

# Engineering

Carl Joshua Quines

March 30, 2020

## 1 What is engineering?

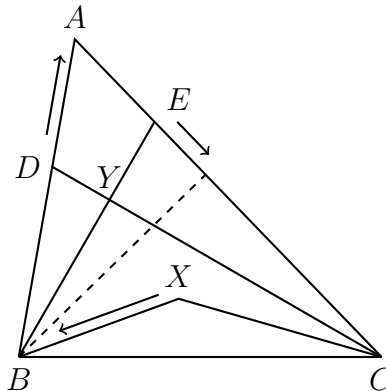
**Engineering is finding the answer to a problem without solving it.** It's the art of solving without solving, of finding the answer with less effort than it would take us if we did the problem "properly". The following techniques are engineering:

- *Engineer's induction*: finding the answer for some cases, then generalizing.
- Carefully drawing a diagram and measuring the answer.
- Using extra degrees of freedom to find a numerical answer.
- Substituting choices of a multiple-choice problem to solve it.
- Using the answer format to assist in any of the above techniques.

**Example 1.1** (PMO 2019 Qualifying III.4). In triangle  $ABC$ ,  $D$  and  $E$  are points on sides  $AB$  and  $AC$  respectively, such that  $BE$  is perpendicular to  $CD$ . Let  $X$  be a point inside the triangle such that  $\angle XBC = \angle EBA$  and  $\angle XCB = \angle DCA$ . If  $\angle A = 54^\circ$ , what is the measure of  $\angle EXD$ ?

It's possible to carefully draw a diagram and then measure it to get the answer. You'll get a measurement of about  $35^\circ$ . The nearest angle that's a nice angle would be  $36^\circ$ , which just happens to be  $90^\circ - 54^\circ$ , so we guess  $36^\circ$  and move on. This is engineering, but it's risky.

Instead, here's some better engineering: **abuse degrees of freedom**. The answer is a number, so it must be the same no matter where the points are positioned. Knowing this, we take the special case when  $D$  is very very close to  $A$ .



Then  $E$  would be very close the foot of the perpendicular from  $B$  to  $AC$ , and  $X$  would be very close to  $B$ . Then  $\angle EXD$  would be about  $\angle EBA$ . But triangle  $EBA$  is right, so  $\angle EBA = 90^\circ - \angle EAB = 36^\circ$ , which is indeed the correct answer.

Some important philosophical notes. I think engineering is a useful skill. It leads to making exciting conjectures, and involves things like estimation that are important outside of competitions. There are certainly lots of people who should use it more, but I also think that there are people who should use it *less*.

**Engineering is not a replacement for learning math.** It shouldn't be *satisfying* to get an answer with engineering. Engineering a problem is like taking out a loan, that you need to repay with learning the *actual* solution to the problem.

With that in mind, I think **engineering is a skill that's best learned when you're more experienced with math**, rather than something you learn when you're starting out. Intuition is something that you need when engineering anyway, as I'll emphasize throughout the article, and intuition is something you pick up over time. This handout can do more harm than good, and don't say I didn't warn you.

## 2 Engineer's induction

*Engineering* takes its name from *engineer's induction*, which is the primary engineering technique. In its simplest form, engineer's induction is about guessing the formula for a sequence, and then convincing yourself that it's true based on some small cases. So the simplest kind of engineer's induction is used in problems where there's a sequence, and you have to guess the formula:

**Example 2.1** (PUMaC Algebra A 2018/2). If  $a_1, a_2, \dots$  is a sequence of real numbers such that for all  $n$ ,  $\sum_{k=1}^n a_k \left(\frac{k}{n}\right)^2 = 1$ , find the smallest  $n$  such that  $a_n < \frac{1}{2018}$ .

We find the first few terms. Substituting  $n = 1$ , we get

$$a_1 \left(\frac{1}{1}\right)^2 = 1 \implies a_1 = 1.$$

Then substituting  $n = 2$  gives

$$a_1 \left(\frac{1}{2}\right)^2 + a_2 \left(\frac{2}{2}\right)^2 = 1 \implies a_2 = \frac{3}{4}.$$

We finally substitute  $n = 3$  to get

$$a_1 \left(\frac{1}{3}\right)^2 + a_2 \left(\frac{2}{3}\right)^2 + a_3 \left(\frac{3}{3}\right)^2 = 1 \implies a_3 = \frac{5}{9}.$$

Looking at the numerators, we have 1, 3, 5, which looks like the odd numbers. Looking at the denominators, we have 1, 4, 9, which looks like the perfect squares. So we would guess that the formula is

$$a_n = \frac{2n-1}{n^2}.$$

In this case, it's not hard to prove the formula through actual induction. (If you don't immediately see how to do this, you should **stop reading** this handout.) From having written lots of inductions before, we see *why* this is true, and that's what engineer's induction is about—convincing yourself that something is true using examples, rather than a proof.

**Exercise 2.2.** We could always check another case to be sure. Do something similar to what we did to find  $a_4$ , and verify that it matches our guess for the formula.

To finish the rest of the problem, we need the least  $n$  such that

$$\frac{2n-1}{n^2} < \frac{1}{2018}.$$

The fraction is approximately  $\frac{2n}{n^2}$ , or  $\frac{2}{n}$ . In this case,  $n = 4036$  would make the fraction equal to  $\frac{1}{2018}$ . But  $2n$  is larger than  $2n-1$ , so the actual fraction would be smaller. So  $a_{4036}$  is indeed smaller than  $\frac{1}{2018}$ .

Is the answer 4035? Let's check whether  $n = 4035$  would make the fraction smaller than  $\frac{1}{2018}$ . Here, let's replace 2018 with  $\frac{n+1}{2}$ :

$$\frac{2n-1}{n^2} < \frac{1}{\frac{n+1}{2}} \iff (2n-1)(n+1) < 2n^2,$$

The left-hand side is larger than the right-hand side, so  $a_{4035} > \frac{1}{2018}$ . So  $a_{4035}$  is larger, which means that  $a_{4036}$  is the first term smaller than  $\frac{1}{2018}$ , and the answer is 4036.

### Exercises

**Problem 2.3** (AMC 12A 2019/9). A sequence of numbers is defined by  $a_1 = 1$ ,  $a_2 = \frac{3}{7}$ , and

$$a_n = \frac{a_{n-2} \cdot a_{n-1}}{2a_{n-2} - a_{n-1}}$$

for all  $n \geq 3$ . Then  $a_{2019}$  can be written as  $\frac{p}{q}$ , where  $p$  and  $q$  are relatively prime positive integers. What is  $p+q$ ?

**Problem 2.4** (ARML Team 2017/6). Let  $\{a_n\}$  be a sequence with  $a_0 = 1$ , and for all  $n > 0$ ,

$$a_n = \frac{1}{2} \sum_{i=0}^{n-1} a_i.$$

Compute the greatest value of  $n$  for which  $a_n < 2017$ . **Hint:** 9

## 2.1 Trying small cases

Engineer's induction can also be applied even if the problem doesn't have a sequence. In this case, the sequence is made by trying small cases of the problem, guessing the formula, and then applying it to the original problem. Here's an example:

**Example 2.5** (OMO Spring 2018/5). A mouse has a wheel of cheese which is cut into 2018 slices. The mouse also has a 2019-sided die, with faces labeled  $0, 1, 2, \dots, 2018$ , and with each face equally likely to come up. Every second, the mouse rolls the dice. If the dice lands on  $k$ , and the mouse has at least  $k$  slices of cheese remaining, then the mouse eats  $k$  slices of cheese; otherwise, the mouse does nothing. What is the expected number of seconds until all the cheese is gone?

The 2018 here is probably an arbitrary choice, and it's asking us to replace it with a smaller number. Let's let this number be  $n$ , and we're trying to look for the answer in terms of  $n$ . That is, we have this new, restated problem:

**Example 2.6.** A mouse has a wheel of cheese which is cut into  $n$  slices. The mouse also has a  $n + 1$ -sided die, with faces labeled  $0, 1, 2, \dots, n$ , and with each face equally likely to come up. Every second, the mouse rolls the dice. If the dice lands on  $k$ , and the mouse has at least  $k$  slices of cheese remaining, then the mouse eats  $k$  slices of cheese; otherwise, the mouse does nothing. What is the expected number of seconds until all the cheese is gone?

Let's call the answer to this problem  $S_n$ . We're looking for  $S_{2018}$ . We will try to guess the formula for  $S_n$  by finding it for small values of  $n$ .

When  $n = 1$ , then the die only has two faces, 0 and 1. It takes exactly  $s$  seconds until all of the cheese is gone only if the first  $s - 1$  rolls are all 0s, and the  $s$ th roll is 1. This happens with probability  $\frac{1}{2^s}$ . So the expected number of seconds is

$$S_1 = 1 \cdot \frac{1}{2^1} + 2 \cdot \frac{1}{2^2} + 3 \cdot \frac{1}{2^3} + \dots$$

To calculate this, we multiply both sides by  $\frac{1}{2}$  and subtract it from  $S_1$ , giving us

$$\begin{aligned} \frac{S_1}{2} &= 1 \cdot \frac{1}{2^2} + 2 \cdot \frac{1}{2^3} + 3 \cdot \frac{1}{2^4} + \dots \\ S_1 - \frac{S_1}{2} &= 1 \cdot \frac{1}{2^1} + (2 - 1) \frac{1}{2^2} + (3 - 2) \frac{1}{2^3} + \dots \\ \frac{S_1}{2} &= \frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1, \end{aligned}$$

which means that  $S_1 = 2$ .

When  $n = 2$ , the problem is slightly harder, but it's still doable. When does it take exactly  $s$  seconds for all of the cheese to be eaten? Either:

- Both slices are eaten at the same time, which happens if the first  $s - 1$  rolls are 0 and the  $s$ th roll is 2. This happens with probability  $\frac{1}{3^s}$ .
- The slices are eaten one at a time. There's some number  $t$  such that the first  $t - 1$  rolls are 0, and then the  $t$ th roll is 1. Then the next  $s - t - 1$  rolls can be 0 or 2, and the  $s$ th roll is 1 again. As  $t$  can be anything from 1 to  $s - 1$ , this happens with probability

$$\left(\frac{1}{3}\right) \left(\frac{2}{3}\right)^{s-2} \left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^{s-3} \left(\frac{1}{3}\right) + \dots + \left(\frac{1}{3}\right)^{s-1} \left(\frac{1}{3}\right) = \frac{2^{s-1} - 1}{3^s}.$$

We can compute this series by noticing that the denominators are all  $3^s$ , while the numerators form a geometric series.

In total, the probability that it takes exactly  $s$  seconds for all the cheese to be eaten is  $\frac{2^{s-1}}{3^s}$ . To compute the expected number of seconds, we do a similar trick on

$$S_2 = 1 \cdot \frac{2^0}{3^1} + 2 \cdot \frac{2^1}{3^2} + 3 \cdot \frac{2^2}{3^3} + \dots$$

We multiply the whole series by  $\frac{2}{3}$  and then subtract it from the original series. This allows us to solve for  $S_2 = 3$ .

So these two answers allow us to make a “sequence” of answers for different  $n$ . The sequence  $S_n$  begins as 2, 3. A good guess for the formula would be  $S_n = n + 1$ , which means that we’d guess  $S_{2018} = 2019$ .

If we really wanted to be sure, we could find  $S_3$ . However, it’s already getting a bit tedious at this point, since we have four cases. Instead, we can guess that the probability that it takes exactly  $s$  seconds for all of the cheese to be eaten is  $\frac{3^{s-1}}{4^s}$ .

**Exercise 2.7.** Let’s try to find  $S_3$  without going through all four cases.

- (a) We guess it takes exactly  $s$  seconds for all of the cheese to be eaten is  $\frac{3^{s-1}}{4^s}$ . Verify that this is true for  $s = 1$  and  $s = 2$ . The fact that we have a simple formula, and that it’s true for  $s = 1$  and  $s = 2$ , should make us believe that the formula is true in general.

- (b) Using a similar technique to what we did earlier, check that the sum

$$S_3 = 1 \cdot \frac{3^0}{4^1} + 2 \cdot \frac{3^1}{4^2} + 3 \cdot \frac{3^2}{4^3} + \dots$$

is indeed equal to 4.

At this point, we’ve seen that  $S_n = n + 1$  for  $n = 1, 2, 3$ , so we should be pretty confident that  $S_{2018} = 2019$ . And this is indeed the correct answer!

Think about this problem. The way we found the answer involved a lot of computation, but **very few tricks were used in our technique**. Solutions that use states require knowing how to set up recursions, and the short intended solution is not easy to come up with. Engineer’s induction may not always give the answer in a fast or elegant way, but it’s often a method that requires knowing the least tricks.

Here’s a different kind of example, taken from the example in the previous section:

**Example 2.8.** Find the least positive integer  $n$  such that

$$\frac{2n - 1}{n^2} < \frac{1}{2018}.$$

Again, 2018 is an arbitrary number, so let’s replace it with small numbers and try to guess the general answer. When we replace it with 1, we get

$$\frac{2n - 1}{n^2} < 1,$$

and the least  $n$  that works is 2. When we replace it with 2, we get

$$\frac{2n - 1}{n^2} < \frac{1}{2},$$

and the least  $n$  that works is 4. And when we replace it with 3, we find that the least  $n$  that works is 6. It always looks to be double the number we’re replacing. So we would guess that in the original problem, with 2018, the least  $n$  that works is 4036, which is correct.

### Exercises

**Problem 2.9** (PUMaC Combinatorics A 2017/3). There is a box containing 100 balls, each of which is either orange or black. The box is equally likely to contain any number of black balls between 0 and 100, inclusive. A random black ball rolls out of the box. The probability that the next ball to roll out of the box is also black can be written in the form  $\frac{p}{q}$  where  $p$  and  $q$  are relatively prime positive integers. Find  $p + q$ .

**Problem 2.10** (PUMaC Combinatorics A 2016/4). A knight is placed at the origin of the Cartesian plane. Each turn, the knight moves in an L-shape (2 units parallel to one axis and 1 unit parallel to the other) to one of eight possible locations, chosen at random. After 2016 such turns, what is the expected value of the square of the distance of the knight from the origin?

**Problem 2.11** (HMMT Guts 2017/20). For positive integers  $a$  and  $N$ , let  $r(a, N)$ , a number in  $\{0, 1, \dots, N - 1\}$  denote the remainder of  $a$  when divided by  $N$ . Determine the number of positive integers  $n \leq 1000000$  for which  $r(n, 1000) > r(n, 1001)$ . **Hint: 13**

#### 2.1.1 Be fooled not by small numbers

Here's one important consideration when considering when to apply engineer's induction. We might only reach for it when the numbers in the problem are large, like the current year, or 1000 or something. But sometimes, engineer's induction can be useful for small numbers too, and we shouldn't dismiss using it just because the given numbers are small. **We can always make small numbers smaller.**

**Example 2.12.** Eight points are equally spaced on a unit circle. What is the product of the distances from one point to each other point?

Here, instead of tackling the problem directly, we use engineer's induction, replacing eight with smaller numbers:

- If there are only two points, then there is only one distance, 2, so the answer is 2.
- If there are three points, then this forms an equilateral triangle with height  $\frac{3}{2}$ , and by constructing a 30-60-90 triangle, we see that it must have side length  $\sqrt{3}$ . So the answer is  $\sqrt{3} \cdot \sqrt{3}$ , which is 3.
- If there are four points, then this forms a square. The diagonal is 2, and by drawing both diagonals, we form a 45-45-90 triangle, which gives us the side length of  $\sqrt{2}$ . Then the answer is  $\sqrt{2} \cdot 2 \cdot \sqrt{2}$ , or 4.

This strongly suggests that for eight points, the answer is 8. We can further check this by checking the case of the hexagon:

**Exercise 2.13.** Check that the answer is 6 for a regular hexagon as well. (Two of the distances are part of an equilateral triangle, so we've done most of the work.)

And 8 is indeed the correct answer, which we found through a much simpler technique than the intended solution of using complex numbers or trigonometry.

## Exercises

**Problem 2.14** (PMO 2017 Areas I.18). A railway passes through four towns  $A, B, C,$  and  $D,$  in that order. The railway forms a complete loop and trains go in both directions. Suppose that a trip between two adjacent towns costs one ticket. Using exactly eight tickets, how many distinct ways are there of travelling from town  $A$  and ending at town  $A$ ? (Note that passing through  $A$  somewhere in the middle of the trip is allowed.)

**Problem 2.15** (OMO Spring 2014/9). Eighteen students participate in a team selection test with three problems, each worth up to seven points. All scores are nonnegative integers. After the competition, the results are posted by Evan in a table with 3 columns: the student's name, score, and rank (allowing ties), respectively. Here, a student's rank is one greater than the number of students with strictly higher scores (for example, if seven students score 0, 0, 7, 8, 8, 14, 21 then their ranks would be 6, 6, 5, 3, 3, 2, 1 respectively).

When Richard comes by to read the results, he accidentally reads the rank column as the score column and vice versa. Coincidentally, the results still made sense! If the scores of the students were  $x_1 \leq x_2 \leq \dots \leq x_{18},$  determine the number of possible values the 18-tuple  $(x_1, x_2, \dots, x_{18}).$  In other words, determine the number of possible multisets (sets with repetition) of scores.

**Problem 2.16** (PUMAC Algebra B 2014/5). Given that

$$a_n a_{n-2} - a_{n-1}^2 + a_n - n a_{n-2} = -n^2 + 3n - 1$$

and  $a_0 = 1, a_1 = 3,$  find  $a_{20}.$  **Hint:** 14

**Problem 2.17** (HMMT Algebra 2013/7). Compute

$$\sum_{a_1=0}^{\infty} \sum_{a_2=0}^{\infty} \dots \sum_{a_7=0}^{\infty} \frac{a_1 + a_2 + \dots + a_7}{3^{a_1 + a_2 + \dots + a_7}}.$$

## 2.2 Choosing what to induct on

Another thing to think about when doing engineer's induction is how *exactly* we induct. What exactly do we change? How exactly are we going to make the sequence? Sometimes it's clear, but sometimes need to be more creative:

**Example 2.18** (AMC 12B 2014/23). The number 2017 is prime. Let  $S = \sum_{k=0}^{62} \binom{2014}{k}.$

What is the remainder when  $S$  is divided by 2017?

Here, there are several choices of what to try. We can see that 2014 and 2017 are related, so we might consider replacing the two numbers, with something like 4 and 7. But then, how are both numbers related to 62? Indeed, the number that we want to change isn't 2014 and 2017, but the 62. Since it doesn't really matter, we should consider what happens when we let it vary.

In particular, let's try finding  $\binom{2014}{k}$  modulo 2017 for small values of  $k:$

- When  $k = 0,$  it's just 1.
- When  $k = 1,$  it's 2014, which would be nicer if we wrote it as  $-3.$
- When  $k = 2,$  it's

$$\frac{2014 \cdot 2013}{2} \equiv \frac{(-3)(-4)}{2} \equiv 6.$$

At this point I'd say that the pattern is strong enough to guess:  $1, -3, 6$  are the triangular numbers, except that the signs alternate.

**Exercise 2.19.** Verify that the pattern continues for  $k = 3$  by using the same trick; you should get  $-10$ .

We find the sums of the first few terms, and we get  $1, -2, 4, -6, 9$ . We can then guess the positive terms are the perfect squares, so the sum from  $k = 0$  to  $k = 62$  would give  $32^2$ , or  $1024$ .

### Exercises

**Problem 2.20** (AMC 12A 2014/16). The product  $(8)(888\dots 8)$ , where the second factor has  $k$  digits, is an integer whose digits have a sum of  $1000$ . What is  $k$ ?

**Problem 2.21** (AIME I 2018/1). Let  $S$  be the number of ordered pairs of integers  $(a, b)$  with  $1 \leq a \leq 100$  and  $b \geq 0$  such that the polynomial  $x^2 + ax + b$  can be factored into the product of two (not necessarily distinct) linear factors with integer coefficients. Find the remainder when  $S$  is divided by  $1000$ .

**Problem 2.22** (AMC 12A 2014/23). The fraction  $\frac{1}{99^2} = 0.\overline{b_{n-1}b_{n-2}\dots b_2b_1b_0}$ , where  $n$  is the length of the period of the repeating decimal expansion. What is the sum  $b_0 + b_1 + \dots + b_{n-1}$ ?  
Hint: 7

### 2.2.1 Two free variables

Sometimes, we may want to have *two* free variables instead of just one. This is slightly harder than only one free variable, because there are way, *way* more possible formulas. One of the tricks is to **only change one variable at a time**.

**Example 2.23** (HMMT Combinatorics 2017/6). Emily starts with an empty bucket. Every second, she either adds a stone or removes a stone from the bucket, each with probability  $\frac{1}{2}$ . If she wants to remove a stone from the bucket and the bucket is currently empty, she does nothing for that second (still with probability  $\frac{1}{2}$ ). What is the probability that after 2017 seconds her bucket contains exactly 1337 stones?

Here, we have two clear numbers that we want to do engineer's induction on, 2017 and 1337. Let's call the number of seconds  $s$ . We can just find the answer for all small  $s$ , since there are only a limited number of possible stones. When  $s = 1$ , we have 0 stones or 1 stone with probability  $\frac{1}{2}$ .

In the former case, half the time we get 0 stones and half the time we get 1 stone. In the latter case, half the time we get 0 stones and the other half we get 2 stones. So in total, for  $s = 2$ , we have 0 stones with probability  $\frac{2}{4}$ , 1 stone with  $\frac{1}{4}$ , and 2 stones with  $\frac{1}{4}$ .

Again, we can use these numbers to find the results for  $s = 3$ . It's 0 stones with probability  $\frac{3}{8}$ , 1 stone with  $\frac{3}{8}$ , 2 stones with  $\frac{1}{8}$ , and 3 stones with  $\frac{1}{8}$ . Now the numerators, 3, 3, 1, 1, are familiar—they're the binomial coefficients of the form  $\binom{3}{x}$ !



**Exercise 2.24.** Verify that for  $s = 4$ , the probabilities for getting  $0, 1, \dots, 5$  stones are

$$\frac{6}{2^4}, \frac{4}{2^4}, \frac{4}{2^4}, \frac{1}{2^4}, \frac{1}{2^4}.$$

Can you guess the formula for the probability of having  $x$  stones after  $s$  seconds?

From here, we see that the answer should be  $\frac{\binom{2017}{340}}{2^{2017}}$ .

### Exercises

**Problem 2.25** (SMT Discrete 2018/7). Let  $S$  be the set of all 1000 element subsets of the set  $\{1, 2, 3, \dots, 2018\}$ . What is the expected value of the minimum element of a set chosen uniformly at random from  $S$ ?

**Problem 2.26** (HMMT Guts 2019/18). 2019 points are chosen independently and uniformly at random on the interval  $[0, 1]$ . Tairitsu picks 1000 of them randomly and colors them black, leaving the remaining ones white. Hikari then computes the sum of the positions of the leftmost white point and the rightmost black point. What is the probability that this sum is at most 1?

**Problem 2.27** (HMMT Combinatorics 2013/9). Given a permutation  $\sigma$  of  $\{1, 2, \dots, 2013\}$ , let  $f(\sigma)$  be the number of fixed points of  $\sigma$  – that is, the number of  $k \in \{1, 2, \dots, 2013\}$  such that  $\sigma(k) = k$ . If  $S$  is the set of all possible permutations  $\sigma$ , compute

$$\sum_{\sigma \in S} f(\sigma)^4.$$

(A permutation  $\sigma$  is a bijective mapping from  $\{1, 2, \dots, 2013\}$  to  $\{1, 2, \dots, 2013\}$ .) **Hint: 4**

## 2.2.2 Trigonometric sums and products

This is a particular kind of engineer's induction that I've seen often enough that it gets its own section. The key idea that I want to emphasize is that **degrees are arbitrary**, and that every time you see a trigonometric sum or product with degrees, you should be thinking about using engineer's induction.

**Example 2.28** (Titu). Evaluate

$$\frac{1}{\cot 9^\circ - 3 \tan 9^\circ} + \frac{3}{\cot 27^\circ - 3 \tan 27^\circ} + \frac{9}{\cot 81^\circ - 3 \tan 81^\circ} + \frac{27}{\cot 243^\circ - 3 \tan 243^\circ}.$$

First, we recognize that the degrees are arbitrary, and probably the only important part is that it triples each time. Replace  $9^\circ$  with  $x$  to get

$$\frac{1}{\cot x - 3 \tan x} + \frac{3}{\cot 3x - 3 \tan 3x} + \frac{9}{\cot 9x - 3 \tan 9x} + \frac{27}{\cot 27x - 3 \tan 27x}.$$

Now we can try different values of  $x$  and see what we get. Convenient ones to try are  $x = 45^\circ$  and  $x = 135^\circ$ . Let's try the first one. This gives us

$$\frac{1}{1-3} + \frac{3}{-1+3} + \frac{9}{1-3} + \frac{27}{-1+3} = \frac{-1+3-9+27}{2} = 10.$$

Similarly, check that if we try  $x = 135^\circ$ , we get  $-10$ .

Now, we know the answer should depend on  $x$  in some way, most likely  $\tan x$  or  $\cot x$ . Can it be something like  $10(\tan x)^2$ ? If it was, then the answer for  $x = 45^\circ$  and  $x = 135^\circ$  would be the same, but it isn't.

So we expect it to be in terms of either  $\tan x$  or  $\cot x$ . The nicest option would be either something like  $10 \tan x$  or  $10 \cot x$ , which matches what we have. What we've done so far can't distinguish between the two of them, but we can use size to figure out which one:

**Exercise 2.29.** Say that  $x$  was very small. Then  $\tan x$  would be very small, and  $\cot x$  would be very large. Convince yourself that, in this case, the sum is closer to  $10 \tan x$  than it is to  $10 \cot x$ .

So we have a good guess of  $10 \tan x$ . What we'd really like is if we had another  $x$  to try it on. Unfortunately, neither  $0$  nor  $30^\circ$  nor  $60^\circ$  give anything defined. Instead, we can use  $15^\circ$ .

**Exercise 2.30.** Recall (or derive) that  $\tan 15^\circ = 2 - \sqrt{3}$ . Then check that for  $x = 15^\circ$ , the above sum really is  $10 \tan x$ .

And indeed, the answer really is  $10 \tan 9^\circ$ . In short-answer type contests, the answer format is typically given, which makes things a bit easier:

**Example 2.31** (AIME II 2000/15). Find the least positive integer  $n$  such that

$$\frac{1}{\sin 45^\circ \sin 46^\circ} + \frac{1}{\sin 47^\circ \sin 48^\circ} + \cdots + \frac{1}{\sin 133^\circ \sin 134^\circ} = \frac{1}{\sin n^\circ}.$$

The denominators form this arithmetic sequence  $45^\circ, 46^\circ, \dots, 134^\circ$ , and the next term would have been  $135^\circ$ . Again, the common difference being  $1^\circ$  is definitely arbitrary—the important facts seem to be that it starts with  $45^\circ$ , ends with  $135^\circ$ , and has an odd number of terms. And also, the  $n^\circ$  on the right-hand side is probably just  $n$  times the common difference.

If the sum had only one term, the sequence would be  $45^\circ, 90^\circ, 135^\circ$ . And then the sum would be

$$\frac{1}{\sin 45^\circ \sin 90^\circ} = \frac{1}{\sin 45n^\circ} \implies n = 1.$$

If we were unsure, we could see what would happen if the sum had more terms. One choice would be the sequence  $45^\circ, 60^\circ, 75^\circ, 90^\circ, 105^\circ, 120^\circ, 135^\circ$ , which would give us the sum

$$\frac{1}{\sin 45^\circ \sin 60^\circ} + \frac{1}{\sin 75^\circ \sin 90^\circ} + \frac{1}{\sin 105^\circ \sin 120^\circ} = \frac{1}{\sin 15n^\circ}$$

**Exercise 2.32.** Evaluate the above sum, and show that  $n = 1$  here as well. It may be easier to note that  $\sin 75^\circ = \sin 105^\circ = \cos 15^\circ$ , not evaluate these until the end, and then use the earlier fact that  $\tan 15^\circ = 2 - \sqrt{3}$ .

Again, if we *really* wanted to be sure, we can also check it for when the sum has two terms. But at this point we should be pretty confident that the answer is just 1, which is correct.

### Exercises

**Problem 2.33.** Find the value of  $\sin\left(\frac{\pi}{2020}\right)\sin\left(\frac{2\pi}{2020}\right)\cdots\sin\left(\frac{2019\pi}{2020}\right)$ .

**Problem 2.34** (HMMT February Guts 2018/17). Find the value of

$$\frac{\cos 30.5^\circ + \cos 31.5^\circ + \cdots + \cos 44.5^\circ}{\sin 30.5^\circ + \sin 31.5^\circ + \cdots + \sin 44.5^\circ}.$$

**Problem 2.35** (AIME I 2015/13). The product  $\prod_{k=1}^{45} \csc^2(2k-1)^\circ = m^n$ , where  $m$  and  $n$  are integers greater than 1. Find  $m+n$ .

**Problem 2.36** (Sipnayan 2018 SHS SF A VD 5). Evaluate  $\sum_{k=1}^{2018} \arcsin\left(\frac{1009-k}{2018}\right)$ . Recall that  $\arcsin(x)$  returns a value in the range  $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$  radians. Give your answer in radians.

## 2.3 Sequence guessing skills

The other component of doing engineer's induction effectively is being able to recognize sequences when they come up. There are a couple sequences that come a lot. **For many common sequences, the first three terms should give an instinctive guess.** For example:

- 1, 1, 2 are the Fibonacci numbers.
- 1, 2, 4 are the powers of two.
- 1, 2, 6 are the factorials.
- 1, 3, 5 are the odd numbers.
- 1, 3, 6 are the triangular numbers.
- 1, 4, 9 are the perfect squares.

Sequences like these should be pretty immediate, although if you see them, do a sanity check to make sure that it makes sense. Other sequences that should be clear after three terms are arithmetic sequences and geometric sequences. For everything else, you may need to do some investigation.

**Example 2.37.** What is the next term in the following sequence: 1, 2, 4, 10, 26, ...?

First, we take the differences, which often helps. Here, it's 1, 2, 6, 16. These terms are pretty composite. In particular, note that the  $n$ th term here is divisible by  $n$ , so we divide it out. This gives us 1, 1, 2, 4, ..., which is the original sequence! So this gives us the recurrence  $a_{n+1} = a_n + na_{n-1}$ .

**Exercise 2.38.** Guess the next term in each of the following sequences:

- 2, 4, 7, 11, 16, ...
- 1, 5, 13, 29, 61, ...
- 1, 3, 11, 43, 171, ...
- 4, 10, 22, 46, 94, ...
- 1, 3, 15, 84, 495, ...
- 1, 2, 6, 20, 70, ...

Here are a list of tricks to try:

- **Taking differences between terms** always helps. These can rule out sequences that are polynomial, because they'll have the same difference after a while. In practice, polynomial sequences tend to be either quadratic or cubic, so you should be able to tell after four or five terms.

To elaborate, consider a sequence like 2, 4, 7, 11, 16. The first differences are 2, 3, 4, 5. Taking the second differences gives 1, 1, 1. Since this is constant, the sequence must be quadratic; indeed it's  $\binom{n}{2} + 1$ . Similarly, if the third differences are constant, the sequence is cubic.

- The multiplicative version of this is **taking ratios of consecutive terms**, which can help you figure out whether the sequence has some exponential component. Consider a sequence like 1, 5, 13, 29, 61. The ratios are all approximately 2, so it's a good idea to compare the sequence to  $2^n$ . From here, we can guess the formula  $2^{n+2} - 3$ .

Similarly, consider a sequence like 1, 3, 11, 43, 171. Each term looks like it's about 4 times the previous one. In fact, it's 4 times the previous term minus 1.

- If all of the terms look "very" composite, try **factoring the terms**. You might be able to rule out some exponential component, like if the  $n$ th term had  $2^n$  as a factor, or if the  $n$ th term had  $n$  as a factor.

The other possibility is that all of your terms are binomial coefficients, which is likely if your terms get "more and more composite" as the sequence goes on. That is, if the terms seem to get more and more factors. A sequence like 1, 3, 15, 84, 495 is a good candidate. More about this in the next section.

- Finally, if you know that the problem is combinatorial and probably involves recursion, you can always **force a linear recurrence**. Four terms are enough to find any linear recurrence of the form  $a_{n+1} = xa_n + ya_{n-1}$ . Again, more about this later.

If the sequence does fall in one of the patterns above, you can usually tell it after three or four terms, maybe five terms. Often, doing the brute force until the fifth term isn't too bad, especially if you don't have an alternative. There isn't really much else to talk about this, so here are the problems:

### Exercises

**Problem 2.39** (HMMT Guts 2017/15). Start by writing the integers 1, 2, 4, 6 on the blackboard. At each step, write the smallest positive integer  $n$  that satisfies both of the following properties on the board.

- $n$  is larger than any integer on the board currently.
- $n$  cannot be written as the sum of 2 distinct integers on the board.

Find the 100th positive integer you write on the board. Recall that at the beginning, there are already 4 integers on the board.

**Problem 2.40** (HMMT Guts 2018/11). Find the value of  $\sum_{k=1}^{60} \sum_{n=1}^k \frac{n^2}{61-2n}$ . **Hint: 11**

**Problem 2.41** (CMIMC Combinatorics 2018/9). Compute the number of rearrangements  $a_1, a_2, \dots, a_{2018}$  of the sequence  $1, 2, \dots, 2018$  such that the inequality  $a_k > k$  is true for *exactly* one value of  $k$ .

**Problem 2.42** (HMMT Combinatorics 2015/5). For positive integers  $x$ , let  $g(x)$  be the number of blocks of consecutive 1's in the binary expansion of  $x$ . For example,  $g(19) = 2$  because  $19 = 10011_2$  has a block of one 1 at the beginning and a block of two 1's at the end, and  $g(7) = 1$  because  $7 = 111_2$  only has a single block of three 1's. Compute  $g(1) + g(2) + g(3) + \dots + g(256)$ .

**Problem 2.43** (HMMT Guts 2010/15). Pick a random integer between 0 and 4095, inclusive. Write it in base 2 (without any leading zeroes). What is the expected number of consecutive digits that are not the same (that is, the expected number of occurrences of either 01 or 10 in the base 2 representation)?

### 2.3.1 Binomial coefficients

As mentioned in the previous section, **if the terms have a lot of factors, they may be binomial coefficients**. This should be especially true if you know it's a combinatorics problem and the answer could involve binomial coefficients. In these cases, it's helpful to draw out the first few rows of Pascal's triangle and start looking. Here's a simple example:

**Example 2.44** (PMO 2020 Areas I.4). Determine the number of ordered quadruples  $(a, b, c, d)$  of odd positive integers that satisfy the equation  $a + b + c + d = 30$ .

Since the minimum sum is 4, and the sum can only ever be even, we can try replacing 30 with 4, 6, 8, and count.

There's only 1 solution to  $a + b + c + d = 4$ . There are 4 solutions to  $a + b + c + d = 6$ , corresponding to  $(3, 1, 1, 1)$  and the 4 ways to arrange it. For  $a + b + c + d = 8$ , there's  $(3, 3, 1, 1)$  which we can arrange in 6 ways, and then  $(5, 1, 1, 1)$  which we can arrange in 4 ways, making a total of 10. So the sequence goes 1, 4, 10. This is already enough to make a guess!

$$\begin{array}{cccccccc}
 & & & & & & & 1 \\
 & & & & & & 1 & & 1 \\
 & & & & & 1 & 2 & 1 & & \\
 & & & 1 & 3 & 6 & 3 & 1 & & \\
 & & 1 & 4 & 10 & 6 & 4 & 1 & & \\
 & 1 & 5 & 10 & 10 & 5 & 1 & & & \\
 1 & 6 & 15 & 20 & 15 & 6 & 1 & & & 
 \end{array}$$

Drawing the first few rows of Pascal's triangle, we can spot 1, 4, 10 as a line going through  $\binom{3}{3}$ ,  $\binom{4}{3}$ , and  $\binom{5}{3}$ . So we'd guess that for  $a + b + c + d = 2n$ , the answer should be  $\binom{n+1}{3}$ . Note that this is also roughly the correct order of magnitude—it's growing like a cubic polynomial, so it isn't too slow nor too fast. And indeed,  $\binom{16}{3} = 560$  is the correct answer for this problem.

Usually the pattern is something like  $\binom{f(n)}{n}$ , or  $\binom{f(n)}{k}$  for some constant  $k$ , up to shifting. In particular, you wouldn't expect the pattern to lie only on a single row; you kind of expect it to lie on a diagonal of Pascal's triangle.

### Exercises

**Problem 2.45.** How many four-digit numbers  $\overline{ABCD}$  satisfy  $A \geq B \geq C \geq D$ ? *Hint: 6*

**Problem 2.46** (HMMT Feb Guts 2016/6). Consider a  $2 \times n$  grid of points and a path consisting of  $2n - 1$  straight line segments connecting all these  $2n$  points, starting from the bottom left corner and ending at the upper right corner. Such a path is called *efficient* if each point is only passed through once and no two line segments intersect. How many efficient paths are there when  $n = 2016$ ?

**Problem 2.47** (HMMT Combinatorics 2007/9). Let  $S$  denote the set of all triples  $(i, j, k)$  of positive integers where  $i + j + k = 17$ . Compute

$$\sum_{(i,j,k) \in S} ijk.$$

**Problem 2.48** (HMMT Guts 2004/34). Find the number of 20-tuples of integers  $x_1, \dots, x_{10}, y_1, \dots, y_{10}$  with the following properties:

- $1 \leq x_i \leq 10$  and  $1 \leq y_i \leq 10$  for each  $i$ ;
- $x_i \leq x_{i+1}$  for  $i = 1, \dots, 9$ ;
- if  $x_i = x_{i+1}$  then  $y_i \leq y_{i+1}$ .

### 2.3.2 Linear recurrences

As a last resort, and *only* as a last resort, if you know that the problem is the kind of problem where you're expected to set up a linear recursion and extend it a couple terms, you could always **solve for a linear recurrence** that fits.

If there are two "base cases", then the recurrence is probably of the form  $a_{n+1} = xa_n + ya_{n-1}$ . Solving for an equation like this requires only four terms. If  $x$  and  $y$  end up being integers, and it matches the fifth term, then you're likely correct. Similarly, if there are three "base cases", then you'll need six terms to solve for the coefficients, and a seventh term to check.

This really is only applicable to a few cases. Many problems that look like they have recursion-ish solutions don't have *linear* recurrences as solutions.

**Example 2.49** (PMO 2019 Areas I.11). A *Vitas word* is a string of letters that satisfies the following conditions:

- It consists of only the letters B, L, R.
- It begins with a B and ends in an L.
- No two consecutive letters are the same.

How many Vitas words are there with 11 letters?

Upon reading the problem, it feels like exactly the kind of problem where you'd expect the answers to form some kind of linear recurrence. How far back would this linear recurrence be? Since the condition only involves two consecutive letters, it only really needs to look back by two. So from  $a_{n+1}$ , it only "cares" about  $a_n$  and  $a_{n-1}$ . This means we'd expect a linear recurrence of the form  $a_{n+1} = xa_n + ya_{n-1}$ .

Now let's begin listing the answers for small  $n$ . For  $n = 1$ , there's nothing, so 0. For  $n = 2$  it's just BL, so 1. Similarly, for  $n = 3$  it's just BLR, so 1. For  $n = 4$ , there are 3: BRBL, BLBL, BLRL. This gives us the sequence 0, 1, 1, 3, and if we solve the system

$$1 = 1x + 0y \quad 3 = 1x + 1y$$

we get  $x = 1$  and  $y = 2$ . So we expect the recursion to be  $a_{n+1} = a_n + 2a_{n-1}$ . We can check this by going one step further:

**Exercise 2.50.** Check that there are 11 different Vitas words with 5 letters.

From here, we can continue the recursion to get 11, 21, 43, 85, 171, 341, which gives us the correct answer.

### Exercises

**Problem 2.51** (PMO 2020 Qualifying III.3). A string of 6 digits, each taken from the set  $\{0, 1, 2\}$ , is to be formed. The string should **not** contain any of the substrings 012, 120, and 201. How many such 6-digit strings can be formed?

**Problem 2.52** (AMC 12A 2007/25). Call a set of integers *spacy* if it contains no more than one out of any three consecutive integers. How many subsets of  $\{1, 2, 3, \dots, 12\}$ , including the empty set, are spacy?

**Problem 2.53** (AIME I 2001/14). A mail carrier delivers mail to the nineteen houses on the east side of Elm Street. The carrier notices that no two adjacent houses ever get mail on the same day, but that there are never more than two houses in a row that get no mail on the same day. How many different patterns of mail delivery are possible?

## 3 And other tricks

Other than engineer's induction, which is the namesake of this handout, there are some other engineering-flavored tricks. Here we present three of them: abusing degrees of freedom, using the answer choices, and using the answer format.

### 3.1 Abusing degrees of freedom

In the introduction, we talked about PMO 2019 Qualifying III.4, where we can abuse the degrees of freedom in the problem. For short-answer contests, we often know that the answer has to be numerical. This means that the answer should be the same no matter what configuration we pick, so we only have to **try a specific case** to find it.

**Example 3.1** (PMO 2017 Areas I.4). Suppose that  $S_k$  is the sum of the first  $k$  terms of an arithmetic sequence with common difference 3. If the value of  $S_{3n}/S_n$  does not depend on  $n$ , what is the 100th term of the sequence?

The morally correct way to solve this would be by writing out  $S_n$  in terms of  $n$ , writing out the ratio  $S_{3n}/S_n$ , and choosing the value of the first term such that the  $n$  cancels out or something.

But we don't have to do this. Since we know the problem is correct anyway, we can just pick two specific values of  $n$  and equate them. In particular, let's equate  $S_3/S_1$  and  $S_9/S_3$ .

**Exercise 3.2.** Letting  $a$  be the first term, write  $S_1$ ,  $S_3$ , and  $S_9$  in terms of  $a$ . Check that  $S_3/S_1 = S_9/S_3$  becomes  $S_3^2 = S_1S_9$ , and the  $9a^2$  on both sides cancel to make a nice linear equation. From here, solve for  $a = \frac{3}{2}$ .

A problem can seem to signal something deeper. It could be that  $S_{3n}/S_n$  is a ratio that doesn't depend on  $n$ , or that given a functional equation, then the value of  $f(x+1) - f(x)$  is always the same. Generally it's of the form, *given this condition, all the solutions share this property*. We don't actually have to show that this is true—we can just take it for granted, find a specific solution, and then use that.

Not really much to talk about here again, so here are problems:

### Exercises

**Problem 3.3** (AMC 10B 2002/20). Let  $a$ ,  $b$ , and  $c$  be real numbers such that  $a - 7b + 8c = 4$  and  $8a + 4b - c = 7$ . What is the value of  $a^2 - b^2 + c^2$ ?

**Problem 3.4** (PUMaC Algebra B 2014/2). The function  $f : \mathbb{Z}^{\geq 0} \rightarrow \mathbb{Z}^{\geq 0}$  satisfies, for all  $n \in \mathbb{Z}^{\geq 0}$ ,  $f(f(n)) + f(n) = 2n + 3$ . Find  $f(2014)$ .

**Problem 3.5** (Mathira Elims 2018/2). Each square of a  $69 \times 69$  board is colored either maroon or green. Each maroon square not adjacent to an edge is adjacent exactly 5 green squares out of its 8 neighbors. Each green square not adjacent to an edge is adjacent to exactly 4 maroon squares out of its 8 neighbors. Find the number of maroon squares.

**Problem 3.6.** Complex numbers  $x$ ,  $y$ , and  $z$  satisfy  $|x| = |y| = |z| = xyz = 1$  and  $x + y + z = 0$ . Find  $|(2+x)(2+y)(2+z)|$ . **Hint:** 15

**Problem 3.7** (AMC 10A 2016/23). A binary operation  $\diamond$  has the properties that  $a \diamond (b \diamond c) = (a \diamond b) \cdot c$  and that  $a \diamond a = 1$  for all nonzero real numbers  $a$ ,  $b$ , and  $c$ . (Here  $\cdot$  represents multiplication). The solution to the equation  $2016 \diamond (6 \diamond x) = 100$  can be written as  $\frac{p}{q}$ , where  $p$  and  $q$  are relatively prime positive integers. What is  $p + q$ ?

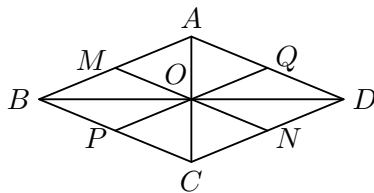
### 3.1.1 Take the boundaries

For geometry problems in particular, this means **taking degenerate cases**. This doesn't happen a lot, since often the conditions in a geometry problem make it so that there's only one possible configuration. But if there are multiple possible configurations, we can exploit this. If you have a point  $P$  that is allowed to be anywhere on a segment  $AB$ , why not take the limit as  $P = A$  or  $P = B$ ? If your triangle can be anything, make it equilateral, or make all the points collinear.

**Example 3.8** (HMMT Guts 2017/6). Let  $ABCD$  be a convex quadrilateral with  $AC = 7$  and  $BD = 17$ . Let  $M$ ,  $P$ ,  $N$ ,  $Q$  be the midpoints of sides  $AB$ ,  $BC$ ,  $CD$ ,  $DA$ , respectively. Compute  $MN^2 + PQ^2$ .



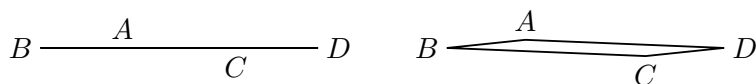
Note that we can pick any quadrilateral where  $AC = 7$  and  $BD = 17$ , so it's best to pick something that's easy to work with. One option is a rhombus, where all the sides are equal. In a rhombus, the diagonals are perpendicular bisectors, and  $MN$  and  $PQ$  both pass through their intersection, which we'll call  $O$ .



So we only have to find  $MO$ . But look at right triangle  $ABO$ . Then  $MO$  is the segment joining the midpoint of the hypotenuse to the opposite vertex, which means it's half the length of the hypotenuse.

**Exercise 3.9.** Show that  $MO = \sqrt{\left(\frac{7}{2}\right)^2 + \left(\frac{17}{2}\right)^2}$ . Don't simplify it yet. Note that the final answer is  $MN^2 + PQ^2$ , and as  $MN = PQ = 2MO$ . From here, show that the final answer is  $\frac{7^2 + 17^2}{2} = 169$ .

Of course, there's an even simpler solution than using a rhombus—just make all the four points collinear! Let the points  $B, A, C$ , and  $D$  be on a line in that order. Then we can pick  $BA = 5$ ,  $AC = 7$ , and  $CD = 5$ , and this will still satisfy the conditions.



The only condition that it doesn't satisfy, technically, is convexity. But we can wiggle around the points so that it becomes convex. The answer would still be the same; the lengths can't change *that* much.

**Exercise 3.10.** Show that, in this case,  $MN^2 + PQ^2$  is still 169.

In fact, it might be even easier to just take  $A = B$ ,  $AC = 7$ , and then  $CD = 10$ , which would still satisfy the conditions. Either way, here are problems:

### Exercises

**Problem 3.11** (AIMO 2019/6). Let  $ABCD$  be a parallelogram. Point  $P$  is on  $AB$  produced such that  $DP$  bisects  $BC$  at  $N$ . Point  $Q$  is on  $BA$  produced such that  $CQ$  bisects  $AD$  at  $M$ . Lines  $DP$  and  $CQ$  meet at  $O$ . If the area of parallelogram  $ABCD$  is 192, find the area of triangle  $POQ$ .

**Problem 3.12** (HMMT Geometry 2017/4). Let  $ABCD$  be a convex quadrilateral with  $AB = 5$ ,  $BC = 6$ ,  $CD = 7$ , and  $DA = 8$ . Let  $M, P, N, Q$  be the midpoints of sides  $AB, BC, CD, DA$ , respectively. Compute  $MN^2 - PQ^2$ .

**Problem 3.13.** Triangle  $ABC$  satisfies  $AB = 5$ ,  $BC = 6$ , and  $CA = 7$ . Let point  $P$  lie on side  $BC$ . Let  $I_1$  and  $I_2$  be the incenters of triangles  $ABP$  and  $ACP$ , respectively. The circumcircle of triangle  $PI_1I_2$  passes through a fixed point  $X$  as  $P$  varies on side  $BC$ . Find the ratio  $BX/CX$ . **Hint:** 2

**Problem 3.14** (HMMT Geometry 2018/6). Let  $ABC$  be an equilateral triangle of side length 1. For a real number  $0 < x < 0.5$ , let  $A_1$  and  $A_2$  be the points on side  $BC$  such that  $A_1B = A_2C = x$ , and let  $T_A = \triangle AA_1A_2$ . Construct the triangles  $T_B = \triangle BB_1B_2$  and  $T_C = \triangle CC_1C_2$  similarly.

There exist positive rational numbers  $b, c$  such that the region of points inside all three triangles  $T_A, T_B, T_C$  is a hexagon with area

$$\frac{8x^2 - bx + c}{(2-x)(x+1)} \cdot \frac{\sqrt{3}}{4}.$$

Find  $(b, c)$ .

### 3.2 Using choices

We can sometimes use answer choices, if they exist, to solve the problem. We can **eliminate incorrect answers** through modulo or divisibility considerations; the answer should be even, the denominator should be divisible by 7, it should be a rational number times  $\sqrt{17}$ . We can also eliminate incorrect choices through estimating size.

And of course, we can also just **substitute each of the choices to see which one works**. Most of the time problems are written to avoid this from happening, but sometimes there are problems where we can just do this.

**Example 3.15** (AMC 12 2001/5). What is the product of all positive odd integers less than 10000?

(A)  $\frac{10000!}{(5000!)^2}$       (B)  $\frac{10000!}{2^{5000}}$       (C)  $\frac{9999!}{2^{5000}}$       (D)  $\frac{10000!}{2^{5000} \cdot 5000!}$       (E)  $\frac{5000!}{2^{5000}}$

Here, let's use a divisibility consideration. Let's say that  $p$  is some prime slightly less than 5000. Then in the final answer, a factor of  $p$  should appear once, and exactly once, because it's a prime. So let's look at all the choices and count how many times  $p$  appears as a factor:

- (A) It appears in the numerator twice, once for  $p$  and once for  $2p$ . It also appears in the denominator twice: once each in  $5000!$ . So it doesn't appear in (A) as a factor at all.
- (B) It appears in the numerator twice, and it doesn't appear in the denominator at all. So it appears twice, and this answer has to be wrong.
- (C) Like (B), it appears in the numerator twice, and doesn't appear in the denominator at all, so this choice has to be wrong.

**Exercise 3.16.** Finish the analysis for the remaining two choices (D) and (E). Does this eliminate one of them as incorrect? If not, what other considerations can you use to get the correct answer?

Often, tricks that eliminate answer choices won't eliminate all of them as cleanly as in this one. But divisibility and estimation are great ways to sanity check an answer anyway, so it's good practice for those.

## Exercises

**Problem 3.17** (PMO 2016 Qualifying I.10). In triangle  $ABC$ ,  $BD$  is the angle bisector of  $\angle ABC$ , and  $AB = BD$ . Moreover,  $E$  is a point on  $AB$  such that  $AE = AD$ . If  $\angle ACB = 36^\circ$ , find  $\angle BDE$ .

- (A)  $24^\circ$     (B)  $21^\circ$     (C)  $18^\circ$     (D)  $15^\circ$     (E)  $12^\circ$

**Problem 3.18** (AMC 12B 2016/17). In  $\triangle ABC$ ,  $AB = 7$ ,  $BC = 8$ ,  $CA = 9$ , and  $\overline{AH}$  is an altitude. Points  $D$  and  $E$  lie on sides  $\overline{AC}$  and  $\overline{AB}$ , respectively, so that  $\overline{BD}$  and  $\overline{CE}$  are angle bisectors, intersecting  $\overline{AH}$  at  $Q$  and  $P$ , respectively. What is  $PQ$ ?

- (A) 1    (B)  $\frac{5}{8}\sqrt{3}$     (C)  $\frac{4}{5}\sqrt{2}$     (D)  $\frac{8}{15}\sqrt{5}$     (E)  $\frac{6}{5}$

**Problem 3.19** (AMC 10A 2019/18). For some positive integer  $k$ , the repeating base- $k$  representation of the (base-ten) fraction  $\frac{7}{51}$  is  $0.\overline{23}_k = 0.232323\dots_k$ . What is  $k$ ?

- (A) 13    (B) 14    (C) 15    (D) 16    (E) 17

**Problem 3.20** (AMC 12A 2018/24). Alice, Bob, and Carol play a game in which each of them chooses a real number between 0 and 1. The winner of the game is the one whose number is between the numbers chosen by the other two players. Alice announces that she will choose her number uniformly at random from all the numbers between 0 and 1, and Bob announces that he will choose his number uniformly at random from all the numbers between  $\frac{1}{2}$  and  $\frac{2}{3}$ . Armed with this information, what number should Carol choose to maximize her chance of winning?

- (A)  $\frac{1}{2}$     (B)  $\frac{13}{24}$     (C)  $\frac{7}{12}$     (D)  $\frac{5}{8}$     (E)  $\frac{2}{3}$

### 3.3 Using the answer format

Ideally, the answer format of a problem shouldn't reveal anything about how to solve it. But when it does, we can abuse the answer format to make a very good guess without actually solving the problem.

**Example 3.21** (Sipnayan Finals JHS 2018/V/TM). Find the sum

$$\sum_{x=0}^{2018} (6x^5 + 45x^4 + 140x^3 + 225x^2 + 186x + 63).$$

Express your answer in the form  $a^b - 1$  where  $a$  is minimized.

It's likely that  $a$  and  $b$  are integers. The fact that they want us to minimize  $a$  is just to prevent us from submitting, say,  $9^2 - 1$  if the answer was  $3^4 - 1$ . In either case, the answer is something that's very close to a perfect power.

A first good step would be to figure out what  $b$  is. Can  $b$  be something huge, like 10? It's unlikely. Remember that if we sum  $1 + 2 + \dots + n$ , we get something that grows like  $n^2$ . Similarly, the sum  $1 + 4 + \dots + n^2$  grows like  $n^3$ . The largest terms in our sum grow like  $1^5 + 2^5 + \dots + n^5$ , which should be around  $n^6$ . So  $b = 6$  would be a good guess.

Now what is  $a$ ? Again, let's estimate the sum, and only care about the largest part:  $6(1^5 + 2^5 + \dots + 2018^5)$ . This is approximately the integral of  $6x^5$  from  $x = 1$  to  $x = 2018$ . So we expect it to be around  $2019^6$ .

**Exercise 3.22.** Convince yourself that the answer is larger than  $2019^6$  by considering the contribution from the other parts. Using the same integral trick gives a pretty good approximation.

Adding in the other parts would give us something slightly above  $2019^6$ . Since we want  $b = 6$ , the next thing above  $2019^6$  would be  $2020^6$ .

Could it be  $2021^6$ ? We can check how *much* larger than  $2019^6$  we want to go by expanding  $(2019 + 1)^6$ . In this case, the terms match up quite well with the estimates you'd get from doing integrals. So the answer  $2020^6 - 1$  should be pretty convincing, and it is in fact the correct answer.

**Exercise 3.23.** You can solve the problem using engineer's induction too. Noting that the answer should be of the form  $a^6 - 1$ , take the sum up to  $x = 0$ , then  $x = 1$ . You should convince yourself that this gives the same answer  $2020^6 - 1$  that we did.

### Exercises

**Problem 3.24** (Mathira 2018/T3-2). We define the Pabibo Najee sequence to consist of positive integers that is determined by the following recursive formula. For  $k \geq 2$ ,  $a_k = 2a_{k-1} + 9a_{k-2}$ , where  $a_0 = 1$  and  $a_1 = 2$ . Find real numbers  $m$  and  $n$  such that  $m > n$  and  $a_k = \frac{m^{k+1} - n^{k+1}}{m - n}$  for all natural numbers  $k$ .

**Problem 3.25** (Sipnayan SHS Elims 2018/V1). Let  $s_0 = 6$ ,  $s_1 = 6$ , and  $s_n = 2s_{n-1} + 8s_{n-2}$  for  $n \geq 2$ . Define

$$A_n = \sum_{i=0}^n s_i.$$

Find  $A_{2018}$ . Express your answer in the form  $a^b + c^d$ , where  $a$ ,  $b$ ,  $c$ , and  $d$  are positive integers.

**Problem 3.26** (Sipnayan SHS Finals 2017/V/FOP). Lake Nick is an enormous underwater lake enclosing 2017 islands arranged at the vertices of a regular 2017-gon. Adjacent islands are joined with exactly two bridges. One day, Spongebob took a wrong bus from Bikini Bottom and ended up in one island in Lake Nick. When Spongebob remembered that the Hash Slingsing Slasher was living in the lake, he started to think of destroying all the bridges. If the island he is on has at least one bridge still joined to it, he randomly selects one such bridge, crosses it, and immediately destroys it. Otherwise, he stops. What is the probability that Spongebob destroys all the bridges before he stops? Express your answer in the form  $a \left(\frac{b}{c}\right)^d$ , where  $a, b, c, d \in \mathbb{Z}^+$  and  $b, c$  are prime numbers. **Hint:** 12

## 4 Grab bag

Here are some more problems that are susceptible to some sort of engineering, sorted roughly by how hard they are to engineer.

### Grab bag

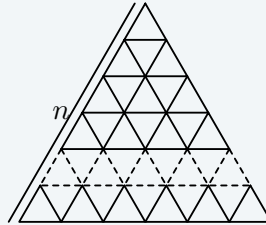
**Problem 4.1** (HMMT Guts 2017/19). Find (in terms of  $n \geq 1$ ) the number of terms with odd

coefficients after expanding the product:

$$\prod_{1 \leq i < j \leq n} (x_i + x_j)$$

e.g., for  $n = 3$  the expanded product is given by  $x_1^2x_2 + x_1^2x_3 + x_2^2x_3 + x_2^2x_1 + x_3^2x_1 + x_3^2x_2 + 2x_1x_2x_3$  and so the answer would be 6.

**Problem 4.2** (Mathira 2019 T10-1). The diagram below shows a grid with  $n$  rows, with the  $k$ th row being composed of  $2k - 1$  identical equilateral triangles for all  $k \in \{1, \dots, n\}$ . If there are 513 different rhombuses each made up of two adjacent smaller triangles in the grid, what must be the value of  $n$ ?



Hint: 10

**Problem 4.3** (Mathira 2017 T12-1). I draw a series of line segments such that  $B$ , the second endpoint of a line segment  $AB$ , is the first endpoint of the next line segment  $BC$ , and  $\angle ABC = 120^\circ$ .  $AB$  is the first line segment,  $BC$  is the second,  $CD$  is the third,  $DE$  is the fourth, and so on, with  $\angle ABC$ ,  $\angle BCD$ , and  $\angle CDE$  all equal to  $120^\circ$  and  $\triangle ABC$  overlaps with  $\triangle BCD$ ,  $\triangle BCD$  overlaps with  $\triangle CDE$  and so on. If every line segment is half of the length of the previous one, the series of line segments eventually becomes a single point  $X$ . What is the ratio of  $AX$  to  $AB$ ? Hint: 8

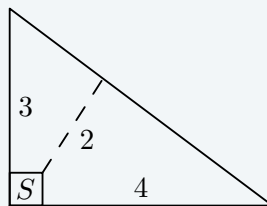
**Problem 4.4** (ARML Individuals 2017/6). A diagonal of a regular 2017-gon is chosen at random. Compute the probability that the chosen diagonal is longer than the median length of all of the diagonals.

**Problem 4.5** (OMO Spring 2014/25). If

$$\sum_{n=1}^{\infty} \frac{\frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n}}{\binom{n+100}{100}} = \frac{p}{q}$$

for relatively prime positive integers  $p, q$ , find  $p + q$ .

**Problem 4.6** (AMC 10A 2018/23). Farmer Pythagoras has a field in the shape of a right triangle. The right triangle's legs have lengths 3 and 4 units. In the corner where those sides meet at a right angle, he leaves a small unplanted square  $S$  so that from the air it looks like the right angle symbol. The rest of the field is planted. The shortest distance from  $S$  to the hypotenuse is 2 units. What fraction of the field is planted?



Hint: 3

**Problem 4.7** (HMMT Algebra and Number Theory 2019/8). There is a unique function  $f : \mathbb{N} \rightarrow \mathbb{R}$  such that  $f(1) > 0$  and such that

$$\sum_{d|n} f(d)f\left(\frac{n}{d}\right) = 1$$

for all  $n \geq 1$ . What is  $f(2018^{2019})$ ? **Hint: 1**

**Problem 4.8** (PUMaC Combinatorics A 2018/4). If  $a$  and  $b$  are selected uniformly from  $\{0, 1, \dots, 511\}$  with replacement, the expected number of 1's in the binary representation of  $a + b$  can be written in simplest form as  $\frac{m}{n}$ . Compute  $m + n$ .

**Problem 4.9** (HMMT Algebra and Number Theory 2017/7). Determine the largest real number  $c$  such that for any 2017 real numbers  $x_1, x_2, \dots, x_{2017}$ , the inequality

$$\sum_{i=1}^{2016} x_i(x_i + x_{i+1}) \geq c \cdot x_{2017}^2$$


holds.

**Problem 4.10** (CMIMC Algebra 2018/7). Compute  $\sum_{k=0}^{2017} \frac{5 + \cos\left(\frac{\pi k}{1009}\right)}{26 + 10 \cos\left(\frac{\pi k}{1009}\right)}$ . **Hint: 5**

## 5 Hints

1. When computing  $f(p)$ ,  $f(p^2)$ ,  $f(p^3)$ , *do not* simplify fractions.
2. Let  $I$  be the incenter. Moving  $P$  to  $B$ , the circle becomes the one with diameter  $BI$ .
3. If  $S$  has side length  $s$ , the shortest distance in the problem is linear in terms of  $s$ .
4. Find it when the exponent is 1 by replacing 2013. Then find it for 2, then 3.
5. This is tricky. Replacing 1009 with  $n$ , notice that the denominators divide  $5^{2n} - 1$ .
6. Engineer on the base. How many are there in base 1, base 2, base 3?
7. What's  $1/9$  in base 10?
8. Draw a good diagram. Your first reasonable guess is probably right.
9. When you have a guess, expand  $(1 + 0.5)^n$ . Most terms are small and can be ignored.
10. Estimate. "Most" triangles are part of 3 rhombi, and each rhombus contains 2 triangles.
11. Replacing 60 with  $n$ , the answer has to be cubic in terms of  $n$ .
12. Tricky. Replace 2017 with  $n$  and engineer. Assume  $b$  and  $c$  are constant over all  $n$ .
13. Replace 1000 with  $n$ , 1001 with  $n + 1$ , and 1000000 with  $n^2$ .
14. Guessing the sequence is hard. Try again after reading [subsection 2.3](#).
15. Pick three nice numbers on the unit circle that sum to zero.

## Acknowledgments

Thanks to Ankan Bhattacharya and Vincent Huang for references, David Altizio for suggestions about the introduction, Raymond Feng for a suggestion about [Example 3.21](#), and Avi Mehra for typo corrections. If you spot a typo, see an error, have a suggestion, or want to ask a question, feel free to contact me at [cj@cjquines.com](mailto:cj@cjquines.com) .