Problem Solving Session 3 – Solutions

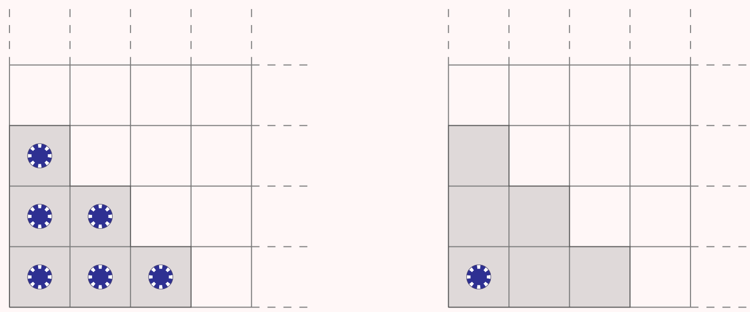
Shreeansh Hota, Naman Mishra

Problem 1. Consider all lattice squares (x, y) with x, y nonnegative integers. Assign to each its lower left corner as a label. We shade the squares $(0, 0)$, $(1, 0)$, $(2, 0)$, $(0, 1)$, $(1, 1)$ and $(0, 2)$.

(Case 1) There is a chip on each of the six squares.

(Case 2) There is only one chip on $(0, 0)$.

In one step, if (x, y) is occupied, but $(x + 1, y)$ and $(x, y + 1)$ are free, you may remove the chip from (x, y) and place a chip on each of $(x + 1, y)$ and $(x, y + 1)$. The goal is to remove the chips from the shaded squares. Is this possible in (Case 1) or (Case 2)?



Source: Numberphile

Numberphile has excellent videos on this problem: Pebbling a Chessboard and More about Pebbling a Chessboard. The solution is heavily inspired from these.

Solution. The idea is to give weights to each square such that the sum of weights of all occupied squares remains invariant.

$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$
$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$
$\frac{1}{1}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$

Whenever a chip at (x, y) is split into a chip each at $(x+1, y)$ and $(x, y+1)$, the weights of the new chips sum to the weight of the old chip.

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

The sum of weights of the first row is

$$1 + \frac{1}{2} + \frac{1}{4} + \cdots = 2.$$

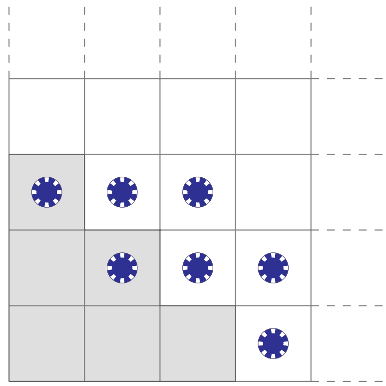
Summing the board row-by-row gives

$$2 + 1 + \frac{1}{2} + \cdots = 4.$$

The sum of the shaded squares is $1 + 2 \times \frac{1}{2} + 3 \times \frac{1}{4} = 2 + \frac{3}{4}$.

(Case 1) The initial sum of weights is $2 + \frac{3}{4}$. The sum of the unshaded region is $1 + \frac{1}{4}$. Since the sum of weights is invariant, it is impossible to move all chips from inside the shaded region to outside.

(Case 2) The initial sum of weights is just 1. The sum of the unshaded region is $1 + \frac{1}{4}$. It appears that it is possible to empty the shaded region. If you try to do so, you might reach the following configuration or similar:



Realize that removing the chip at $(3, 0)$ gives no benefit. It is already outside the shaded region, and it will only introduce more weight into the higher rows. Thus, there is no benefit to filling any of the squares $(4, 0), (5, 0), \ldots$! This is $\frac{1}{16} + \frac{1}{32} + \cdots = \frac{1}{8}$ weight that we miss out on.

Formally, there will only ever be one chip in the bottom row (check). Similarly, there will only ever be one chip in the leftmost column. Thus the total weight that may be occupied by chips in the unshaded region is $1 + \frac{1}{4} - \frac{1}{8} - \frac{1}{8} = 1$. This is exactly the initial weight. However, this would require occupying infinitely many squares, which cannot be done in a finite number of steps. ■

Problem 2. A device generates a random number uniformly distributed on $[0, 1]$ each time a button is pressed. Numbers are generated sequentially until their cumulative sum surpasses 1. Determine:

(1) The probability of requiring more than n button presses.

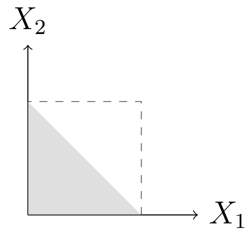
(2) The expected number of button presses needed.

Source: The r/quant subreddit

Solution. Let $X_1, X_2, \cdots \sim \text{Unif}[0, 1]$ be i.i.d. random variables, and let N be the smallest n such that $X_1 + X_2 + \cdots + X_n \geq 1$ (the number of button presses needed). Note that

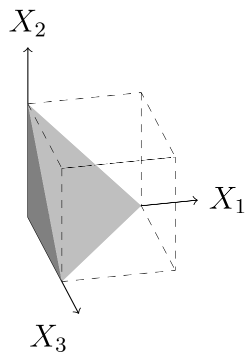
$$\{N > n\} = \{X_1 + X_2 + \cdots + X_n < 1\}.$$

The event that $X_1 + X_2 < 1$ is the region below the line $x + y = 1$.



This has area $\frac{1}{2}$. The entire sample space is $[0,1]^2$, which has area 1.

Similarly, the event that $X_1+X_2+X_3 < 1$ is the region below the plane $x+y+z = 1$.



This has volume $\frac{1}{6}$, and the entire sample space $[0,1]^3$ has volume 1.

In general, the n-simplex

$$T_n = \{(x_1, x_2, \ldots, x_n) \in [0,1]^n \mid x_1 + x_2 + \cdots + x_n < 1\}$$

has volume $\frac{1}{n!}$ (exercise: integrate[1]). Thus

$$\mathbf{P}\{X_1 + X_2 + \cdots + X_n < 1\} = \frac{1}{n!}.$$

Using $\mathbf{E}[N] = \sum_{n=0}^{\infty} \mathbf{P}\{N > n\}$, we have

$$\begin{aligned} \mathbf{E}[N] &= \sum_{n=1}^{\infty} \mathbf{P}\{X_1 + X_2 + \cdots + X_n < 1\} \\ &= \sum_{n=1}^{\infty} \frac{1}{n!} \\ &= e. \end{aligned}$$

■

[1] MSE has pretty proofs without calculus.

Problem 3. On a strange railway line, there is just one infinitely long track, so overtaking is impossible. Any time a train catches up to the one in front of it, they link up to form a single train moving at

(Case 1) the speed of the faster train.

(Case 2) the speed of the slower train.

(Case 3) the average speed of the two trains.

At first, there are n equally spaced trains, each moving at a different speed. You watch, and eventually (after all the linking that will happen has happened), you count the trains. You wonder what would have happened if the trains had started in a different order (but each of the original n trains had kept its same starting speed). On average (averaging over all possible orderings), how many trains will there be after a long time has elapsed in each of the three cases?

Source: PROMYS India 2023 Applications – Problem 10

Solution. Let $v_1, v_2, \ldots, v_n$ be the speeds of the trains, and let $L(v_1, v_2, \ldots, v_n)$ be the number of trains after a long time has elapsed.

(Case 1) The slowest train will link up with some train from its left, and start moving at its speed.

Thus we have two cases:

- If the slowest train is the leftmost, it will never link with any other train, and

$$L(v_1, v_2, \ldots, v_n) = 1 + L(v_2, v_3, \ldots, v_n).$$

- If the slowest train is at position $i > 1$, it has no effect on the rest of the trains.

$$L(v_1, v_2, \ldots, v_n) = L(v_1, \ldots v_{i-1}, v_{i+1}, \ldots, v_n).$$

Let L_n be the expected final number of trains when staring with n trains. By symmetry and linearity of expectation,

$$L_n = \frac{1}{n}(1 + L_{n-1}) + \left(1 - \frac{1}{n}\right)L_{n-1} = L_{n-1} + \frac{1}{n}$$

with $L_1 = 1$. This gives $L_n = \sum_{k=1}^{n} \frac{1}{k}$, the nth harmonic number.

(Case 2) Pick a frame of reference in which every train has negative velocity. This becomes a mirror image of the previous case.

(Case 3) This is ill-defined (and was not present in the original problem). It is not specified how the trains' initial speeds are chosen. In the previous cases, only the relative order of the speeds mattered, not any arithmetic properties. Here, $L(5,1,2)$ gives a different answer from $L(5,1,4)$, even though the relative order of the speeds is the same. When the first two trains link, they move with speed 3. Whether or not they catch up to the third train depends on the the precise values of the speeds. ■

Remark. Using similar arguments, one can derive the exact number of initial permutations of trains that result in a certain final number of trains. For n trains and k final supertrains, this turns out to be $\left[{n \atop k}\right]$, the (n,k)-th Stirling number of the first kind.

> **Problem 4.** Find all integers $n \geq 2$ satisfying $\frac{\sigma(n)}{p(n)-1} = n$, where $\sigma(n)$ denotes the sum of all positive divisors of n, and $p(n)$ denotes the largest prime divisor of n.
>
> *Source: APMO 2023 – Problem 2*

Solution. Let n be a solution. Write $n = p_1^{n_1} \dots p_k^{n_k}$ with $p_1 < p_2 < \cdots < p_n$. Then $p(n) = p_k$ and

$$\sigma(n) = \prod_{i=1}^{k} \frac{p_i^{n_i+1} - 1}{p_i - 1} < \prod_{i=1}^{k} \frac{p_i^{n_i+1}}{p_i - 1} = n \prod_{i=1}^{k} \frac{p_i}{p_i - 1} \leq n \frac{p_k}{p_1 - 1} \tag{1}$$

since $p_i \leq p_{i+1} - 1$. Thus

$$p_k - 1 = \frac{\sigma(n)}{n} < \frac{p_k}{p_1 - 1}.$$

This requires $p_1 - 1 = 1$, so $2 \mid n$. Reject powers of 2, since

$$n = 2^m > 1 \implies \sigma(n) = 2^{m+1} - 1 = 2n - 1 \geq n + 1 > (p(n) - 1)n.$$

Thus we may assume $k > 1 \iff p_k \geq 3$.

Then expanding out the product in equation (1) gives

$$\begin{aligned} p_k - 1 = \frac{\sigma(n)}{n} &\leq \frac{p_1}{p_1 - 1} \frac{p_2}{p_2 - 1} \cdots \frac{p_k}{p_k - 1} \\ &= \frac{2}{p_2 - 1} \frac{p_2}{p_3 - 1} \cdots \frac{p_{k-1}}{p_k - 1} p_k \end{aligned}$$

Each fraction is at most 1. If $p_2 \geq 5$, then the first ratio becomes $\frac{1}{2}$. Then again $p_k - 1 \leq \frac{p_k}{2}$, a contradiction. Thus $p_2 = 3$. Finally,

$$p_k - 1 \leq \frac{2}{3}p_k \implies p_k \leq 3,$$

so that $n = 2^a 3^b$ with $a, b \geq 1$.

$$\begin{aligned}\frac{\sigma(n)}{p(n)-1} &= \frac{(2^{a+1}-1)(3^{b+1}-1)}{4} = 2^a 3^b \\ &\implies \left(1 - \frac{1}{2^{a+1}}\right)\left(3 - \frac{1}{3^b}\right) = 2.\end{aligned}$$

But

$$\left(1 - \frac{1}{2^{a+1}}\right)\left(3 - \frac{1}{3^b}\right) \geq \left(1 - \frac{1}{4}\right)\left(3 - \frac{1}{3}\right) = \left(\frac{3}{4}\right)\left(\frac{8}{3}\right) = 2,$$

with equality holding only for $a = b = 1$. Thus $n = 6$ is the only solution. ■

Problem 5. A set of three nonnegative integers $\{x, y, z\}$ with $x < y < z$ is called *historic* if $\{z - y, y - x\} = \{1776, 2001\}$. Show that the set of all nonnegative integers can be written as the union of pairwise disjoint historic sets.

Source: 2001 IMO Shortlist – Problem C4

The solution is (abridged) from the source above (Art of Problem Solving).

Solution. Let $a = 1776$ and $b = 2001$. The exact values of a and b are not important, only that $a < b$. For us, historic might mean 1947 and 2008.

We describe a greedy algorithm to cover all integers. If an integer has been included in a set, we call it 'colored'.

In one step of the algorithm:

(1) Take the smallest integer k which is not colored.

(2) If $k + a$ is likewise not colored, then the new historic set is $\{k, k + a, k + a + b\}$.

(3) If $k + a$ is colored, then the new historic set is $\{k, k + b, k + a + b\}$.

If we manage to show that the three numbers chosen in either case were not previously colored, then the algorithm would only color each integer at most once. Further, since each step of the algorithm colors the smallest integer not yet colored, any integer will eventually be colored.

Claim. *If k is the smallest integer not yet colored by the algorithm, then $k+a+b$ is not yet colored.*

Proof of claim. The algorithm picks the smallest integer not yet colored. Thus any k' chosen previously by the algorithm must be smaller than k, otherwise k would have been chosen in its place.

But the largest integer colored by the algorithm is $a+b$ more than the chosen k' irrespective of whether case (2) or (3) was used. Thus the largest number colored by the algorithm is $k'+a+b$ for some $k'<k$. This is less than $k+a+b$. □

Claim. *If k is the smallest integer not yet colored by the algorithm, and $k+a$ is likewise not colored, then $k+b$ is not yet colored.*

Proof of claim. Assume $k+b$ was colored by the algorithm when some k' was chosen. Then $k+b \in \{k', k'+a, k'+a+b\}$ or $k+b \in \{k', k'+b, k'+a+b\}$, depending on the case. Thus $k \in \{k'-b, k'+(a-b), k', k'+a\}$. The first three elements are $\leq k'$ (since $a<b$), so they are less than k. Thus $k=k'+a$. But then k would have been colored when k' was chosen, a contradiction. □

We have shown that $k+a+b$ cannot be previously colored in either case, and that $k+b$ cannot be colored in the second case. $k+a$ cannot be colored in the first case simply because that is the requirement for the case to be chosen.

Thus we have managed to show what we needed. ■

Problem 6. Let $\mathbb{R}^+$ be the set of all positive real numbers. Find all functions $f\colon \mathbb{R}^+ \to \mathbb{R}^+$ satisfying

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

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

Source: 2023 India IMO Training Camp – Practice Test 2, Problem 2

Solution. It is easy to check that $f(x)=kx$ is a solution for all $k \in \mathbb{R}^+$. We will show that these are the only solutions.

Claim. *f is strictly increasing, hence injective.*

Proof of claim. Fix an x_0 and let $x_1 = x_0 + t > x_0$. Since $f(x_0) \neq 0$, we can choose $y = \sqrt{\frac{t}{f(x_0)^2}}$, so that $x_1 = x_0 + y^2 f(x_0^2)$. Then

$$f(x_1) = f(x_0 y)^2 + f(x_0) > f(x_0).$$ □

Claim. *$f(x)/x$ is (weakly) increasing.*

Proof of claim. Fix a $z \in \mathbb{R}^+$ and plug $y = x/z$ to get

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

Since f is strictly increasing, the right-hand side is a strictly increasing function of x. Thus the left-hand side must also be stricly increasing in x. This requires

$$x + z^2 \frac{f(x^2)}{x^2}$$

to be strictly increasing in x for *any* z (since z was arbitrary). That is, for $x_1 < x_2$,

$$\begin{aligned} x_1 + z^2 \frac{f(x_1^2)}{x_1^2} &< x_2 + z^2 \frac{f(x_2^2)}{x_2^2} \\ \implies (x_1 - x_2) + z^2 \left(\frac{f(x_1^2)}{x_1^2} - \frac{f(x_2^2)}{x_2^2} \right) &< 0 \end{aligned}$$

for any $z \in \mathbb{R}^+$. If $\frac{f(x_1^2)}{x_1^2} - \frac{f(x_2^2)}{x_2^2} > 0$, then choosing z large enough would violate this inequality. Thus for all $x_1 < x_2$, we have

$$\frac{f(x_1^2)}{x_1^2} \leq \frac{f(x_2^2)}{x_2^2}. \qquad \square$$

Let $f(1) = k$.

Claim. *$f(x) = kx$ for all $x \leq 1$.*

Proof of claim. Plugging $(x, y) = (1, 1)$ gives $f(1 + k) = k^2 + k$. But then

$$\frac{f(1+k)}{1+k} = k = \frac{f(1)}{1}.$$

By the increasing nature of $f(x)/x$, we have $f(x) = kx$ for all $x \in [1, 1 + k]$.

Let $t < 1$. Now plug in $(x, y) = (1, t)$ to get

$$f\left(1 + t^2 k\right) = f(t)^2 + k$$

Now $1 + t^2 k \in [1, 1 + k]$, so

$$\begin{aligned} k + t^2 k^2 &= f(t)^2 + k \\ \implies f(t) &= kt. \end{aligned} \qquad \square$$

We come to the final claim.

Claim. *Suppose there is an $\alpha > 1$ such that $f(t) = kt$ for all $t \le \alpha$. Then $f(t) = kt$ for all $t \le \alpha^2$.*

Proof of claim. Plugging in $(x, y) = (\alpha, 1)$ gives

$$f(\alpha + f(\alpha^2)) = f(\alpha)^2 + f(\alpha) = k^2\alpha^2 + k\alpha.$$

Since $f(x)/x$ is increasing, we have

$$\begin{aligned} \frac{f(\alpha + f(\alpha^2))}{\alpha + f(\alpha^2)} &\ge \frac{f(\alpha)}{\alpha} = k \\ \implies k^2\alpha^2 + k\alpha &\ge k(\alpha + f(\alpha^2)) \\ \implies k &\ge \frac{f(\alpha^2)}{\alpha^2}. \end{aligned}$$

Thus $\frac{f(\alpha^2)}{\alpha^2} = k$. By the monotonicity of $f(x)/x$, we have $f(t) = kt$ for all $t \le \alpha^2$. □

We already know this for $t \le 1 + k$. Thus we have $f(t) = kt$ for all $t \le (1 + k)^2$. By induction, $f(t) = kt$ for all $t \in \mathbb{R}^+$, since $(1 + k)^{2^n} \to \infty$. ■

Problem – Bonus. Let $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$ be sequences of positive real numbers such that $a_1 = b_1 = 1$ and $b_n = b_{n-1}a_n - 2$ for $n = 2, 3, \ldots$. Assume that the sequence (b_j) is bounded. Prove that

$$S = \sum_{n=1}^{\infty} \frac{1}{a_1 a_2 \ldots a_n}$$

converges, and evaluate S.

Source: Putnam 2011 – Problem A2

Solution. This blog post is too well-written to be paraphrased. ■