together.

(g) Follow-up: for any $r \geq 2$, what is the smallest n such that $3^n \equiv 1 \pmod{143^r}$?

In general, if $p > 3$ is a prime, we might be inclined to try and say something about $\nu_p(3^e-1)$ where e is the order of 3 (mod p). A prime p for which this quantity is greater than 1 is called an 3-Wieferich prime. Very little is known about Wieferich primes in general.

Problem 152 (Asian-Pacific Olympiad 2017)**.** Let a, b, c be positive rational numbers with $abc = 1$. Suppose there exist positive integers x, y, z for which $a^x + b^y + c^z$ is an integer. Prove that when a, b, c are written as a fractions in lowest terms, the numerators are perfect powers.

Walkthrough. Pick any prime p and look at ν_p's.

(a) Note $\nu_p(a) + \nu_p(b) + \nu_p(c) = 0$.

(b) Show that it's impossible to have $\nu_p(a) \geq 0$, $\nu_p(b) \geq 0$, and $\nu_p(c) < 0$. (Some readers prefer to think of this as a corollary of (d), so if you want you can skip there directly.)

(c) Conclude that if $\nu_p(a) > 0$, then $\nu_p(b) < 0$ and $\nu_p(c) < 0$. We'll assume this in what follows, and try to show $\nu_p(a)$ is divisible by some fixed number.

(d) Prove that in that case we must have $\nu_p(b^y) = \nu_p(c^z)$.

(e) As an example, if $(y, z) = (5, 6)$ (so that $\nu_p(b) : \nu_p(c) = 6 : 5$), what must $\nu_p(a)$ be divisible by?

(f) Show that $\nu_p(a)$ is divisible by $\frac{y+z}{\gcd(y,z)}$.

(g) Conclude.

Problem 153 (USA TST 2008)**.** Prove that $n^7 + 7$ is not a perfect square for any integer n.

Walkthrough. We begin by taking mods, to get some preliminary facts.

(a) Resolve the edge cases where $n \leq 0$, so we can assume $n \geq 1$ in what follows.

(b) Show that $n \equiv 1 \pmod 4$.

Now, consider the equation $n^7 + 7 = a^2$, and assume for contradiction it is satisfied by some $n > 0$.

(c) Add a certain three-digit positive integer to both sides that gives you something to work with. (You'll know when you have the right constant.)

(d) Prove that $n + 2$ must be divisible by 11. (Possible hint: Fermat's Christmas theorem.)

(e) Show that $\nu_{11}(n^7 + 2^7) = \nu_{11}(n + 2)$.

(f) Consider the exponent of 11 carefully to get a contradiction.

§14.4 Problems

Problem 154. For which primes p is $(p-1)^p + 1$ a power of p?

Problem 155 (IMO Shortlist 1991)**.** Find the largest integer k for which 1991^k divides

$$1990^{1991^{1992}} + 1992^{1991^{1990}}.$$

Problem 156 (Putnam 2003 B3)**.** Prove that for any positive integer n, we have

$$\prod_{k=1}^{n} \operatorname{lcm}\left(1, 2, \ldots, \left\lfloor \frac{n}{k} \right\rfloor\right) = n!.$$

Problem 157 (Bay Area Olympiad 2018)**.** Let a, b, c be positive integers. Show that if $a/b + b/c + c/a$ is an integer, then $\sqrt[3]{abc}$ is an integer as well.

Problem 158 (USAMO 2016)**.** Prove that for any positive integer k,

$$(k^2)! \cdot \prod_{j=0}^{k-1} \frac{j!}{(j+k)!}$$

is an integer.

Problem 159 (IMO Shortlist 2017)**.** A rational number is *short* if its decimal representation has finitely many nonzero digits. A triple (t, c, m) of positive integers is *fantastic* if $c \le 2017$ and $\frac{10^t - 1}{cm}$ is short, but $\frac{10^i - 1}{cm}$ is not short for $1 \le i < t$.

For each positive integer m, let

$$S(m) = \{t \in \mathbb{Z}_{>0} \mid (t, c, m) \text{ fantastic for some } c\}.$$

Find the largest possible value of $|S(m)|$ over all m.

Problem 160 (IMO 1990)**.** Find all positive integers n for which n^2 divides $2^n + 1$.

Problem 161 (IMO Shortlist 2014)**.** Find all primes p and positive integers (x, y) such that $x^{p-1} + y$ and $y^{p-1} + x$ are powers of p.

§14.5 Solutions

Solution 149. Let p be any prime. Then $\nu_p(a) \le 2\nu_p(b)$, $3\nu_p(a) \le 4\nu_p(b)$, etc and in general

$$\nu_p(a) \le \frac{2n}{2n-1}\nu_p(b) = \left(1 + \frac{1}{2n-1}\right)\nu_p(b)$$

for any positive integer n. Since this should hold for any positive integer n, by taking $n \to \infty$ we find $\nu_p(a) \le \nu_p(b)$.

Since this holds for any prime p, we have $a \mid b$. A similar argument shows $b \mid a$.

Solution 150. The answer is $\nu_3(2^{3^n}+1) = n+1$. We will prove this by induction on $n \ge 0$.

When $n = 0$, we have $2^{3^0}+1 = 3$ which indeed has 3^1.

For the inductive step, it suffices to check that

$$\frac{2^{3^{n+1}}+1}{2^{3^n}+1}$$

is divisible by 3 but not 9. Using the fact it is a sum of cubes, it equals $A^2 - A + 1$ where $A = 2^{3^n}$.

Since $A \equiv -1 \pmod 9$ for every positive integer n (again by induction on n), we get

$$A^2 - A + 1 \equiv 1 - (-1) + 1 \equiv 3 \pmod 9$$

which proves the result.

Remark. You can think of this without induction by writing

$$2^{3^n}+1 = (2+1)(2^2-2+1)(8^2-8+1)(512^2-512+1)\dots$$

and then showing that $\nu_3(A^2 - A + 1) = 1$, as above.

Solution 151. The answer is $3 \cdot 5 \cdot 13 = 195$.

First note $3^n \equiv 1 \pmod{11} \iff 5 \mid n$ and $3^n \equiv 1 \pmod{13} \iff 3 \mid n$. Now,

$$\nu_{11}(3^n - 1) = \nu_{11}(243^{n/5} - 1) = 2 + \nu_{11}(n/5)$$
$$\nu_{13}(3^n - 1) = \nu_{13}(27^{n/3} - 1) = 1 + \nu_{13}(n/3).$$

The answer follows from this.

In general, the smallest n such that $3^n \equiv 1 \pmod{143^r}$ is $11^{r-2} \cdot 13^{r-1} \cdot 15$ for $r \ge 2$.

Remark. In general, if $p > 3$ is a prime, we might be inclined to try and say something about $\nu_p(3^e - 1)$ where e is the order of 3 (mod p). A prime p for which this quantity is greater than 1 is called an 3-Wieferich prime. Very little is known about Wieferich primes in general.

Solution 152. It is sufficient to prove the following claim.

Claim. Let p be a prime. If $\nu_p(a) > 0$ then $\nu_p(a)$ is divisible by $\frac{y+z}{\gcd(y,z)}$.

Proof. Note $\nu_p(a) + \nu_p(b) + \nu_p(c) = 0$. WLOG assume $\nu_p(c) < 0$ (hence $\nu_p(c^z) < 0$). Then since

$$\nu_p(a^x + b^y + c^z) \geq 0$$

and $\nu_p(a^x) > 0$, we must have

$$\nu_p(b^y) = \nu_p(c^z) \implies y\nu_p(b) = z\nu_p(c).$$

Thus we may set $\nu_p(b) = -z'k$ and $\nu_p(c) = -y'k$ where $y' = \frac{y}{\gcd(y,z)}$ and $z' = \frac{z}{\gcd(y,z)}$. Then

$$\nu_p(a) = -\nu_p(b) - \nu_p(c) = k \cdot (y' + z')$$

as needed. □

Therefore the numerator of a is a perfect $\frac{y+z}{\gcd(y,z)}$th power.

Solution 153. Assume $n > 0$ since $n = 0$ and $n = -1$ are easy and for $n \leq -2$ the expression is negative.

Suppose $n^7 + 7 = a^2$. Then

$$n^7 + 2^7 = a^2 + 11^2.$$

Taking modulo 4 gives $n \equiv 1 \pmod 4$, but $n + 2 \mid a^2 + 11^2$, and $n + 2 \equiv 3 \pmod 4$.

Note that $a^2 + 11^2$ has no 3 mod 4 prime factors except possibly an 11^2, by Fermat's Christmas theorem. Since $n + 2 \equiv 3 \pmod 4$ we would need to have $\nu_{11}(n+2) = 1$ as a result, since $\nu_{11}(n+2)$ should be odd and at most 2. However, we then get

$$\begin{aligned} \nu_{11}\left(a^2 + 11^2\right) &= \nu_{11}\left(n^7 + 2^7\right) \\ &= \nu_{11}(n+2) + \nu_{11}(7) = 1 + 0 = 1 \end{aligned}$$

by the exponent lifting lemma, which is impossible.

Solution 154. The answer is $p = 2$ or $p = 3$ only, which are checked to work.

For $p > 3$ many solutions are possible:

- By Zsigmondy theorem, $(p-1)^p+1$ cannot be a power of p.
- Catalan conjecture implies the problem.
- Lifting the exponent implies $(p-1)^p+1=p^2$.
- Taking modulo p^3 works too: we have

$$(p-1)^p+1=-\binom{p}{1}\cdot p+\binom{p}{2}\cdot p^2\equiv -p^2 \pmod{p^3}$$

and so $(p-1)^p+1$ must be either p or p^2, which ensures $p<5$.

Solution 155. Let us write the expression as

$$N=A^{1991^{1990}}+B^{1991^{1990}}$$

where $A=1990^{1991^2}$ and $B=1992$. Observe that

$$\begin{aligned} A+B &= (1991-1)^{1991^2}+1992 \\ &\equiv 1991\binom{1991^2}{1}+(-1)+1992 \pmod{1991^2} \\ &\equiv 1991 \pmod{1991^2}. \end{aligned}$$

Factor $1991=11\cdot 181$, so let $p\in\{11,181\}$ be an odd prime. We then have $\nu_p(A+B)=1$. Since this is positive, $p\nmid AB$, and $p>2$, the exponent lifting lemma implies

$$\nu_p(N)=\nu_p(A+B)+\nu_p\left(1991^{1990}\right)=1990+1=1991.$$

Since this holds both for $p=11$ and $p=181$, we conclude the answer $k=1991$.

Solution 156. Actually, we'll prove the stronger claim that for any prime power q, the number of terms on each side divisible by q is exactly the same. By the fundamental theorem of arithmetic, that will imply the desired equality.

Clearly, the number of terms on the right-hand side which are divisible by q is $\lfloor n/q\rfloor$.

As for the left-hand side, the number of lcm's on the right which are divisible by q is given by the number of k for which $\left\lfloor \frac{n}{k}\right\rfloor \geq q$, which is exactly $k=1,\dots,\lfloor n/q\rfloor$.

Solution 157. Fix any prime $p\mid abc$ and let $x=\nu_p(a)$,s $y=\nu_p(b)$, $z=\nu_p(c)$. (Thus $x+y+z>0$.) It is enough to prove $3\mid x+y+z$.

If $x = y = z$ we are done, so assume not. Then $\nu_p(a/b) = x - y$, $\nu_p(b/c) = y - z$, $\nu_p(c/a) = z - x$. At least one of these numbers is negative. Thus from $\nu_p(a/b + b/c + c/a) \geq 0$ we conclude that the two smallest numbers among $\{x - y, y - z, z - x\}$ must be equal.

But if $x - y = y - z$, say, then $2y = x + z$ and so $x + y + z = 3y \equiv 0 \pmod 3$. Similarly for the other two cases.

Solution 158. We show the exponent of any given prime p is nonnegative in the expression. Recall that the exponent of p in $n!$ is equal to $\sum_{i \geq 1} \lfloor n/p^i \rfloor$. In light of this, it suffices to show that for any prime power q, we have

$$\left\lfloor \frac{k^2}{q} \right\rfloor + \sum_{j=0}^{k-1} \left\lfloor \frac{j}{q} \right\rfloor \geq \sum_{j=0}^{k-1} \left\lfloor \frac{j+k}{q} \right\rfloor$$

Since both sides are integers, we show

$$\left\lfloor \frac{k^2}{q} \right\rfloor + \sum_{j=0}^{k-1} \left\lfloor \frac{j}{q} \right\rfloor > -1 + \sum_{j=0}^{k-1} \left\lfloor \frac{j+k}{q} \right\rfloor.$$

If we denote by $\{x\}$ the fractional part of x, then $\lfloor x \rfloor = x - \{x\}$ so it's equivalent to

$$\left\{ \frac{k^2}{q} \right\} + \sum_{j=0}^{k-1} \left\{ \frac{j}{q} \right\} < 1 + \sum_{j=0}^{k-1} \left\{ \frac{j+k}{q} \right\}.$$

However, the sum of remainders when $0, 1, \ldots, k-1$ are taken modulo q is easily seen to be less than the sum of remainders when $k, k+1, \ldots, 2k-1$ are taken modulo q. So

$$\sum_{j=0}^{k-1} \left\{ \frac{j}{q} \right\} \leq \sum_{j=0}^{k-1} \left\{ \frac{j+k}{q} \right\}$$

follows, and we are done upon noting $\left\{k^2/q\right\} < 1$.

Solution 159. The answer is 807.

We restrict our attention to c and m such that $\gcd(c, 10) = \gcd(m, 10) = 1$, since stripping factors of 2 or 5 doesn't change anything. In that case, since t is determined by c and m in a fantastic triple (the order of $10 \pmod{cm}$), we have

$$\begin{aligned} \#S(m) &\leq \#\{1 \leq c \leq 2017 \mid \gcd(c, 10) = 1\} \\ &= 2017 - \left\lfloor \frac{2017}{2} \right\rfloor - \left\lfloor \frac{2017}{5} \right\rfloor + \left\lfloor \frac{2017}{10} \right\rfloor \\ &= 807. \end{aligned}$$

The main point of the problem is to achieve this bound.

Let T be a large composite integer such that $M = 10^T - 1$ is divisible by every prime at most 2017 other than 2 and 5. (Thus T is the order of 10 $\pmod M$.)

Claim. The order of 10 $\pmod{cM}$ is cT, for $1 \le c \le 2017$ with $\gcd(c, 10) = 1$.

Proof. This essentially follows by exponent lifting lemma. Indeed, the order of 10 $\pmod{cM}$ must be divisible by T. Now pick a prime $p \mid c$. If T' is the order of 10 $\pmod{cM}$, then T' must be divisible by T; now compute

$$\begin{aligned} \nu_p(c) + \nu_p(M) &\le \nu_p(10^{T'} - 1) \\ &= \nu_p\left((10^T)^{T'/T} - 1\right) \\ &= \nu_p(10^T - 1) + \nu_p(T') - \nu_p(T) \\ \iff \nu_p(T') &\ge \nu_p(cT). \end{aligned}$$

This completes the proof. □

Hence, the relevant fantastic triples are (cT, c, M) for each $c \le 2017$ relatively prime to 10.

Solution 160. Answer: $n = 1$ or $n = 3$, which clearly work. So we prove they are the only ones.

Assume now $n > 1$, and let $p \mid n$ be a minimal prime. Note that $p \neq 2$. As $2^{2n} \equiv 1 \pmod p$ and $2^{p-1} \equiv 1 \pmod p$ we must have

$$p \mid 2^{\gcd(2n, p-1)} - 1 \mid 2^2 - 1$$

and so $p = 3$.

Now, by lifting the exponent,

$$2\nu_3(n) = \nu_3(n^2) \le \nu_3(2^n + 1) = \nu_3(2+1) + \nu_3(n) = 1 + \nu_3(n) \implies \nu_3(n) \le 1.$$

Now assume for contradiction $n > 3$, and let $q \mid n/3$ be a minimal prime. We know $q \notin \{2, 3\}$, and yet

$$q \mid 2^{\gcd(2n, q-1)} - 1 \mid 2^6 - 1 = 63$$

which would require $q = 7$, but $2^n + 1 \not\equiv 0 \pmod 7$ for any n, contradiction.

Solution 161. If $p = 2$ then any (x, y) with $x + y$ a power of two is okay. We claim the only other answer is $(x, y, p) = (5, 2, 3)$ and $(x, y, p) = (2, 5, 3)$.

Henceforth assume $p > 2$. Then if $\nu_p(x) \ge \nu_p(y) > 0$ we get an immediate contradiction, thus we may assume $p \nmid x, y$ (ergo $\gcd(x, y) = 1$). So by Fermat's Little Theorem, x and y are $-1 \pmod p$.

It is easy to check that when $p > 2$ we cannot have $x = y$, since otherwise $x(x^{p-2}+1)$ is a power of p, which is clearly impossible when $p > 2$. Moreover, if $p > 2$ then $x^{p-1}+y \neq y^{p-1}+x$, since otherwise $(x-y)(x^{p-2}+\cdots+y^{p-2}) = x-y$, which is impossible unless $x = y$.

Thus, suppose $y^{p-1} + x < x^{p-1} + y$, which is equivalent to $y < x$. Then in particular $y^{p-1} + x$ divides $x^{p-1} + y$, so

$$y^{p-1} + x \mid (-y^{p-1})^p + y \implies y^{p-1} + x \mid y^{p(p-2)} + 1.$$

By Lifting the Exponent, we thus deduce that

$$\nu_p(y^{p-1} + x) \leq 1 + \nu_p(y + 1).$$

Actually, since LHS is a power of p, this informs us that

$$y^{p-1} + x \mid p \cdot (y + 1) \implies y^{p-1} + x \leq p \cdot (y + 1).$$

Since $x > y$, this forces

$$y^{p-1} + y \leq p \cdot (y + 1).$$

Also, $y \geq p-1$ since $y \equiv -1 \pmod{p}$. This can only occur if $y = 2$ and $p = 3$.

Now, $y^{p-1} + x \mid p \cdot (y + 1) \implies 4 + x \mid 9$, hence $x = 5$, yielding the solution set.

---

15 Advanced techniques

In this chapter we discuss three more advanced bits of theory.

§15.1 Pell equations

The theory of Pell's equation

$$a^2 - nb^2 = 1$$

involves some algebraic number theory to motivate properly. We will not discuss this in detail here, but merely mention (which is enough for our purposes) how to *generate* solutions. A more comprehensive treatment can be found in the bonus chapter on Pell's equation in [Che19].

Definition 15.1. Let n be a positive integer. Given $\alpha \in \mathbb{Q}(\sqrt{n})$ we define

$$\|\alpha\| = \left\|a + b\sqrt{n}\right\| = a^2 - nb^2.$$

Theorem 15.2. *This norm is multiplicative.*

Proof. Check it directly:

$$\left(a^2 - nb^2\right)\left(c^2 - nd^2\right) = (ac + nbd)^2 - n\,(ad + bc)^2\,. \qquad \square$$

Let us see an example of how this can be used. Suppose we want to generate solutions to $x^2 - 2y^2 = 1$. We start by observing that $(3, 2)$ is a solution; this is the same as saying $3 + 2\sqrt{2}$ has norm 1. Then we can consider

$$\left(3 + 2\sqrt{2}\right)^2 = 17 + 12\sqrt{2}$$

which will also have norm $1^2 = 1$; and indeed $(17, 12)$ is a solution too. Going further,

$$\left(3 + 2\sqrt{2}\right)^3 = 99 + 70\sqrt{2}$$

reveals the solution $(99, 70)$.

It is a theorem (which we will not prove) that in fact $(3+2\sqrt{2})^n$ will generate all solutions. More generally:

Theorem 15.3 (Pell equations generated by a unit)**.** *Let n be a positive squarefree integer and consider the Pell equation $x^2 - ny^2 = 1$.*

Then there exists a pair (x_1, y_1) of positive integers satisfying $x_1^2 - ny_1^2 = 1$ and such that all other solutions (x, y) are obtained by writing

$$x + y\sqrt{n} = \left(x_1 + y_1\sqrt{n}\right)^k$$

for some positive integer k.

§15.2 Jacobi symbol and quadratic reciprocity

The quadratic reciprocity formula specifies how to check if $a \pmod p$ is a quadratic residue.

Definition 15.4. For a prime p and integer a, set

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & p \mid a \\ 1 & a \not\equiv 0 \text{ is a quadratic residue} \\ -1 & a \not\equiv 0 \text{ is not a quadratic residue.} \end{cases}$$

This is called a **Legendre symbol**.

Proposition 15.5 (Legendre's definition)**.** *For odd primes p,*

$$\left(\frac{a}{p}\right) \equiv a^{\frac{1}{2}(p-1)} \pmod p.$$

It follows that the Legendre symbol $\left(\frac{\bullet}{p}\right)$ is multiplicative in the top.

The Jacobi symbol is cooler than the Legendre symbol.

Definition 15.6. The **Jacobi symbol** $\left(\frac{a}{n}\right)$ is defined by extending the Legendre symbol multiplicatively in the bottom.

Hence the Jacobi symbol is completely multiplicative in both parts. It also satisfies:

- $\left(\frac{a}{n}\right) = \left(\frac{b}{n}\right)$ when $a \equiv b \pmod n$.
- $\left(\frac{a}{n}\right) = 0$ if and only if $\gcd(a, n) > 1$ (and is otherwise ± 1).
- $\left(\frac{a}{2}\right) \in \{0, 1\}$ for all a.

Remark 15.7. Warning: $\left(\frac{a}{n}\right)$ doesn't detect quadratic residues modulo n anymore if n is not prime. For example, 2 isn't a quadratic residue modulo *either* 3 or 5, so it is definitely not a quadratic residue modulo 15 either. But $\left(\frac{2}{15}\right) = \left(\frac{2}{3}\right)\left(\frac{2}{5}\right) = (-1)^2 = +1$.

Most importantly, quadratic reciprocity is usually stated for primes, but the statement for Jacobi symbols is cooler.

Theorem 15.8 (Quadratic Reciprocity, with Jacobi symbols)**.** *Let m and n be relatively prime positive odd integers. Then*

$$\left(\frac{-1}{n}\right) = (-1)^{\frac{1}{2}(n-1)}, \qquad \left(\frac{2}{n}\right) = (-1)^{\frac{1}{8}(n^2-1)}$$

and

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = (-1)^{\frac{1}{4}(m-1)(n-1)}.$$

The upshot of this is that you can use the quadratic reciprocity law without having to factor the numerator.

Example 162. Is 481 a quadratic residue modulo 2017?

Solution 162. We have

$$\left(\frac{481}{2017}\right) \overset{\text{QR}}{=} \left(\frac{2017}{481}\right) = \left(\frac{93}{481}\right) \overset{\text{QR}}{=} \left(\frac{481}{93}\right) = \left(\frac{16}{93}\right) = +1.$$

As 2017 is prime, the answer is "yes". ■

Notice how we did not have to factor the numerator (which by the way is $13 \cdot 37$). Thus using the Jacobi symbol instead of the Legendre one makes quadratic reciprocity more powerful (and is indeed one main reason for introducing it).

§15.3 Vieta jumping

Vieta jumping first appeared in the infamous closing problem to IMO 1988, which appears as a walkthrough in Problem 163. This is best seen by example, so we will not say more here.

§15.4 Walkthroughs

Problem 163 (IMO 1988)**.** Let a and b be positive integers. Prove that if

$$\frac{a^2+b^2}{ab+1}$$

is also an integer, then it is a perfect square.

Walkthrough. Let $k = \frac{a^2+b^2}{ab+1}$ be fixed. We will show k is a perfect square.

I want to start with an observation that we'll need later on. The reason I put it here this early is to make sure you realize that it's trivial (and does not require Vieta jumping), before we get lost in the meat of the solution.

(a) Prove that any solution to this equation must satisfy $ab \geq 0$.

The idea behind Vieta jumping is to write this as a quadratic equation

$$a^2 - kb \cdot a + b^2 - k = 0$$

in a; thus for a fixed value of b, we can then "flip" the quadratic in a to get the other value. One might write this as

$$(a,b) \mapsto (k \cdot b - a, b) = \left(\frac{b^2-k}{a}, b\right).$$

Let's do some concrete practice so you can see what I mean.

(b) Let $k = 4$ and observe that $(a, b) = (30, 8)$. Write the quadratic and see how you could realize that $(a, b) = (2, 8)$ was also a solution.

(c) Flip in the other direction: find the other value of b which works with $a = 30$.

(d) Now let's take $(a, b) = (2, 8)$ and flip again, holding $a = 2$ fixed and changing the value of b. What do we get for the other value of b this time?

Thus we see in this problem that every (a, b) automatically has two natural neighbors, one obtained by flipping a and flipping b.

Our goal is to now do this flipping operation in such a way that the pair gets smaller, and see what happens if we keep doing this until we get stuck. (Local, anyone?)

(e) Show that if (a, b) is a solution with $a > b > 0$, then by Vieta jumping we can produce a solution (a', b) with $a' < b$ (but not necessarily $a' > 0$).

(f) Reconcile (a) and (e) to show that we eventually may arrive at a pair in which one component is zero.

(g) Conclude that k is a perfect square.

Problem 164. Prove that $2^n + 1$ has no prime factors of the form $p = 8k + 7$.

Walkthrough. This is a showcase of quadratic reciprocity.

(a) Show that if n is even then all prime divisors of n are $1 \pmod 4$.

(b) Show that if n is odd, then -2 is a quadratic residue modulo p.

(c) Compute $\left(\frac{-2}{p}\right)$ for all primes p.

(d) Use (b) and (c) to finish the problem.

§15.5 Problems

Problem 165. Find all integers $n \geq 1$ such that n divides $2^{n-1} + 3^{n-1}$.

Problem 166 (Bay Area Olympiad 2011)**.** Decide whether there exists a row of Pascal's triangle containing four pairwise distinct numbers a, b, c, d such that $a = 2b$ and $c = 2d$.

Problem 167 (EGMO 2016)**.** Let S be the set of all positive integers n such that n^4 has a divisor in the range $n^2 + 1$, $n^2 + 2$, ..., $n^2 + 2n$. Prove that there are infinitely many elements of S of each of the forms $7m$, $7m + 1$, $7m + 2$, $7m + 5$, $7m + 6$ and no elements of S of the form $7m + 3$ and $7m + 4$, where m is an integer.

Problem 168 (USA TST 2009)**.** Find all pairs of positive integers (m, n) such that $mn - 1$ divides $(n^2 - n + 1)^2$.

Problem 169 (Asian-Pacific Olympiad 1997)**.** Find an integer $100 \leq n \leq 1997$ such that n divides $2^n + 2$.

Problem 170 (IMO Shortlist 2017)**.** Find the smallest positive integer n such that the following holds: there exist infinitely many n-tuples $(a_1, \ldots, a_n)$ of positive rational numbers for which

$$a_1 + \cdots + a_n \qquad \text{and} \qquad \frac{1}{a_1} + \cdots + \frac{1}{a_n}$$

are both integers.

Problem 171. Exhibit a function $s \colon \mathbb{Z}_{>0} \to \mathbb{Z}$ with the following property: if a and b are positive integers such that $p = a^2 + b^2$ is an odd prime, then

$$s(a) \equiv a^{\frac{p-1}{2}} \pmod{p}.$$

Problem 172 (IMO Shortlist 2016)**.** Let a be a positive integer which is not a perfect square, and consider the equation

$$k = \frac{x^2 - a}{x^2 - y^2}.$$

Let A be the set of positive integers k for which the equation admits an integer solution with $x > \sqrt{a}$, and let B be the set of positive integers for which the equation admits an integer solution with $0 \leq x < \sqrt{a}$. Show that $A = B$.

Problem 173 (USA TST 2014)**.** Let a_1, a_2, a_3, ...be a sequence of integers, with the property that every consecutive group of a_i's averages to a perfect square. More precisely, for all positive integers n and k, the quantity

$$\frac{a_n + a_{n+1} + \cdots + a_{n+k-1}}{k}$$

is always the square of an integer. Prove that the sequence must be constant (all a_i are equal to the same perfect square).

§15.6 Solutions

Solution 163. Let $k = \frac{a^2+b^2}{ab+1}$. Then rewrite it as:

$$a^2 - kb \cdot a + b^2 - k = 0.$$

Then we can do a standard Vieta jumping argument. For example, when $k = 4$ the chain goes

$$\cdots \to (112, 30) \to (30, 8) \to (8, 2) \to (2, 0).$$

So, suppose $a > b > 0$ is a solution. Then

$$(a, b) \to (k \cdot b - a, b) = \left(\frac{b^2 - k}{a}, b \right).$$

Notice that $\frac{b^2-k}{a} < a$, so flipping the larger one always decreases.

We have to rule out the possibility of negative numbers in chain. Indeed $k > 0$, so looking at $k = \frac{a^2+b^2}{ab+1}$ shows its impossible for exactly one term to be negative, so eventually one coordinate is zero.

Visibly if $b = 0$ then $k = a^2$ as desired.

Solution 164. Suppose $2^n \equiv -1 \pmod{p}$.

If n is even, then -1 is a quadratic residue, hence $p \equiv 1 \pmod{4}$.

If n is odd, then -2 is a quadratic residue, so

$$+1 = \left(\frac{-2}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right) = \begin{cases} 1 & p \equiv 1, 3 \pmod 8 \\ -1 & p \equiv 5, 7 \pmod 8. \end{cases}$$

So p cannot be $7 \pmod 8$.

Solution 165. The answer is $n = 1$ only. Assume $n > 1$ in the sequel.

Clearly n is odd. Thus for every prime $p \mid 2^{n-1} + 3^{n-1}$, we must have $p \equiv 1 \pmod 4$ (sum of squares). Consequently, $n \equiv 1 \pmod 4$ as well.

Continuing, we then see every prime dividing $2^{n-1} + 3^{n-1}$ is $1 \pmod 8$, and inductively $p \equiv 1 \pmod{2^k}$ whenever $p \mid 2^{n-1} + 3^{n-1}$. This is clearly impossible.

Solution 166. An example is $\binom{203}{68} = 2\binom{203}{67}$ and $\binom{203}{85} = 2\binom{203}{83}$.

To get this, the idea is to look for two adjacent entries and two entries off by one, and solving the corresponding equations. The first one is simple:

$$\binom{n}{j} = 2\binom{n}{j-1} \implies n = 3j - 1.$$

The second one is more involved:

$$\begin{aligned}
\binom{n}{k} &= 2\binom{n}{k-2} \\
\implies (n-k+1)(n-k+2) &= 2k(k-1) \\
\implies 4(n-k+1)(n-k+2) &= 8k(k-1) \\
\implies (2n-2k+3)^2 - 1 &= 2\left((2k-1)^2-1\right) \\
\implies (2n-2k+3)^2 - 2(2k-1)^2 &= -1
\end{aligned}$$

Using standard methods for the Pell equation:

- $(7+5\sqrt{2})(3+2\sqrt{2}) = 41+29\sqrt{2}$. So $k=15$, $n=34$, doesn't work.
- $(41+29\sqrt{2})(3+2\sqrt{2}) = 239+169\sqrt{2}$. Then $k=85$, $n=203$.

Solution 167. I'll just sketch this, since it's not very hard. Note that $n^2+k \mid n^4 \iff n^2+k \mid k^2$ but since $1 \le k \le 2n$ we arrive at only three cases: $n^2+k=k^2$, $2(n^2+k)=k^2$, $3(n^2+k)=k^2$.

The first has no solutions with $k \ge 1$ since we can put $(k-1)^2 < n^2 < k^2$. The other two are Pell equations, and one can check that $n^2 \equiv 2 \pmod 7$ has no solutions at all for $k \pmod 7$ in either case. The assertion about infinitely many solutions then follows by using the Pell recursion, and taking modulo 7.

Solution 168. The answers are $(2,2)$, $(c^2+1,(c+1)^2+1)$, and $((c+1)^2+1,c^2+1)$, for each integer $c \ge 0$. It's easy to see they work and we prove they are the only ones.

Note the condition is equivalent to

$$mn-1 \mid (mn^2-mn+m)^2 \iff mn-1 \mid (m+n-1)^2$$

which in particular is symmetric in m and n.

Now we proceed by Vieta jumping. Fix a $k>0$ for which there exists $m \ge n \ge 1$ obeying

$$(m+n-1)^2 = k(mn-1) \qquad (\star).$$

We prove $k=3$ or $k=4$. The given rearranges as

$$\begin{aligned}
&\iff 0 = m^2+n^2+(2-k)mn-2m-2n+(1-k) \\
&\iff 0 = m^2-(n(k-2)+2)m+(n-1)^2+k.
\end{aligned}$$

Thus given a solution (m,n) with $m \ge n$ we may flip

$$(m,n) \mapsto \left(n, \frac{(n-1)^2+k}{m}\right).$$

Observe that all solutions obtained this way always have both $m, n > 0$. Thus if we flip repeatedly we ought to eventually, for our fixed k, arrive at a solution with $\frac{(n-1)^2+k}{m} \geq m$, otherwise the sum would decrease.

In that situation, we would have

$$\begin{aligned} \frac{(m+n-1)^2}{mn-1} &= k \geq m^2 - (n-1)^2 = (m+n-1)(m-n+1) \\ &\implies m+n-1 \geq (m-n+1)(mn-1). \end{aligned}$$

This last equation is not difficult now (but annoying); we find that it has two cases:

- Suppose $n = 1$. Putting $n = 1$ in $(\star)$ gives $m^2 = k(m-1) \implies k = m + 1 + \frac{1}{m-1}$, which forces $m = 2$ and hence $k = 4$.

  In that case, by flipping, all solutions eventually reach $(2, 1)$ after Vieta jumping. Reversing the Vieta jumping procedure, we work backwards to obtain the curve $\big((c+1)^2 + 1, c^2 + 1\big)$.

- Assume $n \geq 2$; then $m + n - 1 \geq mn - 1 \geq 2m - 1 \implies n \geq m$, so $m = n = 2$. In that case, $k = 3$, and all solutions should arrive here by Vieta jumping. However, $(2, 2)$ is stable, and so this is the only solution in the $k = 3$ situations.

Solution 169. The number $n = 946 = 2 \cdot 11 \cdot 43$ works.

The way you construct is: we try to look for examples of the form $n = 2pq$, where p and q are distinct odd primes. This amounts to $p \mid 2^{2q-1} + 1$ and $q \mid 2^{2p-1} + 1$.

For the second divisibility to hold, it would be nice if we could arrange for

$$2p - 1 = \frac{q-1}{2} \qquad \text{and} \qquad \left(\frac{2}{q}\right) = -1.$$

The first equation rewrites as $q = 4p - 1$. In fact, since this implies $q \equiv 3 \pmod 8$, the second equation would always be true by quadratic reciprocity. This means that $q \mid 2^{2p-1} + 1$ is always true for $q = 4p - 1$.

Going back to the first divisibility, we are hoping for $p \mid 2^{8p-3} + 1 \implies p \mid 2^5 + 1 = 33$. This gives solutions $(p, q) = (3, 11)$ and $(p, q) = (11, 43)$. These give $n = 66$ and $n = 946$ as candidates; the latter works.

Solution 170. The answer is $n = 3$.

First, $n = 1$ clearly fails.

We show $n = 2$ fails: if $a + b = p$ and $\frac{1}{a} + \frac{1}{b} = q$ for integers p and q. Let $a = x/y$ with $\gcd(x, y) = 1$; then

$$q = \frac{1}{a} + \frac{1}{p-a} = \frac{p}{a(p-a)} = \frac{py^2}{x(py-x)}.$$

Then $x \mid p$, so we may write $p = xk$ and obtain

$$q = \frac{xk \cdot y^2}{x^2(ky-1)} = \frac{ky^2}{x(ky-1)}.$$

As $\gcd(ky-1, ky^2) = 1$, this forces $ky - 1 = 1$, or $ky = 2$. If $y = 1$ then $a \in \mathbb{Z}$, so $b \in \mathbb{Z}$, and so either $a = b = 1$ or $a = b = 2$. If $y = 2$, then a and b are both half-integers, and so we conclude $a = b = \frac{1}{2}$.

Now to show $n = 3$ works, we take a triple of the form

$$\left(\frac{1}{1+x+y}, \frac{x}{1+x+y}, \frac{y}{1+x+y} \right)$$

where x, y, z are positive *integers* (in fact if we pick $x, y, z \in \mathbb{Q}$ this is WLOG). Then it suffices that

$$\frac{1+y}{x} + \frac{1+x}{y} \in \mathbb{Z}$$

which is a famous MOP 2007 problem solved by Vieta jumping (there are in fact infinitely many with $\frac{1+y}{x} + \frac{1+x}{y} = 3$).

Solution 171. Note $\gcd(a, p) = 1$, and so interpret $\left(\frac{a}{p}\right)$ as a Legendre symbol. We claim that

$$\left(\frac{a}{p}\right) = \begin{cases} +1 & a \equiv 1 \pmod 2 \\ +1 & a \equiv 0 \pmod 4 \\ -1 & a \equiv 2 \pmod 4. \end{cases}$$

If $p = 2$ this is clear so henceforth assume $p \equiv 1 \pmod 4$. The proof is using the Jacobi symbol.

First, assume a is odd. Then

$$\left(\frac{a}{p}\right) = \left(\frac{p}{a}\right) = \left(\frac{a^2+b^2}{a}\right) = \left(\frac{b^2}{a}\right) = +1.$$

Next, assume $a = 2x$ for x odd. Then $p \equiv 5 \pmod 8$, so $\left(\frac{2}{p}\right) = -1$. Then

$$\left(\frac{a}{p}\right) = \left(\frac{2}{p}\right)\left(\frac{x}{p}\right) = -1 \cdot \left(\frac{p}{x}\right) = -1.$$

Finally, assume $a = 2^e y$ for $e \geq 2$, and y odd. Then $p \equiv 1 \pmod 8$, so $\left(\frac{2}{p}\right) = 1$. Finally

$$\left(\frac{a}{p}\right) = \left(\frac{2}{p}\right)^e \left(\frac{y}{p}\right) = \left(\frac{p}{y}\right) = +1.$$

Remark. Assuming there are infinitely many primes of the form $a^2 + b^2$ for any fixed $a > 0$ (which seems almost certainly true, although it is open), then the function s we gave above is the only one.

Solution 172. The equation (for fixed k, a) is a Pell equation

$$ky^2 - (k-1)x^2 = a.$$

(Doing the case $k = 2$ is good motivation for this.) Based on this we observe the main claim:

Claim. If (u, v) is a solution, then so is

$$((2k-1)u \pm 2kv, (2k-1)v \pm (2k-2)u)$$

where the $\pm$ signs correspond.

Thus one can get larger solutions from small solutions. This implies $B \subseteq A$; we now show the reverse direction by proving that we can descend to solutions with $u < \sqrt{a}$. To be precise:

Claim. Fix a and k, and suppose (u, v) is a solution with $u, v > 0$ and with u minimal. Then $u < \sqrt{a}$.

Proof. We will show that that

$$u > \sqrt{a} \implies |(2k-1)u - 2kv| < u.$$

This consists of two directions, only the latter which requires $u > \sqrt{a}$.

- First, observe unconditionally that

  $$kv^2 - (k-1)u^2 = a > 0 \implies \sqrt{k-1}u < \sqrt{k}v.$$

  Consequently,

  $$kv > \sqrt{k(k-1)}u > (k-1)u$$

  and thus $2kv > (2k-2)u$, so $(2k-1)u - 2kv < u$.

- Assume now $u > \sqrt{a}$. Then the original Pell equation implies $u > v$, since $k(v^2 - u^2) = a - u^2 < 0$. So $(2k-1)u - 2kv > -u$ as well, which is the desired contradiction.

Thus no minimal solution can obey $u > \sqrt{a}$. □

Solution 173. Let $\nu_p(n)$ denote the largest exponent of p dividing n. The problem follows from the following proposition.

Proposition. *Let (a_n) be a sequence of integers and let p be a prime. Suppose that every consecutive group of a_i's with length at most p averages to a perfect square. Then $\nu_p(a_i)$ is independent of i.*

We proceed by induction on the smallest value of $\nu_p(a_i)$ as i ranges (which must be even, as each of the a_i are themselves a square). First we prove two claims.

Claim. If $j \equiv k \pmod{p}$ then $a_j \equiv a_k \pmod{p}$.

Proof. Taking groups of length p in our given, we find that $p \mid a_j + \cdots + a_{j+p-1}$ and $p \mid a_{j+1} + \cdots + a_{j+p}$ for any j. So $a_j \equiv a_{j+p} \pmod{p}$ and the conclusion follows. □

Claim. If some a_i is divisible by p then all of them are.

Proof. The case $p = 2$ is trivial so assume $p \geq 3$. Without loss of generality (via shifting indices) assume that $a_1 \equiv 0 \pmod{p}$, and define

$$S_n = a_1 + a_2 + \cdots + a_n \equiv a_2 + \cdots + a_n \pmod{p}.$$

Call an integer k with $2 \leq k < p$ a **pivot** if $1 - k^{-1}$ is a quadratic nonresidue modulo p.

We claim that for any pivot k, $S_k \equiv 0 \pmod{p}$. If not, then

$$\frac{a_1 + a_2 + \cdots + a_k}{k} \text{ and } \frac{a_2 + \cdots + a_k}{k-1}$$

are both qudaratic residues. Division implies that $\frac{k-1}{k} = 1 - k^{-1}$ is a quadratic residue, contradiction.

Next we claim that there is an integer m with $S_m \equiv S_{m+1} \equiv 0 \pmod{p}$, which implies $p \mid a_{m+1}$. If 2 is a pivot, then we simply take $m = 1$. Otherwise, there are $\frac{1}{2}(p-1)$ pivots, one for each nonresidue (which includes neither 0 nor 1), and all pivots lie in $[3, p-1]$, so we can find an m such that m and $m+1$ are both pivots.

Repeating this procedure starting with a_{m+1} shows that $a_{2m+1}, a_{3m+1}, \ldots$ must all be divisible by p. Combined with the first claim and the fact that $m < p$, we find that all the a_i are divisible by p. □

The second claim establishes the base case of our induction. Now assume all a_i are divisible by p and hence p^2. Then all the averages in our proposition (with length at most p) are divisible by p and hence p^2. Thus the map $a_i \mapsto \frac{1}{p^2} a_i$ gives a new sequence satisfying the proposition, and our inductive hypothesis completes the proof.

Remark. There is a subtle bug that arises if one omits the condition that $k \le p$ in the proposition. When $k = p^2$ the average $\frac{a_1+\cdots+a_{p^2}}{p^2}$ is not necessarily divisible by p even if all the a_i are. Hence it is not valid to divide through by p. This is why the condition $k \le p$ was added.

---

16 Constructions in Number Theory

§16.1 Synopsis

Unlike some previous number theory chapters, here there is more room for you to make choices (e.g. in constructions). As we as saw in the Free chapter, you can often work on a problem in two directions: "experimental" or "restrictive". This dichotomy will be useful to keep in mind.

In addition you will often require number theory skill in order to carry out the correct deductions. (So: globally, it feels like doing a combinatorics problem, but locally, it feels like doing a number theory problem.) This has the weird property that sometimes you'd like to rely on statement that is obviously true ("$n^2 + 1$ is prime infinitely often"), but either hard to prove or open; if you don't know, then you have to make a judgment call. (Whereas in combinatorics, simple true statements are usually easy to prove.)

Two common tropes in this chapter will include:

- Picking really big numbers with lots of prime factors.
- Chinese Remainder Theorem: add modular conditions with reckless abandon, then let the Chinese Remainder Theorem collate them for you.

§16.2 Walkthroughs

Problem 174 (TSTST 2015)**.** Let $\varphi(n)$ denote the number of positive integers less than n that are relatively prime to n. Prove that there exists a positive integer m for which the equation $\varphi(n) = m$ has at least 2015 solutions in n.

Walkthrough. There's a couple of ways to approach this problem. The analytic way to go after it is to try and count the number of obtained φ values. Here's a much more concrete approach. Let's start with some informative examples:

(a) Show that $\varphi(3 \cdot 5000) = \varphi(2 \cdot 5000)$.

(b) Show that $\varphi(11 \cdot 1000) = \varphi(10 \cdot 1000)$.

(c) Find another value of n for which $\varphi(n) = \varphi(10000)$.

The idea is that we have a cushion of primes 2^*5^*. This can work, but we can be much more free with the cushion.

(d) Let $N = 210^{100000000}$. Find some examples of n such that $\varphi(n) = \varphi(N)$, in the spirit of (c).

(e) Construct a set S of 11 prime numbers p for which $p - 1 \mid N$.

(f) Finish the problem.

Problem 175 (USA TST 2015)**.** Prove that for every positive integer n, there exists a set S of n positive integers such that for any two distinct $a, b \in S$, $a - b$ divides a and b but none of the other elements of S.

Walkthrough. The idea is to write

$$S = \{a, a + d_1, a + d_1 + d_2, \dots, a + d_1 + \cdots + d_{n-1}\}$$

and focus on the difference set first, and only then work on the value of a using an application of Chinese remainder theorem.

(a) Find a set S of the form $S = \{a, a + 2, a + 5\}$. (Here $d_1 = 2$, $d_2 = 3$.)

(b) Characterize all the sets S of the form in (a), i.e. those with $(d_1, d_2) = (2, 3)$.

(c) Show that one can find S of the form $S = \{a, a + 600, a + 1500\}$.

(d) Show that one can find S of the form $S = \{a, a+600, a+1500, a+1507\}$.

(e) Suppose there is a set S of size n with differences $(d_1, \dots, d_{n-1})$. Prove that we can find an integer M and prime p, such that there is a set S of size $n + 1$ with difference sequence $(Md_1, \dots, Md_{n-1}, p)$.

(f) Conclude.

§16.3 Problems

Problem 176 (USAMO 2017)**.** Prove that there exist infinitely many pairs of relatively prime positive integers $a, b > 1$ for which $a + b$ divides $a^b + b^a$.

Problem 177 (USAMO 2008)**.** Prove that for each positive integer n, there are pairwise relatively prime integers k_0, ..., k_n, all strictly greater than 1, such that $k_0k_1 \dots k_n - 1$ is the product of two consecutive integers.

Problem 178 (IMO Shortlist 2010)**.** Find the least positive integer n for which there exists a set $\{s_1 < s_2 < \dots < s_n\}$ consisting of n distinct positive integers satisfying

$$\left(1 - \frac{1}{s_1}\right)\left(1 - \frac{1}{s_2}\right) \dots \left(1 - \frac{1}{s_n}\right) = \frac{51}{2010}.$$

Problem 179 (USA TST 2007)**.** Determine whether or not there exist positive integers a and b such that a does not divide $b^n - n$ for all positive integers n.

Problem 180 (EGMO 2018)**.** Consider the set

$$A = \left\{ 1 + \frac{1}{k} : k = 1, 2, 3, \dots \right\}.$$

For every integer $x \geq 2$, let $f(x)$ denote the minimum integer such that x can be written as the product of $f(x)$ elements of A (not necessarily distinct). Prove that there are infinitely many pairs of integers $x \geq 2$ and $y \geq 2$ for which

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

Problem 181 (USAJMO 2016)**.** Prove that there exists a positive integer $n < 10^6$ such that 5^n has six consecutive zeros in its decimal representation.

Problem 182 (EGMO 2014)**.** We denote the number of positive divisors of a positive integer m by $d(m)$ and the number of distinct prime divisors of m by $\omega(m)$. Let k be a positive integer. Prove that there exist infinitely many positive integers n such that $\omega(n) = k$ and $d(n)$ does not divide $d(a^2 + b^2)$ for any positive integers a, b satisfying $a + b = n$.

Problem 183 (USAMO 2013)**.** Let m and n be positive integers. Prove that there exists a positive integer c such that cm and cn have the same nonzero decimal digits.

Problem 184 (TSTST 2016)**.** Decide whether or not there exists a nonconstant polynomial $Q(x)$ with integer coefficients with the following property: for every positive integer $n > 2$, the numbers

$$Q(0),\ Q(1), Q(2),\ \dots,\ Q(n-1)$$

produce at most $0.499n$ distinct residues when taken modulo n.

Problem 185 (IMO 2003)**.** Let p be a prime number. Prove that there exists a prime number q such that for every integer n, the number $n^p - p$ is not divisible by q.

§16.4 Solutions

Solution 174. Here are two explicit solutions.

First solution with ad-hoc subsets, by Evan Chen I consider the following eleven prime numbers:

$$S = \{11, 13, 17, 19, 29, 31, 37, 41, 43, 61, 71\}.$$

This has the property that for any $p \in S$, all prime factors of $p-1$ are one digit.

Let $N = (210)^{\text{billion}}$, and consider $M = \varphi(N)$. For any subset $T \subset S$, we have

$$M = \varphi\left(\frac{N}{\prod_{p\in T}(p-1)} \prod_{p\in T} p\right).$$

Since $2^{|S|} > 2015$ we're done.

Remark. This solution is motivated by the deep fact that $\varphi(11 \cdot 1000) = \varphi(10 \cdot 1000)$, for example.

Second solution with smallest primes, by Yang Liu Let $2 = p_1 < p_2 < \cdots < p_{2015}$ be the *smallest* 2015 primes. Then the 2015 numbers

$$\begin{aligned} n_1 &= (p_1 - 1)p_2 \dots p_{2015} \\ n_2 &= p_1(p_2 - 1) \dots p_{2015} \\ &\vdots \\ n_{2015} &= p_1 p_2 \dots (p_{2015} - 1) \end{aligned}$$

all have the same phi value, namely

$$\varphi(p_1 p_2 \dots p_{2015}) = \prod_{i=1}^{2015} (p_i - 1).$$

Solution 175. The idea is to look for a sequence $d_1, \dots, d_{n-1}$ of "differences" such that the following two conditions hold. Let $s_i = d_1 + \cdots + d_{i-1}$, and $t_{i,j} = d_i + \cdots + d_{j-1}$ for $i \le j$.

(i) No two of the $t_{i,j}$ divide each other.

(ii) There exists an integer a satisfying the CRT equivalences

$$a \equiv -s_i \pmod{t_{i,j}} \quad \forall i \le j$$

Then the sequence $a + s_1$, $a + s_2$, ..., $a + s_n$ will work. For example, when $n = 3$ we can take $(d_1, d_2) = (2, 3)$ giving

$$10 \overbrace{\underbrace{}_{2} 12 \underbrace{}_{3}}^{5} 15$$

because the only conditions we need satisfy are

$$\begin{aligned} a &\equiv 0 \pmod 2 \\ a &\equiv 0 \pmod 5 \\ a &\equiv -2 \pmod 3. \end{aligned}$$

But with this setup we can just construct the d_i inductively. To go from n to $n+1$, take a $d_1, \dots, d_{n-1}$ and let p be a prime not dividing any of the d_i. Moreover, let $M = \prod_{i=1}^{n-1} d_i$. Then we claim that $d_1 M, d_2 M, \dots, d_{n-1} M, p$ is such a difference sequence. For example, the previous example extends as follows.

$$a \underbrace{}_{600} b \overbrace{\underbrace{}_{900} c \underbrace{}_{7}}^{907} d \overbrace{}^{1507}$$

The new numbers p, $p + Md_{n-1}$, $p + Md_{n-2}$, ...are all relatively prime to everything else. Hence (i) still holds. To see that (ii) still holds, just note that we can still get a family of solutions for the first n terms, and then the last $(n+1)$st term can be made to work by Chinese Remainder Theorem since all the new $p + Md_k$ are coprime to everything.

Solution 176. One construction: let $d \equiv 1 \pmod 4$, $d > 1$. Let $x = \frac{d^d + 2^d}{d+2}$. Then set

$$a = \frac{x+d}{2}, \qquad b = \frac{x-d}{2}.$$

To see this works, first check that b is odd and a is even. Let $d = a - b$ be odd. Then:

$$\begin{aligned} a + b \mid a^b + b^a &\iff (-b)^b + b^a \equiv 0 \pmod{a+b} \\ &\iff b^{a-b} \equiv 1 \pmod{a+b} \\ &\iff b^d \equiv 1 \pmod{d+2b} \\ &\iff (-2)^d \equiv d^d \pmod{d+2b} \\ &\iff d + 2b \mid d^d + 2^d. \end{aligned}$$

So it would be enough that

$$d + 2b = \frac{d^d + 2^d}{d+2} \implies b = \frac{1}{2}\left(\frac{d^d + 2^d}{d+2} - d \right)$$

which is what we constructed. Also, since $\gcd(x,d)=1$ it follows $\gcd(a,b)=\gcd(d,b)=1$.

Remark. Ryan Kim points out that in fact, $(a,b)=(2n-1,2n+1)$ is always a solution.

Solution 177. In other words, if we let

$$P(x)=x(x+1)+1$$

then we would like there to be infinitely many primes dividing some $P(t)$ for some integer t.

In fact, this result is true in much greater generality. We first state:

Theorem 16.1 (Schur's theorem)**.** *If $P(x)\in\mathbb{Z}[x]$ is nonconstant and $P(0)=1$, then there are infinitely many primes which divide $P(t)$ for some integer t.*

Proof. If $P(0)=0$, this is clear. So assume $P(0)=c\neq 0$.

Let S be any finite set of prime numbers. Consider then the value

$$P\left(k\prod_{p\in S}p\right)$$

for some integer k. It is $1 \pmod p$ for each prime p, and if k is large enough it should not be equal to 1 (because P is not constant). Therefore, it has a prime divisor not in S. □

Remark. In fact the result holds without the assumption $P(0)\neq 1$. The proof requires only small modifications, and a good exercise would be to write down a similar proof that works first for $P(0)=20$, and then for any $P(0)\neq 0$. (The $P(0)=0$ case is vacuous, since then $P(x)$ is divisible by x.)

To finish the proof, let p_1, ..., p_n be primes and x_i be integers such that

$$\begin{aligned} P(x_1)&\equiv 0 \pmod{p_1}\\ P(x_2)&\equiv 0 \pmod{p_2}\\ &\vdots\\ P(x_n)&\equiv 0 \pmod{p_n}\end{aligned}$$

as promised by Schur's theorem. Then, by Chinese remainder theorem, we can find x such that $x\equiv x_i \pmod{p_i}$ for each i, whence $P(x)$ has at least n prime factor.

Solution 178. The answer is $n = 39$.

To see this is optimal, assume $s_i > 1$ for all i forever after. Then for any n,

$$\prod \left(1 - \frac{1}{s_i}\right) \geq \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right) \dots \left(1 - \frac{1}{n+1}\right) = \frac{1}{n+1}$$

and since $\frac{51}{2010} < \frac{1}{39}$, we need $n + 1 > 39$, or $n > 38$.

As for a construction when $n = 39$, note that

$$\left(\frac{1}{2} \cdot \frac{2}{3} \cdot \dots \cdot \frac{32}{33}\right) \cdot \left(\frac{34}{35} \cdot \dots \cdot \frac{39}{40}\right) \cdot \frac{66}{67}$$

works, since it equals $\frac{1/40}{33/34} \cdot \frac{66}{67} = \frac{17}{670} = \frac{51}{2010}$.

Solution 179. The answer is no.

In fact, for any fixed integer b, the sequence

$$b,\ b^b,\ b^{b^b},\ \dots$$

is eventually constant modulo any integer. (This follows by induction on the exponent: for it to be eventually constant modulo a, it is enough to be eventually constant modulo $\varphi(a)$, hence modulo $\varphi(\varphi(a))$, etc.)

Therefore if n is a suitably tall power-tower of b's, then we will have $b^n \equiv n \pmod{a}$.

Solution 180. One of many constructions: let $n = 2^e + 1$ for $e \equiv 5 \pmod{10}$ and let $x = 11$, $y = n/11$ be our two integers.

We prove two lemmas:

Claim. For any $m \geq 2$ we have $f(m) \geq \lceil \log_2 m \rceil$.

Proof. This is obvious. □

It follows that $f(n) = e + 1$, since $n = \frac{n}{n-1} \cdot 2^e$.

Claim. $f(11) = 5$.

Proof. We have $11 = \frac{33}{32} \cdot \frac{4}{3} \cdot 2^3$. So it suffices to prove $f(11) > 4$.

Note that a decomposition of 11 must contain a fraction at most $\frac{11}{10} = 1.1$. But $2^3 \cdot 1.1 = 8.8 < 11$, contradiction. □

To finish, note that

$$f(11) + f(n/11) \geq 5 + \log_2(n/11) = 1 + \log_2(16n/11) > 1 + e = 1 + f(n).$$

Remark. Most solutions seem to involve picking n such that $f(n)$ is easy to compute. Indeed, it's hard to get nontrivial lower bounds other than the log, and even harder to actually come up with complicated constructions. It might be said the key to this problem is doing as little number theory as possible.

Solution 181. We will prove that $\boxed{n = 20 + 2^{19} = 524308}$ fits the bill.

First, we claim that

$$5^n \equiv 5^{20} \pmod{5^{20}} \qquad \text{and} \qquad 5^n \equiv 5^{20} \pmod{2^{20}}.$$

Indeed, the first equality holds since both sides are $0 \pmod{5^{20}}$, and the second by $\varphi(2^{20}) = 2^{19}$ and Euler's theorem. Hence

$$5^n \equiv 5^{20} \pmod{10^{20}}.$$

In other words, the last 20 digits of 5^n will match the decimal representation of 5^{20}, with leading zeros. However, we have

$$5^{20} = \frac{1}{2^{20}} \cdot 10^{20} < \frac{1}{1000^2} \cdot 10^{20} = 10^{-6} \cdot 10^{20}$$

and hence those first six of those 20 digits will all be zero. This completes the proof! (To be concrete, it turns out that $5^{20} = 95367431640625$ and so the last 20 digits of 5^n will be 00000095367431640625.)

Remark. Many of the first posts in the JMO 2016 discussion thread (see `https://aops.com/community/c5h1230514`) claimed that the problem was "super easy". In fact, the problem was solved by only about 10% of contestants.

Solution 182. Let $n = 2^{p-1}t$, where $t \equiv 5 \pmod 6$, $\omega(t) = k - 1$, and $p \gg t$ is a sufficiently large prime. Let $a + b = n$ and $a^2 + b^2 = c$. We claim that $p \nmid d(c)$, which solves the problem since $p \mid d(n)$.

First, note that $3 \nmid a^2 + b^2$, since $3 \nmid n$. Next, note that $c < 2n^2 < 5^{p-1}$ (since $p \gg t$) so no exponent of an odd prime in c exceeds $p - 2$. Moreover, $c < 2^{3p-1}$.

So, it remains to check that $\nu_2(c) \notin \{p - 1, 2p - 1\}$. On the one hand, if $\nu_2(a) < \nu_2(b)$, then $\nu_2(a) = p - 1$ and $\nu_2(c) = 2\nu_2(a) = 2p - 2$. On the other hand, if $\nu_2(a) = \nu_2(b)$ then $\nu_2(a) \le p - 2$, and $\nu_2(c) = 2\nu_2(a) + 1$ is odd and less than $2p - 1$.

Remark. Weird problem. The condition is very artificial, although the construction is kind of fun. I'm guessing the low scores during the actual contest were actually due to an unusually tricky P2.

Solution 183. One-line spoiler: 142857. More verbosely, the idea is to look at the decimal representation of $1/D$, m/D, n/D for a suitable denominator D, which have a "cyclic shift" property in which the digits of n/D are the digits of m/D shifted by 3.

Remark (An example to follow along)**.** Here is an example to follow along in the subsequent proof If $m = 4$ and $n = 23$ then the magic numbers $e = 3$ and $D = 41$ obey

$$10^3 \cdot \frac{4}{41} = 97 + \frac{23}{41}.$$

The idea is that

$$\begin{aligned} \frac{1}{41} &= 0.\overline{02439} \\ \frac{4}{41} &= 0.\overline{09756} \\ \frac{23}{41} &= 0.\overline{56097} \end{aligned}$$

and so $c = 2349$ works; we get $4c = 9756$ and $23c = 56097$ which are cyclic shifts of each other by 3 places (with some leading zeros appended).

Here is the one to use:

Claim. There exists positive integers D and e such that $\gcd(D, 10) = 1$, $D > \max(m, n)$, and moreover

$$\frac{10^e m - n}{D} \in \mathbb{Z}.$$

Proof. Suppose we pick some exponent e and define the number

$$A = 10^e n - m.$$

Let $r = \nu_2(m)$ and $s = \nu_5(m)$. As long as $e > \max(r, s)$ we have $\nu_2(A) = r$ and $\nu_5(A) = s$, too. Now choose any $e > \max(r, s)$ big enough that $A > 2^r 5^s \max(m, n)$ also holds. Then the number $D = \frac{A}{2^r 5^s}$ works; the first two properties hold by construction and

$$10^e \cdot \frac{n}{D} - \frac{m}{D} = \frac{A}{D} = 2^r 5^s$$

is an integer. □

Remark (For people who like obscure theorems)**.** Kobayashi's theorem implies we can actually pick D to be prime.

Now we take c to be the number under the bar of $1/D$ (leading zeros removed). Then the decimal representation of $\frac{m}{D}$ is the decimal representation of cm repeated (possibly including leading zeros). Similarly, $\frac{n}{D}$ has the decimal representation of cm repeated (possibly including leading zeros). Finally, since

$$10^e \cdot \frac{m}{D} - \frac{n}{D} \text{ is an integer}$$

it follows that these repeating decimal representations are rotations of each other by e places, so in particular they have the same number of nonzero digits.

Remark. Many students tried to find a D satisfying the stronger hypothesis that $1/D$, $2/D$, ..., $(D-1)/D$ are cyclic shifts of each other. For example, this holds in the famous $D = 7$ case.

The official USAMO 2013 solutions try to do this by proving that 10 is a primitive root modulo 7^e for each $e \geq 1$, by Hensel lifting lemma. I think this argument is actually *incorrect*, because it breaks if either m or n are divisible by 7. Put bluntly, $\frac{7}{49}$ and $\frac{8}{49}$ are not shifts of each other.

One may be tempted to resort to using large primes D rather than powers of 7 to deal with this issue. However it is an open conjecture (a special case of Artin's primitive root conjecture) whether or not 10 (mod p) is primitive infinitely often, which is the necessary conjecture so this is harder than it seems.

Solution 184. We claim that

$$Q(x) = 420(x^2 - 1)^2$$

works. Clearly, it suffices to prove the result when $n = 4$ and when n is an odd prime p. The case $n = 4$ is trivial, so assume now $n = p$ is an odd prime.

First, we prove the following easy claim.

Claim. For any odd prime p, there are at least $\frac{1}{2}(p-3)$ values of a for which $\left(\frac{1-a^2}{p}\right) = +1$.

Proof. Note that if $k \neq 0$, $k \neq \pm 1$, $k^2 \neq -1$, then $a = 2(k + k^{-1})^{-1}$ works. Also $a = 0$ works. □

Let $F(x) = (x^2 - 1)^2$. The range of F modulo p is contained within the $\frac{1}{2}(p+1)$ quadratic residues modulo p. On the other hand, if for some t neither of $1 \pm t$ is a quadratic residue, then t^2 is omitted from the range of F as well. Call such a value of t *useful*, and let N be the number of useful residues. We aim to show $N \geq \frac{1}{4}p - 2$.

We compute a lower bound on the number N of useful t by writing

$$\begin{aligned}
N &= \frac{1}{4}\left(\sum_t \left[\left(1-\left(\frac{1-t}{p}\right)\right)\left(1-\left(\frac{1+t}{p}\right)\right)\right] - \left(1-\left(\frac{2}{p}\right)\right) - \left(1-\left(\frac{-2}{p}\right)\right)\right) \\
&\geq \frac{1}{4}\sum_t \left[\left(1-\left(\frac{1-t}{p}\right)\right)\left(1-\left(\frac{1+t}{p}\right)\right)\right] - 1 \\
&= \frac{1}{4}\left(p + \sum_t \left(\frac{1-t^2}{p}\right)\right) - 1 \\
&\geq \frac{1}{4}\left(p + (+1)\cdot \tfrac{1}{2}(p-3) + 0\cdot 2 + (-1)\cdot((p-2) - \tfrac{1}{2}(p-3))\right) - 1 \\
&\geq \frac{1}{4}(p-5).
\end{aligned}$$

Thus, the range of F has size at most

$$\frac{1}{2}(p+1) - \frac{1}{2}N \leq \frac{3}{8}(p+3).$$

This is less than $0.499p$ for any $p \geq 11$.

Remark. In fact, the computation above is essentially an equality. There are only two points where terms are dropped: one, when $p \equiv 3 \pmod 4$ there are no $k^2 = -1$ in the lemma, and secondly, the terms $1-(2/p)$ and $1-(-2/p)$ are dropped in the initial estimate for N. With suitable modifications, one can show that in fact, the range of F is exactly equal to

$$\frac{1}{2}(p+1) - \frac{1}{2}N = \begin{cases} \frac{1}{8}(3p+5) & p \equiv 1 \pmod 8 \\ \frac{1}{8}(3p+7) & p \equiv 3 \pmod 8 \\ \frac{1}{8}(3p+9) & p \equiv 5 \pmod 8 \\ \frac{1}{8}(3p+3) & p \equiv 7 \pmod 8. \end{cases}$$

Solution 185. By orders, we must have $q = pk+1$ for this to be possible. So we just need $n^p \not\equiv p \iff p^k \not\equiv 1 \pmod q$.

So we need a prime $q \equiv 1 \pmod p$ such that $p^k \not\equiv 1 \pmod q$. To do this, we first recall the following lemma.

Lemma. *Let $\Phi_p(X) = 1 + X + X^2 + \cdots + X^{p-1}$. For any integer a, if q is a prime divisor of $\Phi_p(a)$ other than p, then $a \pmod q$ has order p. (In particular, $q \equiv 1 \pmod p$.)*

Proof. We have $a^p - 1 \equiv 0 \pmod q$, so either the order is 1 or p. If it is 1, then $a \equiv 1 \pmod q$, so $q \mid \Phi_p(1) = p$, hence $q = p$. □

Now the idea is to extract a prime factor q from the cyclotomic polynomial

$$\Phi_p(p) = \frac{p^p - 1}{p-1} \equiv 1 + p \pmod{p^2}$$

such that $q \not\equiv 1 \pmod{p^2}$; hence $k \not\equiv 0 \pmod{p}$, and as $p \pmod{q}$ has order p we have $p^k \not\equiv 1 \pmod{q}$.

---

17 Selected Number Theory from USA TST

§17.1 Problems

Problem 186 (USAMO 2018)**.** Let p be a prime, and let a_1, ..., a_p be integers. Show that there exists an integer k such that the numbers

$$a_1 + k,\ a_2 + 2k,\ \ldots,\ a_p + pk$$

produce at least $\frac{1}{2}p$ distinct remainders upon division by p.

Problem 187 (TSTST 2018)**.** As usual, let $\mathbb{Z}[x]$ denote the set of single-variable polynomials in x with integer coefficients. Find all functions $\theta\colon \mathbb{Z}[x] \to \mathbb{Z}$ such that for any polynomials $p, q \in \mathbb{Z}[x]$,

- $\theta(p+1) = \theta(p) + 1$, and
- if $\theta(p) \neq 0$ then $\theta(p)$ divides $\theta(p \cdot q)$.

Problem 188 (TSTST 2018)**.** For which positive integers $b > 2$ do there exist infinitely many positive integers n such that n^2 divides $b^n + 1$?

Problem 189 (USA TST 2018)**.** Let $n \geq 2$ be a positive integer, and let $\sigma(n)$ denote the sum of the positive divisors of n. Prove that the n^{th} smallest positive integer relatively prime to n is at least $\sigma(n)$, and determine for which n equality holds.

Problem 190 (USA TST 2017)**.** Prove that there are infinitely many triples (a, b, p) of integers, with p prime and $0 < a \leq b < p$, for which p^5 divides $(a+b)^p - a^p - b^p$.

Problem 191 (TSTST 2015)**.** Let P be the set of all primes, and let M be a non-empty subset of P. Suppose that for any non-empty subset $\{p_1, p_2, \ldots, p_k\}$ of M, all prime factors of $p_1 p_2 \cdots p_k + 1$ are also in M. Prove that $M = P$.

Problem 192 (USA TST 2019)**.** Let $\mathbb{Z}/n\mathbb{Z}$ denote the set of integers considered modulo n (hence $\mathbb{Z}/n\mathbb{Z}$ has n elements). Find all positive integers n for which there exists a bijective function $g\colon \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, such that the 101 functions

$$g(x), \quad g(x) + x, \quad g(x) + 2x, \quad \ldots, \quad g(x) + 100x$$

are all bijections on $\mathbb{Z}/n\mathbb{Z}$.

§17.2 Solutions

Solution 186. For each $k = 0, \dots, p-1$ let G_k be the graph on on $\{1, \dots, p\}$ where we join $\{i, j\}$ if and only if

$$a_i + ik \equiv a_j + jk \pmod p \iff k \equiv -\frac{a_i - a_j}{i-j} \pmod p.$$

So we want a graph G_k with at least $\frac{1}{2}p$ connected components.

However, each $\{i, j\}$ appears in exactly one graph G_k, so some graph has at most $\frac{1}{p}\binom{p}{2} = \frac{1}{2}(p-1)$ edges (by "pigeonhole"). This graph has at least $\frac{1}{2}(p+1)$ connected components, as desired.

Remark. Here is an example for $p = 5$ showing equality can occur:

$$\begin{bmatrix} 0 & 0 & 3 & 4 & 3 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & 2 & 2 & 0 & 1 \\ 0 & 3 & 4 & 3 & 0 \\ 0 & 4 & 1 & 1 & 4 \end{bmatrix}.$$

Ankan Bhattacharya points out more generally that $a_i = i^2$ is sharp in general.

Solution 187. The answer is $\theta : p \mapsto p(c)$, for each choice of $c \in \mathbb{Z}$. Obviously these work, so we prove these are the only ones. In what follows, $x \in \mathbb{Z}[x]$ is the identity polynomial, and $c = \theta(x)$.

First solution (Merlijn Staps) Consider an integer $n \neq c$. Because $x - n \mid p(x) - p(n)$, we have

$$\theta(x-n) \mid \theta(p(x) - p(n)) \implies c - n \mid \theta(p(x)) - p(n).$$

On the other hand, $c - n \mid p(c) - p(n)$. Combining the previous two gives $c - n \mid \theta(p(x)) - p(c)$, and by letting n large we conclude $\theta(p(x)) - p(c) = 0$, so $\theta(p(x)) = p(c)$.

Second solution First, we settle the case $\deg p = 0$. In that case, from the second property, $\theta(m) = m + \theta(0)$ for every integer $m \in \mathbb{Z}$ (viewed as a constant polynomial). Thus $m + \theta(0) \mid 2m + \theta(0)$, hence $m + \theta(0) \mid -\theta(0)$, so $\theta(0) = 0$ by taking m large. Thus $\theta(m) = m$ for $m \in \mathbb{Z}$.

Next, we address the case of $\deg p = 1$. We know $\theta(x + b) = c + b$ for $b \in \mathbb{Z}$. Now for each particular $a \in \mathbb{Z}$, we have

$$c + k \mid \theta(x + k) \mid \theta(ax + ak) = \theta(ax) + ak \implies c + k \mid \theta(ax) - ac.$$

for any $k \neq -c$. Since this is true for large enough k, we conclude $\theta(ax) = ac$. Thus $\theta(ax+b) = ac+b$.

We now proceed by induction on $\deg p$. Fix a polynomial p and assume it's true for all p of smaller degree. Choose a large integer n (to be determined later) for which $p(n) \neq p(c)$. We then have

$$\frac{p(c)-p(n)}{c-n} = \theta\left(\frac{p-p(n)}{x-n}\right) \mid \theta\left(p-p(n)\right) = \theta(p) - p(n).$$

Subtracting off $c-n$ times the left-hand side gives

$$\frac{p(c)-p(n)}{c-n} \mid \theta(p) - p(c).$$

The left-hand side can be made arbitrarily large by letting $n \to \infty$, since $\deg p \geq 2$. Thus $\theta(p) = p(c)$, concluding the proof.

Solution 188. This problem is sort of the union of IMO 1990/3 and IMO 2000/5.

The answer is any b such that $b+1$ is not a power of 2. In the forwards direction, we first prove more carefully the following claim.

Claim. If $b+1$ is a power of 2, then the only n which is valid is $n = 1$.

Proof. Assume $n > 1$ and let p be the smallest prime dividing n. We cannot have $p = 2$, since then $4 \mid b^n + 1 \equiv 2 \pmod 4$. Thus,

$$b^{2n} \equiv 1 \pmod p$$

so the order of $b \pmod p$ divides $\gcd(2n, p-1) = 2$. Hence $p \mid b^2 - 1 = (b-1)(b+1)$.

But since $b+1$ was a power of 2, this forces $p \mid b-1$. Then $0 \equiv b^n + 1 \equiv 2 \pmod p$, contradiction. □

On the other hand, suppose that $b+1$ is not a power of 2 (and that $b > 2$). We will inductively construct an infinite sequence of distinct primes $p_0, p_1, \ldots$, such that the following two properties hold for each $k \geq 0$:

- $p_0^2 \ldots p_{k-1}^2 p_k \mid b^{p_0 \cdots p_{k-1}} + 1$,
- and hence $p_0^2 \ldots p_{k-1}^2 p_k^2 \mid b^{p_0 \cdots p_{k-1} p_k} + 1$ by exponent lifting lemma.

This will solve the problem.

Initially, let p_0 be any odd prime dividing $b+1$. For the inductive step, we contend there exists an *odd* prime $q \notin \{p_0, \ldots, p_k\}$ such that $q \mid b^{p_0 \cdots p_k} + 1$. Indeed, this follows immediately by Zsigmondy theorem since $p_0 \ldots p_k$ divides $b^{p_0 \cdots p_{k-1}} + 1$. Since $(b^{p_0 \cdots p_k})^q \equiv b^{p_0 \cdots p_k} \pmod q$, it follows we can then take $p_{k+1} = q$. This finishes the induction.

To avoid the use of Zsigmondy, one can instead argue as follows: let $p = p_k$ for brevity, and let $c = b^{p_0 \cdots p_{k-1}}$. Then $\frac{c^p+1}{c+1} = c^{p-1} - c^{p-2} + \cdots + 1$ has GCD exactly p with $c+1$. Moreover, this quotient is always odd. Thus as long as $c^p + 1 > p \cdot (c+1)$, there will be some new prime dividing $c^p + 1$ but not $c+1$. This is true unless $p = 3$ and $c = 2$, but we assumed $b > 2$ so this case does not appear.

Remark (On new primes)**.** In going from $n^2 \mid b^n+1$ to $(nq)^2 \mid b^{nq}+1$, one does not necessarily need to pick a q such that $q \nmid n$, as long as $\nu_q(n^2) < \nu_q(b^n+1)$. In other words it suffices to just check that $\frac{b^n+1}{n^2}$ is not a power of 2 in this process.

However, this calculation is a little more involved with this approach. One proceeds by noting that n is odd, hence $\nu_2(b^n + 1) = \nu_2(b + 1)$, and thus $\frac{b^n+1}{n^2} = 2^{\nu_2(b+1)} \le b + 1$, which is a little harder to bound than the analogous $c^p + 1 > p \cdot (c + 1)$ from the previous solution.

Solution 189. The equality case is $n = p^e$ for p prime and a positive integer e. It is easy to check that this works.

First solution In what follows, by $[a, b]$ we mean $\{a, a + 1, \dots, b\}$. First, we make the following easy observation.

Claim. If a and d are positive integers, then precisely $\varphi(d)$ elements of $[a, a + d - 1]$ are relatively prime to d.

Let d_1, d_2, ..., d_k denote denote the divisors of n in some order. Consider the intervals

$$\begin{aligned} I_1 &= [1, d_1], \\ I_2 &= [d_1 + 1, d_1 + d_2] \\ &\vdots \\ I_k &= [d_1 + \cdots + d_{k-1} + 1, d_1 + \cdots + d_k]. \end{aligned}$$

of length $d_1, \dots, d_k$ respectively. The jth interval will have exactly $\varphi(d_j)$ elements which are relatively prime d_j, hence at most $\varphi(d_j)$ which are relatively prime to n. Consequently, in $I = \bigcup_{j=1}^{k} I_k$ there are at most

$$\sum_{j=1}^{k} \varphi(d_j) = \sum_{d \mid n} \varphi(d) = n$$

integers relatively prime to n. On the other hand $I = [1, \sigma(n)]$ so this implies the inequality.

We see that the equality holds for $n = p^e$. Assume now $p < q$ are distinct primes dividing n. Reorder the divisors d_i so that $d_1 = q$. Then $p, q \in I_1$, and so I_1 should contain strictly fewer than $\varphi(d_1) = q - 1$ elements relatively prime to n, hence the inequality is strict.

Second solution (Ivan Borsenco and Evan Chen) Let $n = p_1^{e_1} \dots p_k^{e_k}$, where $p_1 < p_2 < \dots$. We are going to assume $k \geq 2$, since the $k = 1$ case was resolved in the very beginning, and prove the strict inequality.

For a general N, the number of relatively prime integers in $[1, N]$ is given exactly by

$$f(N) = N - \sum_i \left\lfloor \frac{N}{p_i} \right\rfloor + \sum_{i<j} \left\lfloor \frac{N}{p_i p_j} \right\rfloor - \dots$$

according to the inclusion-exclusion principle. So, we wish to show that $f(\sigma(n)) < n$ (as $k \geq 2$). Discarding the error terms from the floors (noting that we get at most 1 from the negative floors) gives

$$\begin{aligned} f(N) &< 2^{k-1} + N - \sum_i \frac{N}{p_i} + \sum_{i<j} \frac{N}{p_i p_j} - \dots \\ &= 2^{k-1} + N \prod_i \left(1 - p_i^{-1}\right) \\ &= 2^{k-1} + \prod_i \left(1 - p_i^{-1}\right)\left(1 + p_i + p_i^2 + \dots + p_i^{e_i}\right) \\ &= 2^{k-1} + \prod_i \left(p_i^{e_i} - p_i^{-1}\right). \end{aligned}$$

The proof is now divided into two cases. If $k = 2$ we have

$$\begin{aligned} f(N) &< 2 + \left(p_1^{e_1} - p_1^{-1}\right)\left(p_2^{e_2} - p_2^{-1}\right) \\ &= 2 + n - \frac{p_2^{e_2}}{p_1} - \frac{p_1^{e_1}}{p_2} + \frac{1}{p_1 p_2} \\ &\leq 2 + n - \frac{p_2}{p_1} - \frac{p_1}{p_2} + \frac{1}{p_1 p_2} \\ &= n + \frac{1 - (p_1 - p_2)^2}{p_1 p_2} \leq n. \end{aligned}$$

On the other hand if $k \geq 3$ we may now write

$$\begin{aligned} f(N) &< 2^{k-1} + \left[\prod_{i=2}^{k-1} \left(p_i^{e_i}\right)\right]\left(p_1^{e_1} - p_1^{-1}\right) \\ &= 2^{k-1} + n - \frac{p_2^{e_2} \dots p_k^{e_k}}{p_1} \\ &\leq 2^{k-1} + n - \frac{p_2 p_3 \dots p_k}{p_1}. \end{aligned}$$

If $p_1 = 2$, then one can show by induction that $p_2 p_3 \dots p_k \geq 2^{k+1} - 1$, which implies the result. If $p_1 > 2$, then one can again show by induction $p_3 \dots p_k \geq 2^k - 1$ (since $p_3 \geq 7$), which also implies the result.

Solution 190. The key claim is that if $p \equiv 1 \pmod 3$, then

$$p(x^2 + xy + y^2)^2 \text{ divides } (x+y)^p - x^p - y^p$$

as polynomials in x and y. Since it's known that one can select a and b such that $p^2 \mid a^2 + ab + b^2$, the conclusion follows. (The theory of quadratic forms tells us we can do it with $p^2 = a^2 + ab + b^2$; Thue's lemma lets us do it by solving $x^2 + x + 1 \equiv 0 \pmod{p^2}$.)

To prove this, it is the same to show that

$$(x^2 + x + 1)^2 \text{ divides } F(x) \coloneqq (x+1)^p - x^p - 1.$$

since the binomial coefficients $\binom{p}{k}$ are clearly divisible by p. Let ζ be a third root of unity. Then $F(\zeta) = (1+\zeta)^p - \zeta^p - 1 = -\zeta^2 - \zeta - 1 = 0$. Moreover, $F'(x) = p(x+1)^{p-1} - px^{p-1}$, so $F'(\zeta) = p - p = 0$. Hence ζ is a double root of F as needed.

(Incidentally, $p = 2017$ works!)

Remark. One possible motivation for this solution is the case $p = 7$. It is nontrivial even to prove that p^2 can divide the expression if we exclude the situation $a + b = p$ (which provably never achieves p^3). As $p = 3, 5$ fails considering the $p = 7$ polynomial gives

$$(x+1)^7 - x^7 - 1 = 7x(x+1)\left(x^4 + 2x^3 + 3x^2 + 2x + 1\right).$$

The key is now to notice that the last factor is $(x^2 + x + 1)^2$, which suggests the entire solution.

In fact, even if $p \equiv 2 \pmod 3$, the polynomial $x^2 + x + 1$ still divides $(x+1)^p - x^p - 1$. So even the $p = 5$ case can motivate the main idea.

Solution 191. The following solution was found by user `Aiscrim` on AOPS.

Obviously $|M| = \infty$. Assume for contradiction $p \notin M$. We say a prime $q \in M$ is *sparse* if there are only finitely many elements of M which are $q \pmod p$ (in particular there are finitely many sparse primes).

Now let C be the product of all sparse primes (note $p \nmid C$). First, set $a_0 = 1$. For $k \geq 0$, consider then the prime factorization of the number

$$Ca_k + 1.$$

No prime in its factorization is sparse, so consider the number a_{k+1} obtained by **replacing each prime in its factorization with some arbitrary representative of that prime's residue class**. In this way we select a number a_{k+1} such that

- $a_{k+1} \equiv Ca_k + 1 \pmod{p}$, and
- a_{k+1} is a product of distinct primes in M.

In particular, $a_k \equiv C^k + C^{k-1} + \cdots + 1 \pmod{p}$

But since $C \not\equiv 0 \pmod{p}$, we can find a k such that $a_k \equiv 0 \pmod{p}$ (namely, $k = p-1$ if $C \equiv 1$ and $k = p-2$ else) which is clearly impossible since a_k is a product of primes in M!

Solution 192. Call a function g *valiant* if it obeys this condition. We claim the answer is all numbers relatively prime to 101!.

The construction is to just let g be the identity function.

Before proceeding to the converse solution, we make a long motivational remark.

Remark (Motivation for both parts)**.** The following solution is dense, and it is easier to think about some small cases first, to motivate the ideas. We consider the result where 101 is replaced by 2 or 3.

- If we replaced 101 with 2, you can show $2 \nmid n$ easily: write
$$\sum_x x \equiv \sum_x g(x) \equiv \sum_x (g(x) + x) \pmod{n}$$
which implies
$$0 \equiv \sum_x x = \frac{1}{2} n(n+1) \pmod{n}$$
which means $\frac{1}{2}n(n+1) \equiv 0 \pmod{n}$, hence n odd.
- If we replaced 101 with 3, then you can try a similar approach using squares, since
$$\begin{aligned} 0 &\equiv \sum_x \left[(g(x) + 2x)^2 - 2\left(g(x) + x\right)^2 + g(x)^2 \right] \pmod{n} \\ &= \sum_x 2x^2 = 2 \cdot \frac{n(n+1)(2n+1)}{6} \end{aligned}$$
which is enough to force $3 \nmid n$.

We now present several different proofs of the converse, all of which generalize the ideas contained here. In everything that follows we assume $n > 1$ for convenience.

First solution (original one) The proof is split into two essentially orthogonal claims, which we state as lemmas.

Lemma (Lemma I: elimination of g)**.** *Assume valiant $g\colon \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ exists. Then*

$$k! \sum_{x \in \mathbb{Z}/n\mathbb{Z}} x^k \equiv 0 \pmod{n}$$

for $k = 0, 1, \dots, 100$.

Proof. Define $g_x(T) = g(x) + Tx$ for any integer T. If we view $g_x(T)^k$ as a polynomial in $\mathbb{Z}[T]$ of degree k with leading coefficient x^k, then taking the kth finite difference implies that, for any x,

$$k!x^k = \binom{k}{0} g_x(k)^k - \binom{k}{1} g_x(k-1)^k + \binom{k}{2} g_x(k-2)^k - \cdots + (-1)^k \binom{k}{k} g_x(0)^k.$$

On the other hand, for any $1 \le k \le 100$ we should have

$$\begin{aligned} \sum_x g_x(0)^k \equiv \sum_x g_x(1)^k &\equiv \cdots \equiv \sum_x g_x(k)^k \\ &\equiv S_k \coloneqq 0^k + \cdots + (n-1)^k \pmod{n} \end{aligned}$$

by the hypothesis. Thus we find

$$k! \sum_x x^k \equiv \left[\binom{k}{0} - \binom{k}{1} + \binom{k}{2} - \cdots \right] S_k \equiv 0 \pmod{n}$$

for any $1 \le k \le 100$, but also obviously for $k = 0$. □

We now prove the following self-contained lemma.

Lemma (Lemma II: power sum calculation)**.** *Let p be a prime, and let n, M be positive integers such that*

$$M \quad \textit{divides} \quad 1^k + 2^k + \cdots + n^k$$

for $k = 0, 1, \dots, p-1$. If $p \mid n$ then $\nu_p(M) < \nu_p(n)$.

Proof. The hypothesis means that that any polynomial $f(T) \in \mathbb{Z}[T]$ with $\deg f \le p-1$ will have $\sum_{x=1}^n f(x) \equiv 0 \pmod{M}$. In particular, we have

$$\begin{aligned} 0 &\equiv \sum_{x=1}^n (x-1)(x-2)\cdots(x-(p-1)) \\ &= (p-1)! \sum_{x=1}^n \binom{x-1}{p-1} = (p-1)! \binom{n}{p} \pmod{M}. \end{aligned}$$

But now $\nu_p(M) \le \nu_p(\binom{n}{p}) = \nu_p(n) - 1$. □

Now assume for contradiction that valiant $g\colon \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ exists, and $p \le 101$ is the *smallest* prime dividing n. Lemma I implies that $k! \sum_x x^k \equiv 0 \pmod{n}$ for $k = 1, \dots, p-1$ and hence $\sum_x x^k \equiv 0 \pmod{n}$ too. Thus $M = n$ holds in the previous lemma, impossible.

A second solution Both lemmas above admit variations where we focus on working modulo p^e rather than working modulo n.

Lemma (Lemma I'). *Assume valiant $g\colon \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ exists. Let $p \le 101$ be a prime, and $e = \nu_p(n)$. Then*

$$\sum_{x \in \mathbb{Z}/n\mathbb{Z}} x^k \equiv 0 \pmod{p^e}$$

for $k = 0, 1, \dots, p-1$.

Proof. This is weaker than Lemma I, but we give an independent specialized proof. Begin by writing

$$\sum_x (g(x) + Tx)^k \equiv \sum_x x^k \pmod{p^e}.$$

Both sides are integer polynomials in T, which vanish at $T = 0, 1, \dots, p-1$ by hypothesis (since $p - 1 \le 100$).

We now prove the following more general fact: if $f(T) \in \mathbb{Z}[T]$ is an integer polynomial with $\deg f \le p-1$, such that $f(0) \equiv \cdots \equiv f(p-1) \equiv 0 \pmod{p^e}$, then all coefficients of f are divisible by p^e. The proof is by induction on $e \ge 1$. When $e = 1$, this is just the assertion that the polynomial has at most $\deg f$ roots modulo p. When $e \ge 2$, we note that the previous result implies all coefficients are divisible by p, and then we divide all coefficients by p.

Applied here, we have that all coefficients of

$$f(T) \coloneqq \sum_x (g(x) + Tx)^k - \sum_x x^k$$

are divisible by p^e. The leading T^k coefficient is $\sum_k x^k$ as desired. □

Lemma (Lemma II'). *If $e \ge 1$ is an integer, and p is a prime, then*

$$\nu_p\left(1^{p-1} + 2^{p-1} + \cdots + (p^e - 1)^{p-1}\right) = e - 1.$$

Proof. First, note that the cases where $p = 2$ or $e = 1$ are easy; since if $p = 2$ we have $\sum_{x=0}^{2^e-1} x \equiv 2^{e-1}(2^e - 1) \equiv -2^{e-1} \pmod{2^e}$, while if $e = 1$ we have $1^{p-1} + \cdots + (p-1)^{p-1} \equiv -1 \pmod{p}$. Henceforth assume that $p > 2$, $e > 1$.

Let g be an integer which is a primitive root modulo p^e. Then, we can sum the terms which are relatively prime to p as

$$S_0 \coloneqq \sum_{\gcd(x,p)=1} x^{p-1} \equiv \sum_{i=1}^{\varphi(p^e)} g^{(p-1)\cdot i} \equiv \frac{g^{p^{e-1}(p-1)^2} - 1}{g^{p-1} - 1} \pmod{p^e}$$

which implies $\nu_p(S_0) = e - 1$, by lifting the exponent. More generally, for $r \ge 1$ we may set

$$S_r \coloneqq \sum_{\nu_p(x)=r} x^{p-1} \equiv (p^r)^{p-1} \sum_{i=1}^{\varphi(p^{e-r})} g_r^{(p-1)\cdot i} \pmod{p^e}$$

where g_r is a primitive root modulo p^{e-r}. Repeating the exponent-lifting calculation shows that $\nu_p(S_r) = r(p-1) + ((e-r)-1) > e$, as needed. □

Assume to the contrary that $p \le 101$ is a prime dividing n, and a valiant $g\colon \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ exists. Take $k = p-1$ in Lemma I' to contradict Lemma II'

A third remixed solution We use Lemma I and Lemma II' from before. As before, assume $g\colon \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ is valiant, and n has a prime divisor $p \le 101$. Also, let $e = \nu_p(n)$.

Then $(p-1)! \sum_x x^{p-1} \equiv 0 \pmod{n}$ by Lemma I, and now

$$\begin{aligned} 0 &\equiv \sum_x x^{p-1} \pmod{p^e} \\ &\equiv \frac{n}{p^e} \sum_{x=1}^{p^e-1} x^{p-1} \not\equiv 0 \pmod{p^e} \end{aligned}$$

by Lemma II', contradiction.

A fourth remixed solution We also can combine Lemma I' and Lemma II. As before, assume $g\colon \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ is valiant, and let p be the smallest prime divisor of n.

Assume for contradiction $p \le 101$. By Lemma I' we have

$$\sum_x x^k \equiv 0 \pmod{p^e}$$

for $k = 0, \dots, p-1$. This directly contradicts Lemma II with $M = p^e$.

List of problems and examples

Index of named problems

Bibliography

[Che16] Evan Chen. *Euclidean geometry in mathematical olympiads*. MAA Problem Books Series. Mathematical Association of America, Washington, DC, 2016, pp. xv+311.

[Che19] Evan Chen. *An Infinitely Large Napkin*. 2019. URL: http://web.evanchen.cc/napkin.html.

[Ste16] Justin Stevens. *Olympiad Number Theory through Challenging Problems*. 2016. URL: https://numbertheoryguy.com/publications/olympiad-number-theory-book/.

Acknowledgements

Thanks to everyone who read through drafts of this text, locating errors and providing valuable suggestions. All remaining mistakes are the author's responsibility. I'd like to acknowledge in particular the following individuals who provided suggestions and corrections on early versions of this text:

- Ankan Bhattacharya
- Daniel Sheremeta
- Derek Liu
- Gopal Goel
- Jeffrey Kwan
- Karen Ge
- Mason Fang
- Niyanth Rao
- William Yue
- Zack Chroman

Finally, one final thanks to all my students over the years, past or present, without whom these excerpts would not exist.