The 5$^{\text{th}}$ US Ersatz Math Olympiad

Solutions and Results

Evan Chen

30 June 2024

Contents

1 Summary

The fifth USEMO was held on October 21 – 22, 2023. A total of 70 students submitted at least one paper.

Although not as brutal as last year, this was still a challenging exam. About two-thirds of the students were able to solve at least one problem. Happily, every problem in the exam was solved by at least two students, although problems 2, 3, 5, 6 only had a single-digit number of solves each. (Amusingly, two students solved the final problem 6 but no other problems.) This makes the showings of the top few students extremely impressive.

§1.1 Thanks

I am once again grateful to many individuals who helped make this competition possible.

§1.1a Proposers of problems

I thank Ankan Bhattacharya, Atharva Sathe, Azat Madimarov, Holden Mui, Kaixin Wang, Karthik Vedula, Kornpholkrit Weraarchakul, Krishna Pothapragada, Maxim Li, Nikolai Beluhov, Oleg Kryzhanovsky, Qiao Sun, Sanjana Das, Sathyaram Basker, Serena An, Siraphop Khawplad, Xiaoyu Chen, Zeyu Yang for contributing 36 problem proposals.

§1.1b Reviewers

I thank Andrew Gu, Ankan Bhattacharya, Maxim Li, Nikolai Beluhov, Oleg Kryzhanovsky, William Yue for reviewing the proposed problems.

§1.1c Graders

Thanks to everyone who graded at least one paper: Aleksij Tasikj, Alex Chui, Ana Boiangiu, Anurag Singh, Arianée, Atul Shatavart Nadig, Carlos Villagordo Espinosa, Cerlat Marius, Dan, Danielle Wang, Demira Nedeva, Elizabeth Lau, Galin Milenov Totev, Hao-Yu Gan, Haruka Kimura, Helio Ng, Hu Man Keat, Joao Vitor Carvalho Almeida, Joshua Im, Kaixin Wang, Kanav Talwar, Kevin Liu, Kevin Shi, Kyan Cheung, Liam Celinski, Lincoln Liu, Luis André Villán Gabriel, Max Mei, Mikel Perez de Gracia, Miroslav Marinov, Muhammad Alhafi, Novak Despotović, Orestis Lignos, Paixiao Seeluangsawat, Petko Lazarov, Rushil Mathur, Shikhar Sehgal, Shreeansh Hota, Szymon Tobiasz, Tache David Stefan, Tanupat Trakulthongchai, Teya Chobanova, Will Ren.

§1.1d Other supporters

I would like to thank the Art of Problem Solving for offering the software and platform for us to run the competition. Special thanks to Christie Harrison, Rebecca Sodervick, Deven Ware, and Jo Welsh who collaborated with me.

2 Results

If you won one of the seven awards, please reach out to `usemo@evanchen.cc` to claim your prize!

§2.1 Top Scores

Congratulations to the top three scorers, who win the right to propose problems to future instances of USEMO.

1st place Linus Tang (score 36)

2nd place Hannah Fox (score 30)

3rd place Henrick Rabinovitz (score 28)

§2.2 Special awards

See the Rules for a description of how these are awarded. (Note in particular that students already in the top three above aren't considered for special awards.)

For the purposes of awarding monetary prizes, ties are broken more or less arbitrarily by considering the presentation of elegance of solutions (which is obviously subjective). When this occurs, the names of tied students are noted as well.

Youth prize Alexander Wang

Top female Vivian Loh

Top day 1 Aprameya Tripathy

Top day 2 Qiao Zhang

§2.3 Honorable mentions

This year we award Honorable Mention to anyone scoring at least 16 points (who is not in the top three already). The HMs are listed below in alphabetical order.

Alexander Wang

Allen Wang

Aprameya Tripathy

Carlos Rodriguez

Feodor Yevtushenko

Jason Mao

Qiao Zhang

Zongshu Wu

§2.4 Distinction

The Distinction award is awarded for either scoring at least 14 points or in the top 25 of scores, whichever is more inclusive. This year, the 25th place student scored 8 points, so Distinction awards recognize any student with at least 8 points. The Distinction awards are listed below in alphabetical order.

Aditya Pahuja

Ahmad Alkhalawi

Angela Liu

Ashvin Sinha

Benny Wang

Evan Fan

Grant Blitz

Jiahe Liu

Jordan Lefkowitz

Leo Yu

Mingyue Yang

Neal Yan

Oron Wang

Ritwin Narra

Rohan Das

Shruti Arun

Srinivas Arun

Vivian Loh

Wilbert Chu

3 Solutions and marking schemes

§3.1 USEMO 1 — proposed by Oleg Kryzhanovsky

Problem statement

A positive integer n is called *beautiful* if, for every integer $4 \le b \le 10000$, the base-b representation of n contains the consecutive digits 2, 0, 2, 3 (in this order, from left to right). Determine whether the set of all beautiful integers is finite.

§3.1a Solution

We show there are infinitely many beautiful integers. Here are three different approaches.

¶ One constructive approach. We will construct an increasing sequence of positive integers

$$N_4 < N_5 < N_6 < \dots$$

such that for every $k = 4, 5, \dots$, the number N_k contains 2023_b in every base $4 \le b \le k$. This will solve the problem because N_{10000}, N_{10001}, ... will be the requested infinite set.

For the base case, take $N_4 = 2023_4$.

For the inductive step, here is one of many valid recipes. We are going to select

$$N_k = N_{k-1} + c \cdot (k\ell)^e$$

where the ingredients c, ℓ, e are selected to satisfy:

- ℓ is the product of all primes at most k which are relatively prime to k (in particular, $\gcd(k, \ell) = 1$);
- e is large enough that for each $b = 4, 5, \dots, k$, the largest power of b dividing $(k\ell)^e$ is greater than $b \cdot N_{k-1}$;
- c is chosen to satisfy the modular congruence
$$c \cdot \ell^e \equiv 2k^3 + 0k^2 + 2k + 3 \pmod{k^4}$$
which is possible since $\gcd(k^4, \ell^e) = 1$.

With these ingredients, for all the smaller bases 4, 5, ..., $k-1$, the ending of N_k in base-b is the same as in N_{k-1} (since $(k\ell)^e$ is a multiple of a large enough power of b). On the other hand, we've embedded 2023_k into the base-k representation of N_{k-1}, because the coefficients of k^{e+3}, k^{e+2}, k^{e+1}, k^e in the base-k representation are exactly 2, 0, 2, 3.

> **Remark.** The essential difficulty in this approach arises from the fact that different bases may share overlapping prime factors, so the Chinese Remainder Theorem does not immediately apply. (Indeed, the solution above does not use the Chinese remainder theorem at all.) To give a concrete example, if N is an integer whose base-6 representation contains a specific substring, then there are several restrictions on what possible base-9 representations it could have, which are serious enough to limit the control a student has, but not strong enough to

actually determine any of the digits. That's why in the inductive proof, there is sort of a "critical step" in which the common prime factors are factored out via $(k\ell)^e$, after which c is chosen via modular inverses once the common prime factors have been factored out.

¶ An indirect inductive approach. The goal of this approach is to construct a system of congruences of the form

$$\begin{aligned} N &\equiv 2023_4 \cdot 4^{e_4} + t_4 \pmod{4^{e_4+4}} \\ N &\equiv 2023_5 \cdot 5^{e_5} + t_5 \pmod{5^{e_5+4}} \\ N &\equiv 2023_6 \cdot 6^{e_6} + t_6 \pmod{6^{e_6+4}} \\ &\vdots \\ N &\equiv 2023_{10000} \cdot 10000^{e_{10000}} + t_{10000} \pmod{10000^{e_{10000}+4}} \end{aligned}$$

which has at least one simultaneous solution in N, and where $0 \le t_b < b^{e_b}$. The equation involving b^{e_b+4} will automatically ensure N has the desired substring 2023_b, so we concern ourselves only with ensuring the system of equations is consistent.

In order to do this, we will need a more general version of the Chinese remainder theorem to moduli that are not coprime:

Theorem (Generalized Chinese remainder theorem)

Fix integers a, b, m, n with $m, n > 0$. The equations

$$\begin{aligned} x &\equiv a \pmod{m} \\ x &\equiv b \pmod{n} \end{aligned}$$

have a simultaneous solution if and only if $a \equiv b \pmod{\gcd(m,n)}$. Moreover, if there is a solution, that solution is unique modulo $\operatorname{lcm}(m,n)$.

Remark. The usual Chinese remainder theorem is the case where $\gcd(m,n) = 1$, in which case $a \equiv b \pmod{1}$ is always true.

We now proceed by induction, selecting the pairs (t_4, e_4), (t_5, e_5), ... in order. For the base case, it is enough to take $t_4 = e_4 = 0$.

Now suppose we have selected pairs up to (t_{b-1}, e_{b-1}) for some b and it is time to select (t_b, e_b). The inductive hypothesis means that the previous equations up to base $b-1$ can be collated into a single equivalent equation

$$N \equiv C \pmod{L} \qquad \text{where } L := \operatorname{lcm}\left(4^{e_4+4}, 5^{e_5+4}, \dots, (b-1)^{e_{b+1}+4}\right).$$

The critical observation is that as long as e_b is selected large enough so that

$$b^{e_b} \ge \gcd(L, b^{e_b+4})$$

then at least one of the choices of t_b will satisfy

$$2023_b \cdot b^{e_b} + t_b \equiv C \pmod{\gcd(L, b^{e_b+4})}$$

which is the compatibility condition needed to ensure

$$N \equiv 2023_b \cdot b^{e_b} + t_b \pmod{b^{e_b+4}}$$

can be added to the preceding equations. Since this is obviously possible (by taking $e_b > \log_b L$) the induction is complete.

¶ **Density approach (outline).** It's enough to prove the following claim:

Claim — Fix any integer $b \geq 4$. The arithmetic density of nonnegative integers whose base-b representation does not contain 2023_b as a contiguous substring is zero.

Proof. Think about just the last $4n$ digits in base-b, in n groups of 4. For every complete residue class modulo b^{4n}, the number of base-b numbers that don't have 2023_b in their base-b representation is at most $(b^4-1)^n$.

Consequently, if we have a threshold N, the number of 2023_b-avoiding numbers in $\{0, 1, \dots, N\}$ is bounded above by

$$(b^4-1)^n \cdot \left\lfloor \frac{N}{b^{4n}} \right\rfloor + b^{4n}$$

and so the density of 2023_b-avoiding numbers, for large N, is at most

$$\frac{(b^4-1)^n}{b^{4n}} = \left(1 - \frac{1}{b^4}\right)^n .$$

Since this statement has to hold for any $n \geq 1$, the density must be zero. □

Thus, it follows that "most" positive integers are beautiful!

§3.1b Marking scheme

Marking scheme for inductions and constructions (common approach)

The following remark from the solution packet is key to understanding the rubric in what follows:

> *The essential difficulty in this problem arises from the fact that different bases may share overlapping prime factors, so the Chinese Remainder Theorem does not immediately apply. (Indeed, the solution above does not use the Chinese remainder theorem at all.) To give a concrete example, if N is an integer whose base-6 representation contains a specific substring, then there are several restrictions on what possible base-9 representations it could have, which are serious enough to limit the control a student has, but not strong enough to actually determine any of the digits. That's why in the inductive proof, there is sort of a "critical step" in which the common prime factors are factored out via $(k\ell)^e$, after which c is chosen via modular inverses once the common prime factors have been factored out.*

In short, a solution is considered 7^- once it can achieve this critical step where, to deal with overlapping prime factors, a large set of common primes are factored out so that a suitable modular inverse can be used in order to complete the proof.

To achieve this benchmark, these steps must be written in sufficient detail to be checked; they should be deduced using explicit equations specifying the necessary thresholds and moduli. For example the official solution reads:

> *We are going to select $N_k = N_{k-1} + c \cdot (k\ell)^e$ where the ingredients c, ℓ, e are selected to satisfy:*
>
> - *ℓ is the product of all primes at most k which are relatively prime to k (in particular, $\gcd(k, \ell) = 1$);*

- *e is large enough that for each $b = 4, 5, \dots, k$, the largest power of b dividing $(k\ell)^e$ is greater than $b \cdot N_{k-1}$;*
- *c is chosen to satisfy the modular congruence*
$$c \cdot \ell^e \equiv 2k^3 + 0k^2 + 2k + 3 \pmod{k^4}$$
which is possible since $\gcd(k^4, \ell^e) = 1$.

Merely claiming there is *some* congruence of some sort is not sufficient to pass the benchmark.

For solutions that *fail* to achieve this benchmark, the following partial credits are available but **not additive**:

- **0 points** for yes/no answer alone.
- **0 points** for just mentioning induction.
- **0 points** for base cases like $b = 4$.
- **0 points** for a solution that only works on coprime moduli. This includes, e.g. showing there is a number which has 2023_b for every base $4 \le b \le 2023$ which is the power of a prime (that is, $b \in \{4, 5, 7, 8, 9, 11, 13, \dots, 2011\}$).
- **1 point** for showing that the existence of a *single* beautiful number implies the existence of infinitely many, e.g. by adding large powers of 2023!.
- **1 point** for a serious induction attempt or construction but which botches the main difficulty mentioned above.

  In practice, most students who obtain this item will probably have obtained the previous item (possibly implicitly), so it is not particularly relevant.
- **2 points** for a solution that additionally has the idea in the indirect construction solution of ensuring compatibility by picking a certain integer parameter $t_i \in \{0, 1, \dots, b^k - 1\}$ in a *consecutive range* for a sufficiently large k.

Solutions that pass this benchmark earn:

- **7 points** if they are completely correct.
- **6 points** for a minor error such as
  - completely omitting the base case of the induction;
  - using an integer parameter which is not large enough as stated but could easily be changed to be large enough.
- **5 points** for a more serious but non-central flaw in one of the steps of an inductive approach.

The distinction between 5 and 6 will be done more closely by the problem captain.

Marking scheme for arithmetic density approach

- **0 points** for just conjecturing the arithmetic density of beautiful numbers is 1
- **4 points** for showing that the number of residue classes modulo b^{4n} that don't have 2023_b as a contiguous substring is at most $(b^4 - 1)^n$.
- **6 points** for complete solution modulo minor errors.
- **7 points** for complete solution.

§3.2 USEMO 2 — proposed by Holden Mui

Problem statement

Each point in the plane is labeled with a real number. Show that there exist two distinct points P and Q whose labels differ by less than the distance from P to Q.

§3.2a Solution

Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be the labeling, and suppose for contradiction the difference in labels for any points $P, Q \in \mathbb{R}^2$ is at least their distance.

Claim — Let I be a closed interval of length 1. For any $\varepsilon > 0$, the pre-image

$$f^{-1}(I) := \left\{x \in \mathbb{R}^2 \mid f(x) \in I\right\}$$

can be contained inside a set of squares whose total area is at most ε.

Proof. Let $n \geq 1$. Divide I into n closed intervals of length $\frac{1}{n}$. The problem condition implies that the pre-image of each sub-interval is contained inside a square of side length at most $\frac{2}{n}$, and hence with area at most $\left(\frac{2}{n}\right)^2 = \frac{4}{n^2}$. The total area of the square is thus bounded by $n \cdot \frac{4}{n^2} = \frac{4}{n}$. By taking n large enough that $\varepsilon > 4/n$, we're done. $\square$

Divide the codomain $\mathbb{R}$ into closed intervals

$$\begin{aligned} I_1 &= [0,1] \\ I_2 &= [-1,0] \\ I_3 &= [1,2] \\ I_4 &= [-2,-1] \\ I_5 &= [2,3] \\ I_6 &= [-3,-2] \\ &\vdots \end{aligned}$$

By the claim, the pre-image $f^{-1}(I_k)$ could be contained inside squares whose total area is at most, say, 10^{-k}. So the entire pre-image $f^{-1}(\mathbb{R})$ could be contained inside squares whose total area is at most $\sum_{k\geq 1} 10^{-k} = \frac{1}{9}$, which is finite. But this is absurd, since $f^{-1}(\mathbb{R}) = \mathbb{R}^2$.

Remark (Generalization of the problem)**.** One can imagine the same problem with the target condition modified to

$$|f(P) - f(Q)| < |PQ|^c$$

for a general $c > 0$; the present problem is $c = 1$.

For $0 < c < 2$, the above proof works equally well. For $c = 2$, we provide a construction where the statement is no longer true.

It suffices to find a function $f\colon \mathbb{C} \to \mathbb{R}$ such that

$$|f(z_2) - f(z_1)| > |z_2 - z_1|^2$$

for all complex numbers z_1 and z_2.

Let $g(z)\colon [0,1]+[0,1]i \to [0,1]$ be the inverse Hilbert curve. The preimage of any interval $\left[\frac{n}{2^{2k}}, \frac{n+1}{2^{2k}}\right]$ is a square of side length $\frac{1}{2^k}$ that is adjacent to the preimage of $\left[\frac{n+1}{2^{2k}}, \frac{n+2}{2^{2k}}\right]$. This means the preimage of any length ℓ interval is contained in a width $4\sqrt{\ell}$ square. This means

$$8\sqrt{|g(z_2)-g(z_1)|} > |z_2 - z_1|,$$

implying that some sufficiently large multiple of $g(z)$, say $h(z)$, satisfies the desired inequality over its domain.

To extend the domain of this solution to all complex numbers, partition the complex plane into countably many unit squares, copy $h(z)$ onto each unit square, and space the images of each unit square sufficiently far apart on the real number line.

§3.2b Marking scheme

For all solutions, the following are *not awarded marks*:

- Proving the statement for all continuous labellings.
- Assuming there exists a labelling $f : \mathbb{R}^2 \to \mathbb{R}$ such that $|f(x)-f(y)| \geq |x-y|$ for the sake of contradiction.
- Proving f is injective.

For solutions similar to the official solution, the following items are available but **not additive**:

- **1 point** for examining the pre-image of an interval.
- **1 point** for examining a 2-dimensional set (as opposed to a lattice or a countable number of lines) and mentioning area.
- **2 points** for proving $f^{-1}([a,b])$ is contained in a set of area $C(b-a)^2$ for some constant C. For reference, this set will (usually) be either a square or a disk.
- **5 points** for proving the above quadratic bound **and** introducing a harmonic covering of $\mathbb{R}$.
- **5 points** for proving $f^{-1}([a,b])$ is contained in a set of area ε for any $\varepsilon > 0$.
- **7 points** for a complete solution.

Solutions using the advanced theory of Lipschitz functions are scored as follows, with all marks being additive.

- **+1 point** for showing f^{-1} is 1-Lipschitz **and** surjective on its domain S.
- **+2 points** for a suitable extension of f^{-1} to $g : \mathbb{R}^2 \to \mathbb{R}^2$. Graders should look out for implicit assumptions that S is composed of a finite number of intervals/dense or not dense/Borel/Lebesgue-measurable/e.t.c.
- **+1 point** for covering the preimage of $f^{-1}(\mathbb{R})$ with a countable number of discs with finite area, awarded **only if** the contestant hits the above 2 points.

For all solutions which are incomplete with errors, the following deductions apply and are all additive:

- **-1 point** for only partitioning $\mathbb{R}^+$ into intervals.
- **-1 point** for mentions of the "area" of $f^{-1}([a,b])$ (instead of that of a superset).

§3.3 USEMO 3 — proposed by Maxim Li

Problem statement

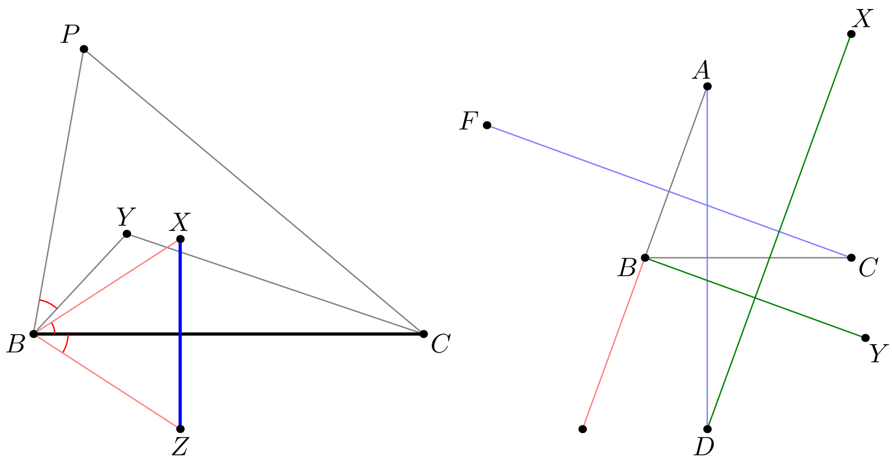
Canmoo is trying to do constructions, but doesn't have a ruler or compass. Instead, Canmoo has a device that, given four distinct points A, B, C, P in the plane, will mark the isogonal conjugate of P with respect to triangle ABC, if it exists. Show that if two points are marked on the plane, then Canmoo can construct their midpoint using this device, a pencil for marking additional points, and no other tools.

(Recall that the *isogonal conjugate* of P with respect to triangle ABC is the point Q such that lines AP and AQ are reflections around the bisector of $\angle BAC$, lines BP and BQ are reflections around the bisector of $\angle CBA$, lines CP and CQ are reflections around the bisector of $\angle ACB$. Additional points marked by the pencil can be assumed to be in general position, meaning they don't lie on any line through two existing points or any circle through three existing points.)

§3.3a Solution

We assume Canmoo can mark points in arbitrarily general position.

We first prove two claims showing that reflection around a point is possible. We will only use the second claim in what proceeds (so with the second claim proven, we can forget about the first one.)



Claim — Given any three points X, B, C, Canmoo can construct the reflection of X over $\overline{BC}$.

Proof. Pick another point P, and let Y be the isogonal conjugate of X in $\triangle PBC$, and Z the isogonal conjugate of X in $\triangle YBC$. Then

$$\measuredangle CBX = \measuredangle YBP = \measuredangle ZBC, \qquad \text{and} \qquad \measuredangle BCX = \measuredangle YCP = \measuredangle ZCB.$$

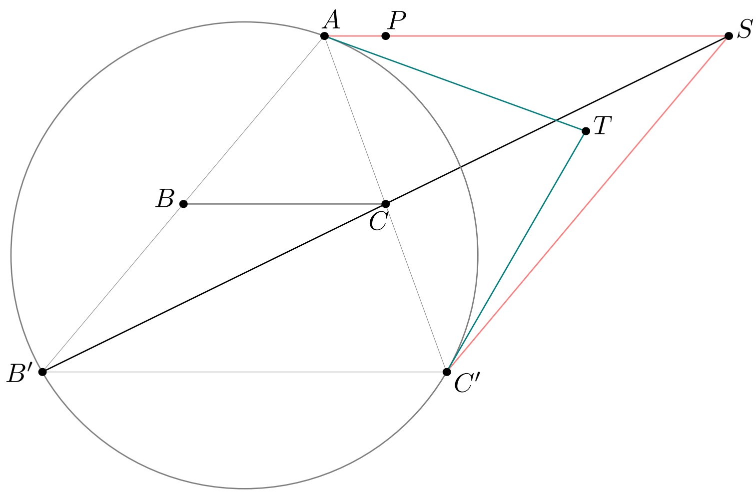
Thus, Z is the reflection of X over $\overline{BC}$. $\square$

Claim — Given any two points A and B, Canmoo can reflect A over B.

Proof. Pick another point C. Let D and F denote the reflections of A and C over $\overline{BC}$ and $\overline{AB}$. Then reflect D over $\overline{CF}$ to X, and B over $\overline{DX}$ to Y, so that $\overline{BY} \perp \overline{AB}$. Finally, the reflection of A over $\overline{BY}$ is the reflection of A over B. $\square$

Claim — Given any three points A, B, C, Canmoo can construct the point P such that $\overline{AP} \parallel \overline{BC}$ and $\angle APC = 90^\circ$.

Proof. WLOG assume $\angle B \neq 90^\circ$. (If not, replace B with its reflection over C.) Reflect A over B and C to get B' and C', respectively. Reflect B' over C to get S, so $AB'C'S$ is a parallelogram which is not a rectangle. Since $\overline{AP} \parallel \overline{B'C'}$, the isogonal of $\overline{AP}$ with respect to $\angle B'AC'$ is the tangent to the circumcenter of $\triangle ABC$ at A; and hence the isogonal conjugate of S is $T := \overline{AA} \cap \overline{C'C'}$, the intersection of the tangents at A and C', as shown below.



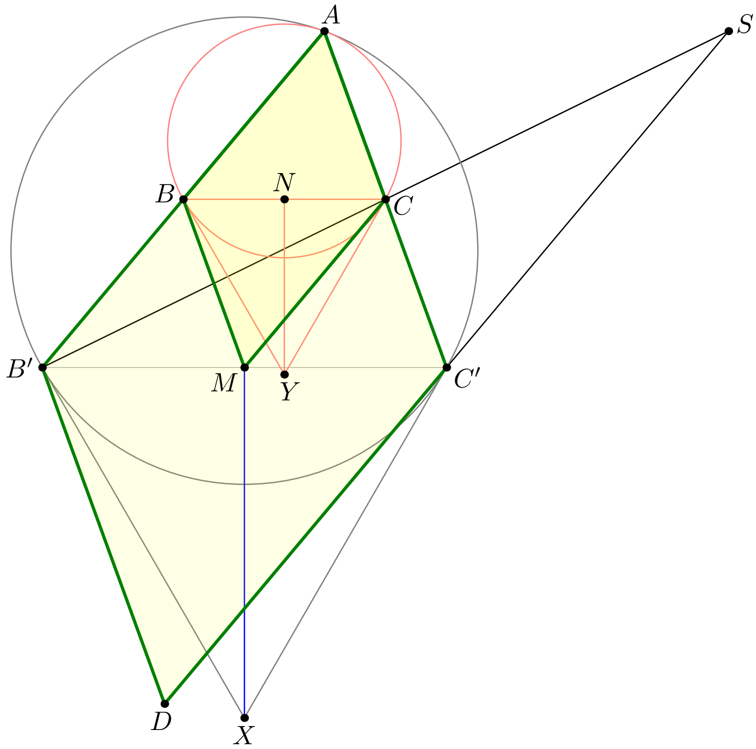
Now take the isogonal conjugate of T with respect to $\triangle ABC$. It still lies on line $\overline{AS}$, but because $\angle TCA = 90^\circ$ we have $\measuredangle PCB = 90^\circ$ too, hence $\measuredangle APC = 90^\circ$. So P is exactly the point desired. $\square$

Claim — Given any three points A, B, C, Canmoo can construct the foot from A to BC.

Proof. Construct the point P in the third claim, and WLOG assume none of the angles in the problem are 90° (otherwise, reflect B or C over each other), so P is distinct from A. Then apply the construction again to BAP to get the desired foot. $\square$

We are ready to tackle the main construction. Let B and C be the initial two given points, and pick a third point A and assume $\angle A \neq 90^\circ$. As in the proof of the third claim, reflect A over B, C to get B' and C', and reflect B' over C to S.

Then let D be the reflection of S across C'; hence $AB'DC'$ is a parallelogram. The isogonal conjugate of D with respect to $\triangle AB'C'$ is therefore the intersection X of the tangents to $(AB'C')$ at B' and C'. Then taking the foot from X to $\overline{B'C'}$ gives the foot from M to $\overline{B'C'}$.



In particular, $ABMC$ is a parallelogram, and so again we may take the isogonal conjugate of M with respect to $\triangle ABC$ to obtain the point $Y = \overline{BB} \cap \overline{CC}$ which is the intersection of the tangents to B and C at (ABC). Finally, taking the foot from Y to $\overline{BC}$ gives the desired midpoint.

§3.3b Marking scheme

For the most part, solutions could be read on a case-by-case basis, because not many solutions have nontrivial progress. We stanardize the following **non-additive** benchmarks:

- **1 point** for showing X can be reflected over YZ.
- **0 points** for showing X can be reflected over a point Y.
- **3 points** for being able to obtain the feet of the altitudes in a triangle
- **3 points** for being able to obtain the circumcenters of triangles.

§3.4 USEMO 4 — proposed by Ankan Bhattacharya

Problem statement

Let ABC be an acute triangle with orthocenter H. Points A_1, B_1, C_1 are chosen in the interiors of sides BC, CA, AB, respectively, such that $\triangle A_1B_1C_1$ has orthocenter H. Define $A_2 = \overline{AH} \cap \overline{B_1C_1}$, $B_2 = \overline{BH} \cap \overline{C_1A_1}$, and $C_2 = \overline{CH} \cap \overline{A_1B_1}$.

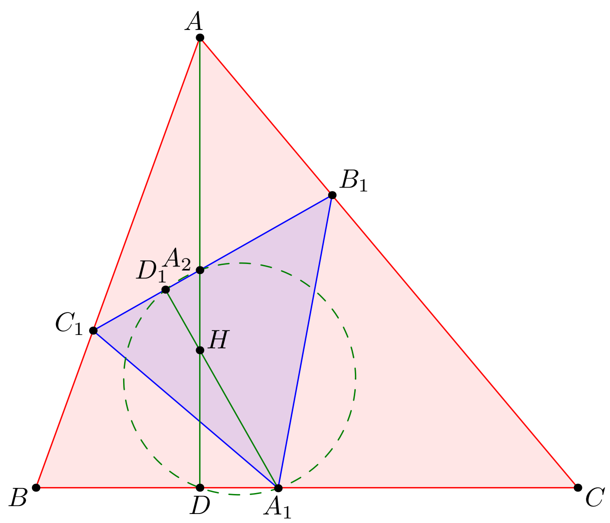
Prove that triangle $A_2B_2C_2$ has orthocenter H.

§3.4a Solution

We present four solutions.

¶ Power of a point solution, by Nikolai Beluhov. In this solution, all lengths are signed. Let $\triangle DEF$ be the orthic triangle of $\triangle ABC$, and $\triangle D_1E_1F_1$ be the orthic triangle of $\triangle A_1B_1C_1$. We define two common quantities, through power of a point:

$$k \coloneqq HA \cdot HD = HB \cdot HE = HC \cdot HF.$$
$$k_1 \coloneqq HA_1 \cdot HD_1 = HB_1 \cdot HE_1 = HC_1 \cdot HF_1.$$



Because quadrilateral $A_2D_1DA_1$ is concyclic (with circumdiameter $\overline{A_1A_2}$), by power of a point, we get

$$HA_2 \cdot HD = HD_1 \cdot HA_1 = k_1$$
$$\implies HA_2 = \frac{k_1}{HD} = \frac{k_1}{k} \cdot HA.$$

Since k_1/k is fixed, a symmetric argument now gives

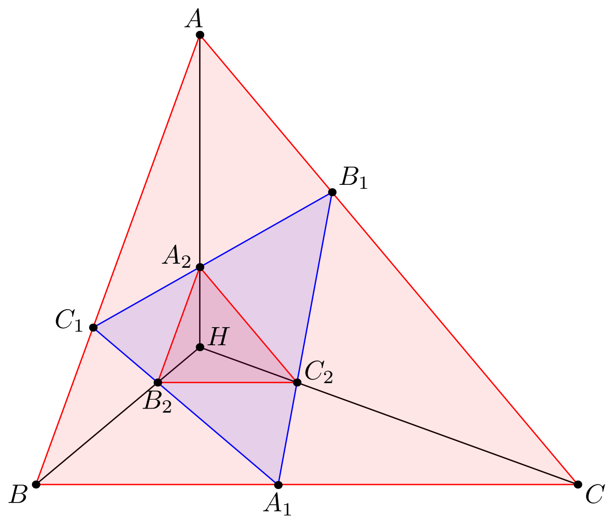
$$\frac{HA_2}{HA} = \frac{HB_2}{HB} = \frac{HC_2}{HC} = \frac{k_1}{k}.$$

Therefore, H is the center of a homothety mapping $\triangle A_2B_2C_2$ to $\triangle ABC$. In particular, it is also the orthocenter of $\triangle A_2B_2C_2$.

¶ Author's ratio-based solution. We are going to prove:

Claim — We have $\overline{B_2C_2} \parallel \overline{BC}$.

Proof. Refer to the diagram below.



Note that

$$\begin{aligned} \frac{C_1A_2}{A_2B_1} &= \frac{[AC_1H]}{[AB_1H]} = \frac{AC_1 \cdot d(H, \overline{AB})}{AB_1 \cdot d(H, \overline{AC})} \\ &= \frac{AC_1/HC}{AB_1/HB} = \frac{HB}{HC} \cdot \frac{\sin \angle AB_1C_1}{\sin \angle AC_1B_1} \\ &= \frac{HB}{HC} \cdot \frac{\sin \angle BHA_1}{\sin \angle CHA_1} = \frac{[HBA_1]}{[HCA_1]} = \frac{BA_1}{A_1C}. \end{aligned}$$

Similarly, $\frac{A_1B_2}{B_2C_1} = \frac{CB_1}{B_1A}$ and $\frac{B_1C_2}{C_2A_1} = \frac{AC_1}{C_1B}$. Hence,

$$[BB_2C] = [BC_1C] \cdot \frac{B_2A_1}{C_1A_1} = [BAC] \cdot \frac{B_2A_1}{C_1A_1} \cdot \frac{C_1B}{AB} = [ABC] \cdot \frac{B_1C}{AC} \cdot \frac{C_1B}{AB}.$$

Similarly, $[BC_2C]$ also equals this quantity, so $\overline{B_2C_2} \parallel \overline{BC}$ and $\overline{A_2H} \perp \overline{B_2C_2}$. □

Repeating this we see that H is the orthocenter of $\triangle A_2B_2C_2$, as wanted.

Remark. In the first equality chain, we obtained

$$[AC_1H] \cdot [CA_1H] = [AB_1H] \cdot [BA_1H].$$

Similarly, $[BC_1H] \cdot [CB_1H]$ also equals this quantity, and so we see that

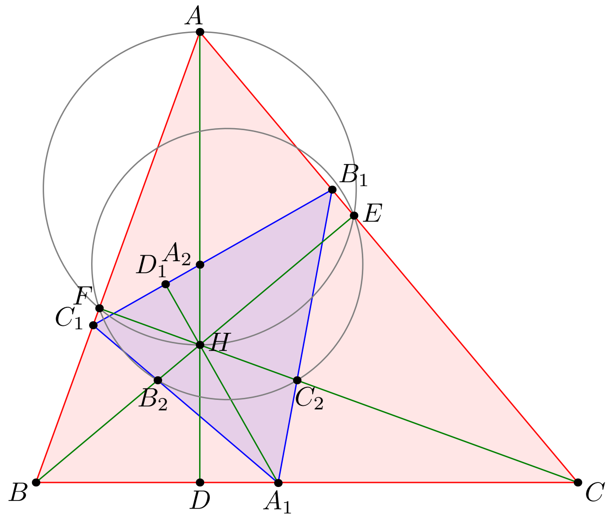
$$\frac{\sin \angle BHC_1 \cdot \sin \angle CHB_1}{AH \cdot A_1H} = \frac{\sin \angle CHA_1 \cdot \sin \angle AHC_1}{BH \cdot B_1H} = \frac{\sin \angle AHB_1 \cdot \sin \angle BHA_1}{CH \cdot C_1H}.$$

Intuitively, this result is symmetric under swapping $\triangle ABC$ and $\triangle A_1B_1C_1$, and doesn't depend upon $\triangle A_1B_1C_1$ being inscribed in $\triangle ABC$, in the sense that scaling $\triangle ABC$ or $\triangle A_1B_1C_1$ by any factor (with center H) preserves this property. Thus, this offers an intuitive explanation for why "swapping" the triangles preserves the common orthocenter.

It might be possible to adapt this into a phantom-point approach to directly settle the problem, but I don't see how to do that.

¶ Alternative finish to first solution by inversion. Define D, D_1, E, E_1, F and F_1 as in the first solution. As before, the quadrilateral $A_2D_1DA_1$ is concyclic (with circumdiameter $\overline{A_1A_2}$).

Now, rather than using power of a point, consider the negative inversion with center H and radius $\sqrt{HA_1 \cdot HD}$. Since $A_2D_1DA_1$ was cyclic and similarly, this inversion maps E to B_2 and F to C_2. Hence the circle (EHF) maps to the line B_2C_2.



But the circle (EHF) is symmetric with respect to line DH (the center of the circle is the midpoint of $\overline{AH}$), so the circle (EHF) must map to a line perpendicular to DH. It follows that $B_2C_2 \perp DH$, or that $A_2H \perp B_2C_2$, as needed.

¶ Alternative finish to first solution using Reim's theorem. Refer to the preceding figure. From the cyclic quadrilaterals $A_2D_1DA_1$ and similar in the first solution we also have

$$HC_2 \cdot HF = HC_1 \cdot HF_1 = HB_1 \cdot HE_1 = HB_2 \cdot HE.$$

From this it follows that B_2C_2EF is cyclic.

We also know $BCEF$ is cyclic and hence by Reim's Theorem we get that $B_2C_2 \parallel BC$, which implies the result.

§3.4b Marking scheme

General rules:

- As usual, incomplete computational approaches earn partial credits only based on the amount of synthetic progress which is made.
- No points are awarded for just drawing a diagram or simple observations.
- No points are deducted for configuration issues (such as not using directed angles) and minor typos.
- If the student has approaches from more than one of the solutions, they get the maximum of all possible markings.

Items worth zero points

0 points for stating the problem is equivalent to prove that $AB \parallel A_2B_2$, or that ABC and $A_2B_2C_2$ are homothetic with center H.

Power of a point solution, by Nikolai Beluhov

The following partial items are available and are **additive**:

(a) +1 point Proving that A_1, D, A_2 and D_1 lie on a circle.

(b) +1 point Showing it is sufficient to prove $\frac{HA_2}{HA}$ is constant (or $\frac{HA_2}{HA} = \frac{HB_2}{HB} = \frac{HC_2}{HC}$, or anything else identical).

Author's ratio-based solution

The following partial items are available and are **additive**:

(a) +2 points Proving that $\dfrac{C_1A_2}{A_2B_1} = \dfrac{BA_1}{A_1C}$.

(b) +1 point Showing it is sufficient to prove $[BB_2C] = [BC_2B]$ (or that the distances from B_2 and C_2 to BC are equal).

Inversion (third solution)

The following partial items are available and are **additive**, though note that altogether they form a complete solution:

(a) +1 point Proving that A_1, D, A_2 and D_1 lie on a circle.

(b) +1 point Stating that the line DH and the circle (EHF) are orthogonal (or that the center of (EHF) lies on DH).

(c) +1 point Reducing the problem (e.g. by inversion) to showing that the line DH and the circle (EHF) are orthogonal- (or that the center of (EHF) lies on DH).

Angle chase with power of a point (fourth solution)

The following partial items are available, but only **(a)** and **(b)** are additive. Note that altogether **(b)** and **(c)** form a complete solution.

(a) 1 point Proving that A_1, D, A_2 and D_1 lie on a circle.

(b) 1 point Showing it is sufficient to prove that B_2, C_2, E and F lie on a circle.

(c) 2 points Proving that B_2, C_2, E and F lie on a circle.

Any complete solution is worth **7 points**.

§3.5 USEMO 5 — proposed by Nikolai Beluhov

Problem statement

Let $n \geq 2$ be an integer. A cube of size $n \times n \times n$ is dissected into n^3 unit cubes. A nonzero real number is written at the center of each unit cube so that the sum of the n^2 numbers in each slab of size $1 \times n \times n$, $n \times 1 \times n$, or $n \times n \times 1$ equals zero. (There are a total of $3n$ such slabs, forming three groups of n slabs each such that slabs in the same group are parallel and slabs in different groups are perpendicular.)

Could it happen that some plane in three-dimensional space separates the positive and the negative written numbers? (The plane should not pass through any of the numbers.)

§3.5a Solution

We show this can never happen.

Suppose, for the sake of contradiction, that such a plane α did exist. Let $Oxyz$ be a Cartesian coordinate system whose origin O lies in α and whose axes are parallel to the edges of our cube. Let the equation of α in this coordinate system be $ax + by + cz = 0$. Without loss of generality, all positive written numbers lie in the half-space $ax+by+cz > 0$ relative to α and all negative written numbers lie in the half-space $ax + by + cz < 0$ relative to α.

For all $i \in \{1, 2, \dots, n^3\}$, let r_i be the number written at point (x_i, y_i, z_i). Then for all i we have that $(ax_i + by_i + cz_i)r_i > 0$.

Therefore,

$$\begin{aligned}
0 &= a \cdot 0 + b \cdot 0 + c \cdot 0 \\
&= \sum_x ax \cdot \left(\sum_{i:x_i=x} r_i \right) + \sum_y by \cdot \left(\sum_{i:y_i=y} r_i \right) + \sum_z cz \cdot \left(\sum_{i:z_i=z} r_i \right) \\
&= \sum_i ax_i r_i + \sum_i by_i r_i + \sum_i cz_i r_i \\
&= \sum_i (ax_i + by_i + cz_i) r_i \\
&> 0.
\end{aligned}$$

We have arrived at a contradiction. The solution is complete.

Remark. The so-called Farkas lemma guarantees that if there wasn't a valid labeling, then we can combine the given inequalities to obtain a contradiction; therefore there is a sense in which an inequality-based solution like the one above is *a priori* promised to exist if the answer is no.

Remark (Neil Kolekar)**.** More generally, if P, Q, R are nonconstant polynomials with real coefficients, then from $0 = \sum_{i:x_i=x} r_i$ being true for each x (and similarly), we should have

$$0 = \sum_x P(x) \cdot \left(\sum_{i:x_i=x} r_i \right) + \sum_y Q(y) \cdot \left(\sum_{i:y_i=y} r_i \right) + \sum_z R(z) \cdot \left(\sum_{i:z_i=z} r_i \right).$$

> This means the same proof would work in general when the hyperplanes $ax + by + cz = 0$ are replaced by more general surfaces of the form
> $$P(x) + Q(y) + R(z) = 0.$$

§3.5b Marking scheme

In what follows, we let $f(x, y, z)$ be the number written at (x, y, z) with $1 \leq x, y, z \leq n$. Also let the required plane be $ax + by + cz - d = 0$.

Partial items for 0+ solutions

The following partial items apply for incomplete solutions and are **additive**.

- **0 points** for claiming the answer
- **0 points** for making the false claim that a plane cannot pass through all slabs
- **0 points** for solving the problem for $n = 2$ or $n = 3$
- **+2 points** for considering an expression of the form
$$\sum x f(x, y, z)$$
and showing that it is equal to zero.
- **+2 points** for considering the product $(ax + by + cz - d)f(x, y, z)$ or something similar

Complete solutions

In case of a complete solution:

- **No deduction** for not explicitly considering at most one of the following cases:
  1. $f(x, y, z)$ and $ax + by + cz - d$ have the *same* sign
  2. $f(x, y, z)$ and $ax + by + cz - d$ have *opposite* sign

§3.6 USEMO 6 — proposed by Kaixin Wang

Problem statement

Let $n \geq 2$ be a fixed integer.

(a) Determine the largest positive integer m (in terms of n) such that there exist complex numbers $r_1, \dots, r_n$, not all zero, for which

$$\prod_{k=1}^{n}(r_k+1) = \prod_{k=1}^{n}(r_k^2+1) = \cdots = \prod_{k=1}^{n}(r_k^m+1) = 1.$$

(b) For this value of m, find all possible values of

$$\prod_{k=1}^{n}(r_k^{m+1}+1).$$

§3.6a Solution

For part (a) the answer is $m = 2^n - 2$; for part (b) the answer is 2^n.

¶ Construction for (a). For $m = 2^n - 2$, fix $\omega := \exp(\frac{2\pi i}{2^n-1})$ and set

$$r_j = \omega^{2^j} \qquad j = 1, 2, \dots, m = 2^n - 2.$$

We can expand to see that the

$$\prod_{k=1}^{n}(r_k^j+1) = \sum_{j=0}^{2^n-1} \omega^j = 1, \qquad j = 1, 2, \dots, 2^n - 2.$$

¶ Bound for (a). It remains to show when $m = 2^n - 1$ we must have $r_1 = \cdots = r_n = 0$. For each nonempty subset S of $\{1, \dots, n\}$, define

$$\Pi_S := \prod_{k \in S} r_k.$$

Then the problem condition, when expanded, states that

$$0 = -1 + \prod_{k=1}^{n}(r_k^j+1) = \sum_{\varnothing \neq S \subseteq \{1,\dots,n\}} \Pi_S^j \qquad j = 1, 2, \dots, 2^n - 1.$$

But view the Π_S as $2^n - 1$ variables. Then the first $2^n - 1$ power sums of the Π_S all vanish, and hence (say, by Newton sums) it follows that every Π_S must be zero. In particular, all the variables are zero as well.

¶ Proof for (b). Fix $m = 2^n - 2$ and

$$\omega := \exp\left(\frac{2\pi i}{2^n-1}\right).$$

We are going to prove that:

Claim — Every r_i is a power of ω.

Proof. If we apply Newton sums as we did before, with the identity

$$0 = -1 + \prod_{k=1}^{n} (r_k^j + 1) = \sum_{\varnothing \neq S \subseteq \{1,\dots,n\}} \Pi_S^j \qquad j = 1, 2, \dots, 2^n - 1.$$

then we get that all the elementary symmetric polynomials in Π_S vanish, except for the last one. In other words, the polynomial identity

$$X^{2^n-1} - c = \prod_{\varnothing \neq S \subseteq \{1,\dots,n\}} (X - \Pi_S).$$

should hold for some c. We saw already (in (a)) that if $c = 0$ then $r_i = 0$ for all i, so assume $c \neq 0$, and that $r_i \neq 0$ for all i.

Let λ be any complex $(2^n - 1)$th root of c. Then the factorization of the left-hand polynomial over $\mathbb{C}$ is given exactly by

$$X^{2^n-1} - c = \prod_{j=0}^{2^n-2} \left(X - \lambda\omega^j\right).$$

Hence, we have the equality of unordered sets

$$\left\{\Pi_S \mid \varnothing \neq S \subseteq \{1, \dots, n\}\right\} = \left\{\lambda\omega^j \mid 0 \le j \le 2^n - 2\right\}.$$

In particular,

$$r_1 = \frac{r_1 r_2}{r_2} = \frac{\Pi_{\{1,2\}}}{\Pi_{\{2\}}} = \frac{\lambda\omega^{n_{12}}}{\lambda\omega^{n_2}} = \omega^{n_{12}-n_2}$$

for some integers n_{12} and n_2 (whose values are unimportant); so r_1 is a $(2^n - 1)$th root of unity, and similarly so is every r_i. $\square$

Hence

$$\prod_{i=1}^{n} (r_i^m + 1) = \prod_{i=1}^{n} \left(r_i^{2^n-1} + 1\right) = \prod_{i=1}^{n} (1 + 1) = 2^n$$

is the only possible value of the product requested in (b).

Remark. An interesting problem is to characterize all $(r_1, \dots, r_n)$. The author has not solved that yet.

§3.6b Marking scheme

Part (a) will be graded out of 5 points and part (b) will be graded out of 2 points. The scores of the two parts will be added for the final score out of 7.

Scoring for (a)

For incomplete solutions, the following items are available but **not additive**.

- **1 point** for the correct answer of $m = 2^n - 2$ in part (a)
- **0 points** for expanding the product and rewriting it as a sum.
- **2 points** for expanding the product and rewriting it as a sum, using Newton sums or Vandermonde Determinant to argue that the answer is at most $2^n - 2$.
- **2 points** For a correct construction and answer.
- **5 points** For completely solving part (a).

For complete solutions, the following deductions apply, and are **additive**.

- **-1 points** for wrong answer.
- **-1 points** For another mistake.

Scoring for (b)

- **1 point** for correct answer
- **2 points** for correct solution.

4 Statistics

§4.1 Summary of scores for USEMO 2023

N	70	1st Q	2	Max	36
μ	8.94	Median	7	Top 3	28
σ	7.96	3rd Q	14	Top 12	15

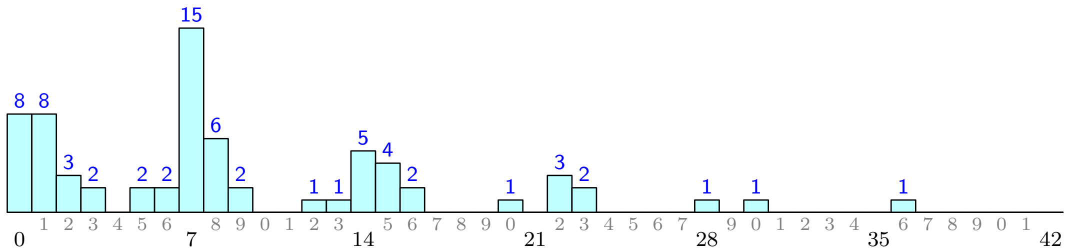
§4.2 Problem statistics for USEMO 2023

	P1	P2	P3	P4	P5	P6
0	15	59	53	37	65	62
1	14	3	15	3	0	5
2	2	2	0	0	0	0
3	1	0	0	0	0	0
4	0	0	0	0	0	0
5	4	0	0	0	0	0
6	2	0	0	1	0	2
7	32	6	2	29	5	1
Avg	3.96	0.70	0.41	3.03	0.50	0.34
QM	5.03	2.09	1.27	4.57	1.87	1.34
#5+	**38**	**6**	**2**	**30**	**5**	**3**
%5+	%54.3	%8.6	%2.9	%42.9	%7.1	%4.3

§4.3 Rankings for USEMO 2023

Sc	Num	Cu	Per	Sc	Num	Cu	Per	Sc	Num	Cu	Per
42	0	0	0.00%	28	1	3	4.29%	14	5	20	28.57%
41	0	0	0.00%	27	0	3	4.29%	13	1	21	30.00%
40	0	0	0.00%	26	0	3	4.29%	12	1	22	31.43%
39	0	0	0.00%	25	0	3	4.29%	11	0	22	31.43%
38	0	0	0.00%	24	0	3	4.29%	10	0	22	31.43%
37	0	0	0.00%	23	2	5	7.14%	9	2	24	34.29%
36	1	1	1.43%	22	3	8	11.43%	8	6	30	42.86%
35	0	1	1.43%	21	0	8	11.43%	7	15	45	64.29%
34	0	1	1.43%	20	1	9	12.86%	6	2	47	67.14%
33	0	1	1.43%	19	0	9	12.86%	5	2	49	70.00%
32	0	1	1.43%	18	0	9	12.86%	4	0	49	70.00%
31	0	1	1.43%	17	0	9	12.86%	3	2	51	72.86%
30	1	2	2.86%	16	2	11	15.71%	2	3	54	77.14%
29	0	2	2.86%	15	4	15	21.43%	1	8	62	88.57%
								0	8	70	100.00%

§4.4 Histogram for USEMO 2023



§4.5 Full stats for USEMO 2023

Rank	P1	P2	P3	P4	P5	P6	Σ
1.	7	7	1	7	7	7	36
2.	7	2	7	7	7	0	30
3.	7	7	0	7	7	0	28
4.	7	2	7	7	0	0	23
4.	6	1	1	7	7	1	23
6.	7	7	1	7	0	0	22
6.	7	7	1	7	0	0	22
6.	7	0	0	7	7	1	22
9.	7	7	0	6	0	0	20
10.	7	0	1	7	0	1	16
10.	7	0	1	7	0	1	16
12.	7	7	1	0	0	0	15
12.	7	0	1	7	0	0	15
12.	7	0	1	7	0	0	15
12.	7	0	1	7	0	0	15
16.	7	0	0	7	0	0	14
16.	7	0	0	7	0	0	14
16.	7	0	0	7	0	0	14
16.	7	0	0	7	0	0	14
16.	7	0	0	7	0	0	14
21.	6	0	0	7	0	0	13
22.	5	0	0	7	0	0	12
23.	7	0	1	0	0	1	9
23.	1	0	1	7	0	0	9
25.	7	1	0	0	0	0	8
25.	1	0	0	7	0	0	8
25.	1	0	0	7	0	0	8
25.	1	0	0	7	0	0	8
25.	1	0	0	7	0	0	8
25.	1	0	0	1	0	6	8
31.	7	0	0	0	0	0	7
31.	7	0	0	0	0	0	7
31.	7	0	0	0	0	0	7
31.	7	0	0	0	0	0	7
31.	7	0	0	0	0	0	7
31.	7	0	0	0	0	0	7

Rank	P1	P2	P3	P4	P5	P6	Σ
31.	7	0	0	0	0	0	7
31.	7	0	0	0	0	0	7
31.	7	0	0	0	0	0	7
31.	7	0	0	0	0	0	7
31.	7	0	0	0	0	0	7
31.	0	0	0	7	0	0	7
31.	0	0	0	7	0	0	7
31.	0	0	0	7	0	0	7
31.	0	0	0	7	0	0	7
46.	5	0	1	0	0	0	6
46.	0	0	0	0	0	6	6
48.	5	0	0	0	0	0	5
48.	5	0	0	0	0	0	5
50.	3	0	0	0	0	0	3
50.	1	1	1	0	0	0	3
52.	2	0	0	0	0	0	2
52.	2	0	0	0	0	0	2
52.	1	0	0	1	0	0	2
55.	1	0	0	0	0	0	1
55.	1	0	0	0	0	0	1
55.	1	0	0	0	0	0	1
55.	1	0	0	0	0	0	1
55.	1	0	0	0	0	0	1
55.	1	0	0	0	0	0	1
55.	0	0	1	0	0	0	1
55.	0	0	0	1	0	0	1
63.	0	0	0	0	0	0	0
63.	0	0	0	0	0	0	0
63.	0	0	0	0	0	0	0
63.	0	0	0	0	0	0	0
63.	0	0	0	0	0	0	0
63.	0	0	0	0	0	0	0
63.	0	0	0	0	0	0	0
63.	0	0	0	0	0	0	0