Themes and Heuristics in Analysis-Flavored Olympiad Problems

Draft Version 1.0, May 6, 2017

Send comments / corrections / suggestions to farrellw@mit.edu

By

Farrell Eldrian Wu

A research paper
submitted in partial fulfillment of
the requirements for the course
Research/Capstone Project
May 2017

Abstract

This paper identifies themes and heuristics involved in the solutions to analytic-flavored olympiad problems, which is a rising trend in mathematical olympiads but has received little attention both in literature and in discussions. It deals with a class of problems that focuses on the long-term and large-scale behavior of sequences as well as other problems with implicit sequences or functions. The author selected thirty-six problems that have at least one solution involving analytic reasoning and examined the main ideas in their solutions. The main ideas were organized into twenty-four heuristics, which were later condensed into eleven and subsequently categorized under three main themes, namely *taking a global view*, *refocusing*, and*investigating dynamics or processes*. This article presents an exposition into these themes, providing an informal, intuition-based discussion to the heuristics involved. The discussion is based on past mathematical olympiad problems that incorporate the specific heuristics as their main ideas.

Keywords: mathematical analysis, problem-solving, problem-solving heuristics, analytical reasoning, mathematical olympiads, sequences, functions, convergence

Contents

List of abbreviations

Throughout this paper, the following abbreviations are used to label problem sources:

APMO X PY Problem Y of the X Asian-Pacific Mathematics Olympiad

Benelux X PY Problem Y of the X Benelux Mathematical Olympiad

Canada MO X PY Problem Y of the X Canadian Mathematics Olympiad

IMO X PY Problem Y of the X International Mathematical Olympiad

ISL X (abbr.) Y International (Mathematical Olympiad) Shortlist Problem Y under **A**lgebra, **N**umber theory, **C**ombinatorics, **G**eometry

MOSC Mathematical Olympiad Summer Camp (Philippines IMO Team Selection)

PEM Handout PY Problem Y of the handout *Problems on Convergence of Sequences* from the September 3, 2016 session, advanced group, *Program for Excellence in Mathematics*

PFTB Example X Example X of *Problems from the Book* Chapter 17

S#, M# Small idea with score #, major idea with score #

USA TST X PY Problem Y of the X United States of America Team Selection Test

USA TSTST X PY Problem Y of the X United States of America Selection Test for the Team Selection Group

USAMO X PY Problem Y of the X United States of America Mathematical Olympiad.

USAMTS X/Y/Z Problem Z of the Yth round of the Xth United States of America Mathematical Talent Search

1 Introduction

A considerable portion of problems (approximately 10%, increasing recently) featured in mathematical olympiads may be described as analysis-flavored. While most such problems are within algebra and involve examining the long-term behavior of sequences that fit the given conditions, similar ideas emerge in problems within number theory or combinatorics, whose solutions identify a pertinent function and understands its large-scale, rather than term-specific, properties.

According to Wolfram MathWorld, analysis "is the systematic study of real and complex-valued continuous functions" (Renze and Wesstein, n.d.). While analysis is a large field composed of numerous sub-topics such as calculus, integration, and topology, analytical concepts in olympiad problems cover a much narrower domain, focusing on the behavior of sequences or functions, commonly known as calculus. An analytical approach to solving an olympiad problem involves understanding a sequence or function from a global perspective, typically foregoing details in order to focus on large-scale behavior. In these solutions, techniques and arguments involved are "primarily simple, bare-hands deductions based on inequalities, size considerations, infinitesimal, or asymptotic heuristics" (Wang, personal communication, March 27, 2017).

Analytical reasoning may be considered as a thought framework where one considers a notion of sizes or estimates (such as a quantity being sufficiently large) that is typically not done in more rigid problems (Chen, personal communication, March 9, 2017). It may involve first understanding the scenario locally and then generalizing to understand the whole picture, also known as the local-global principle.

1.1 Background

The term *analysis-flavored* to describe a particular class of problems is not in wide use: the only two explicit mentions found as of writing are the second section (*Analytic-flavored stuff*) of the article *MOP Experiment* (Wang, 2014) and the seventeenth chapter (*At the Border of Analysis and Number Theory*) of the book *Problems from the Book* (Andreescu and Dopinescu, 2010).

The article *MOP Experiment* compiles problems from themes typically not covered in olympiad handouts, one which is analytical considerations in number theory and sequences. In a personal correspondence, the author categorizes mathematical topics and arguments as analytical, structured (algebra), or visual (geometry). The main idea in analytical reasoning is the understanding of structure versus noise; the latter is ignored in favor of the former. While few problems are purely analytical, some have primary analytic techniques.

Analytical reasoning leans heavily towards global (the subtleties of how the pieces interact with each other) rather than local (how individual points or pieces or a small portion works) methods. In the context of problem solving, this implies that in analysis-flavored problems, the value of a function at $x = 2$, for example, is far less important than its overall rate of change. Such reasoning may be done purely based on intuition, size estimates, and an understanding of the function in general, without having to examine the details (Wang, personal correspondence, March 27, 2017).

The chapter *At the Border of Analysis and Number Theory* in *Problems from the Book* focuses on the application of analytic techniques in number theory. The solutions to the problems featured all exhibit convergent integer sequences, which must eventually reach its limit which is a fixed value. It offers analytical reasoning (particularly, that of bounded or convergent integer sequences) as an alternative to difficult-to-design elementary solutions to several number theory problems. It claims that the difficulty is in manipulating and understanding the problem sufficiently to identify the underlying sequence (Andreescu and Dopinescu, 2010).

1.2 Objectives

This paper intends to deconstruct the class of analytic-flavored olympiad problems by examining and categorizing the heuristics involved in the solutions to thirty-six analytic-flavored problems. Typically, such problems are categorized into detail-based categories such as sequences, functions, or sets; sometimes as miscellaneous problem-solving. According to the article *Some Thoughts on Olympiad Material Design* (Chen, 2017), these categories do not adequately describe nor convey the understanding of the style of thought and arguments involved.

As this category of problems encompasses many different types, from typical sequence-based problems (such as IMO 2010 Problem 6) to combinatorics problems (such as IMO 2012 Problem 3), identifying problems belonging in this category may be difficult. Although on a surface level, the solutions look different, they are related to the small number of main ideas, which are far more important than technical details (Chen, 2017). This variety in problem statement styles may explain the lack of literature as well as the lack of awareness on this approach. Thus, another objective of this article is to expound on the vague notion of analytical thinking in the context of olympiad problems.

In his article, Chen further stresses the importance of proper “olympiad taxonomy”, or the classification of olympiad problems. Typically, problems are sorted based on “particular technical details” that appear in the problem, which do not necessarily represent the main ideas. His recommendation, instead, is to “classify the main ideas into categories and themes” in order to exemplify connections between problems. Therefore, this article classifies the selected analytic-flavored olympiad problems based on their key heuristics and main ideas.

1.3 Organization and preparation

With the above guidelines in consideration, the author first skimmed through problems from the following sources: International Mathematical Olympiad Shortlist 2000–2015, United States of America Mathematical Olympiad 2000–2015, Canadian Mathematical Olympiad 2000–2015, United States of America Mathematical Talent Search 2000–2015, Asian-Pacific Mathematics Olympiad 2000–2015, United States of America Team Selection Test 2000–2015, Benelux Mathematical Olympiad 2011–2015, Philippines IMO Selection Quizzes 2016–2017, *Problems from the Book Chapter 17* (Andreescu and Dopinescu, 2010), *Problems on Convergence of Sequences* (Chan Shio, 2016), and *MOP Experiment* (Wang, 2014). He then selected problems which involve examining a situation for an infinite number of iterations (approximately 150 problems). Afterwards, he read the solutions to these 150 problems and then selected 66 problems that contain analytical ideas, taking note of the role of analytical reasoning in their solutions. He then narrowed the list to thirty-six problems that best exemplify the nature of and heuristics involved in analytical reasoning.

For each of the thirty-six problems, the author examined the solution by recording its key ideas, summarizing them in a sentence with elaboration if necessary. Each key idea was classified as either *small* (mainly a technical detail) or *major* (a pivotal aspect in solving the problem) and assigned a point value from 1 to 10, with the constraint that the sum of points assigned from a problem is exactly 10. He then collated the list of key ideas and matched the ideas under twenty-three heuristics. As some heuristics were overly narrow or better classified as tools, the author combined some heuristics with one another, resulting into eleven broader heuristics. He then grouped the fifteen heuristics under three categories, namely *themes*: (1) taking a global view, (2) refocusing, and (3) investigating dynamics or processes.

The succeeding sections will discuss the three themes and the heuristics within them. Each heuristic will be described by one or two problems, followed by a thematic discussion and a list of problems for further reference (taken from the thirty-six problems that were analyzed). Problems are listed according to index number; heuristic scores and additional comments are listed in each entry. The problem discussion will maintain an informal yet instructional tone, including all thought motivation and guiding the reader through the motivation to the solution of the problem. Some details (not necessarily minor) may be left to the reader as exercises in order to focus the discussion to the heuristics and main ideas at hand. The text is interspersed with parenthetical comments to verify the reader's understanding of analytic techniques.

1.4 Use of bounding

Bounding is commonly used to establish the **relative position** of a term of a sequence or a defined quantity. Bounding may be either one-sided (only one of either $\geq$ or $\leq$) or two-sided (both $\geq$ and $\leq$).

Some problems that involve bounding a quantity are listed below:

A.2. Benelux 2011 P3 (S2) – *squeezing between for bounding and contradiction*

A.31. USAMTS 20/3/5 (S10)

A.33. USAMTS 26/3/3 (S1) – *squeezing powers to solve both sides simultaneously*

Once a discrete (i.e. taking only whole numbers) function is bounded both above and below (one may be a natural bound such as "positive integers"), its possibilities are limited.

The problems below involve bounded discrete functions as a central idea. As they are rather advanced, it is recommended to read the rest of the article first before attempting these problems.

A.7. ISL 2009 A6 (S2)

A.15. ISL 2015 N4 (M5) - *find an aspect in which it is bounded; reinterpreting using modulo may be helpful*

A.27. PFTB Example 12 (M7)

2 Taking a global view

Taking a global view is the essence of analytical reasoning: focusing on the "big picture" rather than point-based properties. A main idea in this heuristic is the concept of local *discrepancies*, which is inconsistent, deviant, or erroneous behavior relative to the norm. Sometimes, local discrepancies would compound as the sequence indices grow and become large, thus contradicting hypotheses of normalcy or boundedness. At other times, these discrepancies may be shown to have negligible effect over the long run and hence (after sufficient justification) may be ignored. A related concept is that of density of a sequence – in simple terms, this refers to the approximate or asymptotic proportion of the numbers from 1 to n, with n sufficiently large, represented by at least one term in the sequence.

The concept of taking a global view is used to finish the details to solutions in problems of succeeding sections, after applying another theme or heuristic.

2.1 Exploiting local discrepancies together with bounds

Although some discrepancies may seem minuscule, they may add up as the sequence iterates, ultimately causing a contradiction. This heuristic is frequently paired with bounded gaps, meaning that even later throughout the sequence, the discrepancy is not allowed to grow unbounded.

Example 2.1 (PEM Handout P2)**.** Let $(a_n)_n$ be a sequence of real numbers that satisfies the recurrence relation $a_{n+1} = \sqrt{a_n^2 + a_n - 1}$ for $n \geq 1$. Prove that a_1 is not within $(0,1)$.

Solution. Since we are showing that a_1 may not fall within $(0,1)$, a natural approach would be to assume otherwise and examine what occurs. Looking backwards, a clear contradiction would occur if one of the a_i's is such that $a_n^2 + a_n - 1 < 0$, as a_{n+1} would not be a real number. This is equivalent to $a_n^2 + a_n < 1$, which is true for all $a_n \leq \frac{1}{2}$ (this is merely a loose bound used to simplify computations; the actual bound is approximately $a_n \leq 0.618$.)

We analyze the relation of a_{n+1} to a_n by trying some examples:

- If $a_n = 0.8$, then $a_n^2 + a_n - 1 = 0.48$, $a_{n+1} = \sqrt{a_n^2 + a_n - 1} \approx 0.68$.
- If $a_n = 0.9$, then $a_{n+1} = \sqrt{a_n^2 + a_n - 1} = \sqrt{0.71} \approx 0.84$.
- If $a_n = 0.95$, then $a_{n+1} = \sqrt{0.95^2 + 0.95 - 1} = \sqrt{0.8525} \approx 0.923$.

It seems that $a_{n+1} < a_n$ whenever a_{n+1} is defined and $a_n < 1$. Indeed, as we are only interested in the case where $a_n^2 + a_n - 1 > 0$, we may square both sides of the equation $\sqrt{a_n^2 + a_n - 1} < a_n$ to get

$$\begin{aligned} a_n^2 + a_n - 1 &< a_n^2 \\ a_n - 1 &< 0 \\ a_n &< 1, \end{aligned}$$

which is true by assumption. Hence among defined a_i, the sequence (a_n) is strictly decreasing if $a_1 < 1$, as $a_2 < a_1 < 1$, $a_3 < a_2 < a_1 < 1$, and so on.

Despite the contradiction being apparent, the proof is not yet complete, as we merely have shown that the sequence is monotonically decreasing – nothing was shown about the rate of decrease. It is possible for a sequence to be monotonically decreasing yet never goes below a certain number; for example, the sequence $0.81, 0.801, 0.8001, \ldots$ is monotonically decreasing yet is always above 0.8. Thus, we must find a stronger claim.

We examine the values of a_{n+1} relative to a_n a bit more: Consider $f(x) = \sqrt{x^2 + x - 1}$. As seen earlier, $f(0.8) \approx 0.69$, $f(0.9) \approx 0.84$, and $f(0.95) \approx 0.923$. It seems that applying f to an x in $(0.7,1)$ decreases it in a somewhat predictable manner: the gap from one is multiplied by slightly more than 1.5. If we are able to prove that $1 - f(x) > 1.5(1 - x)$, this

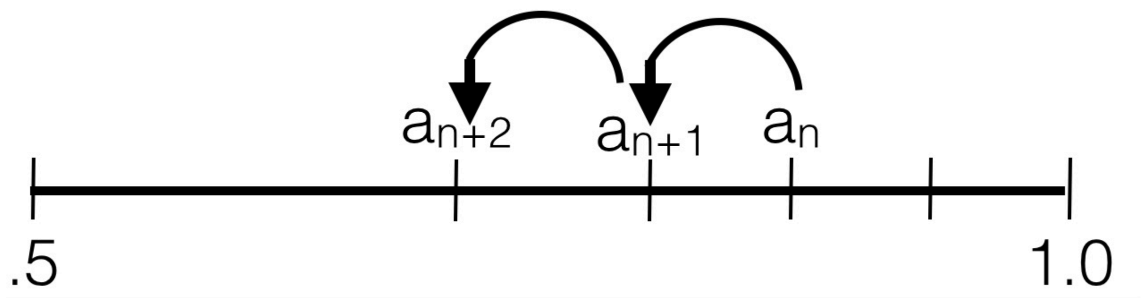

Figure 1: Visualization of a_n, a_{n+1}, and a_{n+2} on the number line.

becomes $1-a_{n+1} > 1.5(1-a_n)$. However, if there exists $a_i < 0.5 \to 1-a_n > 0.5$ then we arrive at a contradiction.

If we let the initial distance from 1 (i.e. $1-a_1$) be d, then

$$1-a_2 > 1.5d, 1-a_3 > 1.5^2 d,$$

and so on. In general, as long as a_i is defined, $1-a_i > 1.5^{i-1}d$. As 1.5^{i-1} is unboundedly large (it is well-known that a^x where $a > 1$ goes to infinity as x grows large), we are able to find i such that $1-a_i > 0.5$; hence, this establishes the contradiction.

It remains to prove that $1-f(x) > 1.5(1-x)$. We arrange, substitute, and manipulate:

$$\begin{aligned} f(x) &< 1.5x - 0.5 \\ \sqrt{x^2+x-1} &< 1.5x-0.5 \\ x^2+x-1 &< 2.25x^2-1.5x+0.25 \\ 1.25x^2-2.5x+1.25 &> 0 \\ 1.25(x-1)^2 &> 0, \end{aligned}$$

which is true. Thus, $1-f(x) > 1.5(1-x)$ and our proof is complete. □

Discussion. This is an example of how a small discrepancy (i.e. the value of $1-f(x)$; with the alternative hypothesis being that it stays near 1 so that there would be no contradiction) grows larger over the terms. The motivation for considering the expanding discrepancy is observation and working backwards to our goal, which is to get a contradiction. Note that the discrepancy need not be exponential (as is the case in this problem), as long as it is unbounded.

Below is another example on exploiting a local discrepancy. This problem is significantly more difficult than Example 2.1, and the motivation for considering "ballooning" local differences is harder to spot. A hint to exploit local discrepancies is the ability to quantify a term's influence on later terms, typically through a straightfoward recurrence-style equation.

Example 2.2 (USAMTS 26/3/3)**.** Let $a_1, a_2, a_3, ...$ be a sequence of positive real numbers such that:

(i) For all positive integers m, n, we have $a_{mn} = a_m a_n$.

(ii) There exists a positive real number B such that for all positive integers m, n with $m < n$, we have $a_m < Ba_n$.

Find all possible values of $\log_{2015}(a_{2015}) - \log_{2014}(a_{2014})$.

Solution. Our first step is to examine condition (i), as this seems to be be the most useful to analyze the long-term behavior of the sequence.

Set $m = n$ in (i) to get $a_{n^2} = (a_n)(a_n)$. Use this equation to set $m = n^2$ in (i) to get

$$a_{n^3} = (a_{n^2})(a_n) = (a_n)(a_n)(a_n) = a_n^3.$$

By induction, $a_{n^i} = a_n^i$ for all $i \geq 1$ (verify this). We seek an invariant for this. Notice that both the index and the exponent grow exponentially (with the same power). Together with the fact that

$$\log_{n^i} a_n^i = \log_n a_n, \tag{1}$$

we get that $\log_k a_k$ is constant for $k \in \{n, n^2, n^3, \cdots\}$. Observe the similarity to what the problem is asking; therefore, it is natural to hypothesize that $\log_k a_k$ is constant for all k, or more narrowly, it is the same for $k = 2015$ and $k = 2014$.

Write $\log_{2015}(a_{2015}) = a$, $\log_{2014}(a_{2014}) = b$, and $f(x) = \log_x a_x$. Our discrepancy right now is a versus b, so we assume that they are different. We have, however, that $\log_{2015^i}(a_{2015^i}) = a$ and $\log_{2014^j}(a_{2014^j}) = b$ by Equation 1. We may set these close together (up to a factor of 2015 by setting j such that $2015^i < 2014^j < 2015^{i+1}$). The reader can verify that for every i, there exists a j. Hence, even at large indices, $f(x)$ still maintains differences that originated from comparably small indices (a_{2014} and a_{2015}). Taking the logarithm of condition (ii) gives

$$\log(a_m) - \log(a_n) < \log(B) \tag{2}$$

for $m < n$.

We apply this first to the left hand side of the equation, giving that $\log(a_{2015^i}) - \log(a_{2014^j})$ is bounded, using this to show that $a \leq b$. Suppose that $a > b$. By Equation 1,

$$\log(a_{2015^i}) - \log(a_{2014^j}) = i\log(a_{2015}) - j\log(a_{2014}) = ia\log 2015 - jb\log 2014,$$

and this quantity is bounded above. As $2015^i < 2014^j < 2015^{i+1}$,

$$i\log 2015 - j\log 2014 > i\log 2015 - (i+1)\log 2015 = -\log 2015 > -4.$$

Hence we know that $i \log 2015$ is smaller by only a constant when compared to $j \log 2014$. However, if we grow i linearly, then j would grow linearly as well (at the order of $(a-b)x$), then $ia \log 2015 - jb \log 2014$ would grow linearly and would be unbounded (positive, as we assumed that $a > b$), contradicting the assumption that this is bounded above. (The reader is encouraged to formalize this argument, as the proof in this paragraph is rather informal.)

Just as the left hand side of $2015^i < 2014^j < 2015^{i+1}$ was used to prove that $a \leq b$, the right hand side of the same compound inequality may be used to prove that $a \geq b$. As the proof is very similar, it is left as an exercise to the reader.

Exercise 2.3. Adapt the proof in the previous paragraph to prove that $a \geq b$.

Combining $a \leq b$ and $a \geq b$ gives $a = b$, which is the desired conclusion; thus, our answer is 0. □

Discussion. The key ingredients to this heuristic are the ability to maintain properties seen in terms with small indices to terms with large indices (here, we repeatedly used the $a_{mn} = a_m a_n$) and the presence of a rigidly-bounded quantity that holds for all terms. Rigid depends on the problem, but is typically something constant. When a growing discrepancy (at least linear) is together with a constant-bounded quantity, there is a likely contradiction, just as seen in the bounding part of the solution.

The ideas in the related problems below are even more subtle than the second example, but nevertheless play a pivotal role in their solutions. If stuck, "forcing something until it breaks" is the proper mindset for this heuristic: examine what occurs if a local discrepancy is forced to remain as indices grow large, and lead this to a contradiction.

Related problems

A.4. ISL 2001 A2 (M7) – *measuring differences (using a specialized function or interpretation) and summing them up*

A.5. ISL 2003 A3 (S3) – *let small quantities repeated many times add up (as long as they are not decreasing by much)*

A.12. ISL 2014 A1 (M7) – *force it until it breaks, find something that would make it break*

A.30. USAMO 2015 P6 (M3) – *discrete conditions forcing 'steps' that would cause contradictions*

A.31. USAMTS 20/3/5 (S3) – *increase 'tightness' of a fixed-length bound by iterating*

2.2 Minimizing local discrepancies' impact by considering large

Just as local discrepancies may have multiplied impact as indices grow, they may also have reduced / divided impact, as seen in the examples below. Typically, this occurs when we are able to show that the discrepancy either stays constant or decreases as the index grows large, and we are concerned with the relative size of the discrepancy with an asymptotically larger quantity.

Example 2.4 (USAMTS 28/3/4)**.** Let $\{A_1,\ldots,A_n\}$ and $\{B_1,\ldots,B_n\}$ be sets of points in the plane. Suppose that for all points x,

$$D(x,A_1)+D(x,A_2)+\cdots+D(x,A_n) \geq D(x,B_1)+D(x,B_2)+\cdots+D(x,B_n)$$

where $D(x,y)$ denotes the distance between x and y. Show that the A_i's and the B_i's share the same center of mass.

Solution. Suppose, for the purpose of contradiction, that the centers of mass are different. Let the centers of mass of the sets $\{A_1,\ldots,A_n\}$ and $\{B_1,\ldots,B_n\}$ be X and Y, respectively. There is no apparent contradiction when we consider a point Z near X and Y, as two-dimensional distances are difficult to compute. As we want the distances to be more quantifiable, we let Z be far away from X and Y along the line YX . Hence, we claim that there exists a point Z on the extension of ray YX such that

$$D(x,A_1)+D(x,A_2)+\cdots+D(x,A_n) \geq D(x,B_1)+D(x,B_2)+\cdots+D(x,B_n)$$

is false.

As Z goes away from X along ray YX, $D(Z,A_i)$ appears to be increasingly solely-affected by the distance from Z to the foot of the perpendicular from A_i to XY. This leads us to consider splitting $D(Z,A_i)$ into two components: the horizontal and residual components. Define the horizontal component as the distance from Z to the foot of the perpendicular from A_i to YX, and label this as $H(Z,A_i)$. On the other hand, define the residual component as $D(Z,A_i)-H(Z,A_i)$, and label this as $R(Z,A_i)$. Define H and R similarly for B_i, $1 \leq i \leq n$.

Based on the above reductions, we quantify the sum of distances given in the problem:

$$\begin{aligned} D(Z,A_1)+D(Z,A_2)+\cdots+D(Z,A_n) =& H(Z,A_1)+H(Z,A_2)+\cdots+H(Z,A_n) \\ &+R(Z,A_1)+R(Z,A_2)+\cdots+R(Z,A_n). \end{aligned}$$

Similarly,

$$\begin{aligned} D(Z,B_1)+D(Z,B_1)+\cdots+D(Z,B_n) =& H(Z,B_1)+H(Z,B_2)+\cdots+H(Z,B_n) \\ &+R(Z,B_1)+R(Z,B_2)+\cdots+R(Z,B_n). \end{aligned}$$

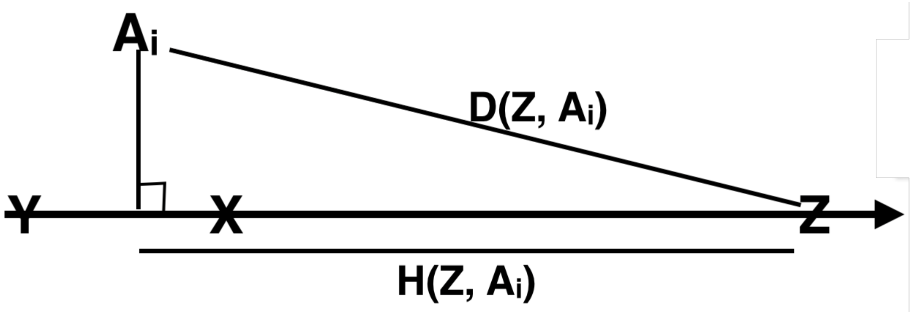

Figure 2: Visualization of $D(Z,A_i)$ and $H_(Z,A_i)$.

Note that $H(Z,A_1)+\cdots+H(Z,A_n)$ is simply nZX because X is the center of mass of $\{A_1,\ldots,A_n\}$. Similarly, $H(Z,B_1)+\cdots+H(Z,B_n)$ is nZY. Note that $ZY-ZX$ is equal to the length of segment YX, so $nZY-nZX$ is equal to n times the length of segment YX.

Thus,

$$(H(Z,B_1)+H(Z,B_2)+\cdots+H(Z,B_n))-(H(Z,A_1)+H(Z,A_2)+\cdots+H(Z,A_n)$$

is constant regardless wherever we place Z on ray YX. From the inequality given in the problem statement, we have to show that

$$(R(Z,A_1)+R(Z,A_2)+\cdots+R(Z,A_n))-(R(Z,B_1)+R(Z,B_2)+\cdots+R(Z,B_n)) \tag{3}$$

is always at least this constant. We have, however, observed that as Z goes far from X and Y, the non-horizontal component has less and less effect. A possible approach is to show that the value of

$$(R(Z,A_1)+R(Z,A_2)+\cdots+R(Z,A_n))-(R(Z,B_1)+R(Z,B_2)+\cdots+R(Z,B_n))$$

tends to 0 as Z goes further from X and Y along the ray YX. If this occurs, then Equation 3 will be less than this constant, implying that the centers of mass must be identical.

Note that it suffices to show that each of $R(Z,A_i)$ and $R(Z,B_i)$ would tend to 0 as the distance from Z to Y grows large – because, if this is so, then the sum of any number of the R's (or their negative) would also tend to 0. We quantify $R(Z,A_i)$ (similar for $R(Z,B_i)$). Let the foot of the perpendicular from A_i to YX be P_i. Then by the Pythagorean Theorem

$$R(Z,A_i)=ZA_i-ZP_i=\sqrt{ZP_i^2+P_iA_i^2}=\frac{P_iA_i^2}{\sqrt{ZP_i^2+P_iA_i}+ZP_i}.$$

Note that the numerator is constant, while the denominator tends to infinity as Z goes arbitarily far from Y (and also arbitrary far from P_i). Hence, $R(Z,A_i)$ tends towards 0. $\square$

Discussion. The main idea here is observing that as we take a point far from all the given points, then the sum of the distances are more predictable due to local discrepancies (i.e. the residual parts) having lessened effect. Note that throughout the solution, we needed to quantify our vague, intuitive observations, hence the choice of Z along ray YX.

The next problem is a number theoretic problem on sets and divisibility. Although there are multiple constructions, a construction which exemplifies the ideas of this section will be presented. By ignoring (with justification) the effect of "small" components in favor to the "large" components, it becomes possible to deal with them separately and ignore interdependencies, greatly simplifying the problem.

Example 2.5 (MOSC 4/27/17 AM Quiz)**.** Give a set of 2016 numbers such that there does not exist two distinct subsets, the sum of elements one of which divides the sum of elements of the other.

Solution. We start by finding simplifications. A potential complication is that there is no restriction on the number of elements of the "divisor" and the "dividend" subsets. It would be more convenient if we restrict the possible number of elements somehow, and this is made possible by limiting the range of the sum of elements when we consider a subset of a particular size.

To do this, it is helpful to add an extremely large "base value" for each element in our set. Let this be K (to be defined later) that is large enough such that everything aside from the "base value" virtually do not matter at all, as long as subset-sum whole number quotients are concerned. This would restrict the subset pair in consideration to a pair of subsets, one of whose cardinality divides that of the other.

Exercise 2.6. Why is this? Provide an intuitive explanation.

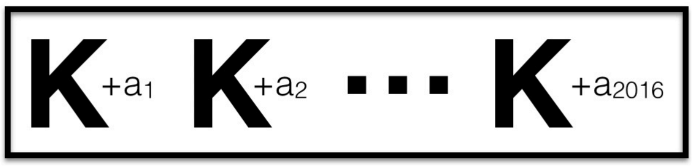

Figure 3: Visualization of K's role in the construction.

As K is dependent on the small, non-base values, we assign these first. It is well-known that to prevent subset-sum collisions, one could take powers of 2. Although we are proving something stronger - not only collisions, but also one being divisible by another - this is an ideal starting point. Hence, we make the small values 2^0, 2^1, $\cdots$, 2^{2015}.

Now, we find a suitable value of K, to be able to state our subset and justify why it satisfies the problem statement's conditions. The sum of a subset of $\{2^0, 2^1, \ldots, 2^{2015}\}$ is a positive integer that is less than or equal to 2^{2016}, as $2^0 + 2^1 + \cdots + 2^{2015} = 2^{2016} - 1$. Hence, if our subset contains m elements, the sum of elements in our subset is between mK and $mK + 2^{2016}$. As the number of elements in a subset ranges from 1 to 2016, the subset-sum-quotient may be as high as 2016, so our potential "swing" may be up to $2016 \cdot 2^{2016}$. It makes sense to let $K = 2016 \cdot 2^{2016}$. For later convenience, however, we arbitrarily increase

K to avoid tight bounds and hard-to-prove details. Write $K = 2^{4032}$, which is clearly greater than $2016 \cdot 2^{2016}$.

With $K = 2^{4032}$, our set S is $\{2^{4032} + 2^0, 2^{4032} + 2^1, \cdots 2^{4032} + 2^{2015}\}$.

The proof that this set has no two subsets A and B such that the sum of the elements in A divides the sum of elements in B is left as an exercise to the reader. It contains two major parts: one to bound large values (i.e. $K = 2^{4032}$) and the other to bound small values (i.e. the 2^is, $0 \leq i \leq 2015$). First, we show that the cardinality of A divides the cardinality of B by contradiction, writing $|B| = p|A| + q$, $0 < q < |A|$. Next, given this, we show that it is impossible to find two disjoint subsets of $\{2^0, 2^1, \cdots 2^{2015}\}$ such that the cardinality and sum of one both divide the cardinality and sum of the other, respectively. □

Discussion. This problem is related to the heuristic of "minimizing local discrepancies' impact by taking large numbers" because we are separating the small (local) and the large (global) in order to reduce interdependencies, which would only complicate the proof. A significant portion of the solution was spent identifying K. Indeed, finding a suitable K involves having a thorough understanding of the sequence its minor discrepancies, so that K would surpass all of the minuscule "bumps".

The two examples on this heuristic show just a small portion of the wide variety of problems that incorporate this heuristic. Although one problem involves finding a contradiction and the other is aimed at proving that a configuration works (i.e. no contradiction), their main idea is the same: distinguishing the large and the small.

Related problems

A.1. APMO 2013 P3 (M6) – *consider the difference between floor and actual values*

A.6. ISL 2004 A2 (M4) – *the anti-asymptotic/mainstream behavior is local and has little impact*

A.11. ISL 2013 A3 (M3)

A.14. ISL 2015 C5 (M3)

A.26. PFTB Example 8 (M5) – *tightening ranges as we make things large; oscillations are asymptotically negligible*

2.3 Considering density of number occurrence

Another heuristic related to taking a global view is considering density of number occurrence. Density may be restated as: "approximately what proportion of integers within a certain range is represented in the sequence?" This may be a linear function (e.g. $\frac{n}{2}$, $\frac{2n}{3}$), a radical quantity (e.g. $\sqrt{n}$, $^3\sqrt{n}$), a logarithmic quantity (e.g. $\ln n$, $\log_2 n$), or a combination of them (e.g. $\frac{n}{\log n}$, the density of primes). A quantity may be constructed to measure the density - a common one is the reciprocal or a modification of it.

The example problem below devises a seemingly-unintuitive quantity that efficiently establishes a contradiction, and the exposition attempts to explain the intuition behind this.

Example 2.7 (PFTB Example 6)**.** Suppose that f is a polynomial with integer coefficients and that (a_n) is a strictly increasing sequence of positive integers such that $a_n \leq f(n)$ for all n. Then the set of prime numbers dividing at least one term of the sequence (a_n) is infinite.

Solution. We proceed by contradiction; suppose that the set of prime numbers S dividing at least one term of the sequence (a_n) is finite. Then all terms of the sequence are distinct and are composed only of prime factors from S. One may recall that that the sum of reciprocals of all powers of a prime p is equal to $1 + \frac{1}{p} + \frac{1}{p^2} + \cdots = \frac{p}{p-1}$. This is generalizable to the scenario when more than one prime is considered through the fact that

$$\sum_{\alpha_1, \alpha_2, \cdots, \alpha_N \geq 0} \frac{1}{p_1^{\alpha_1} \cdots p_N^{\alpha_N}} = \left(\sum_{\alpha_1 \geq 0} \frac{1}{p_1^{\alpha_1}} \right) \cdots \left(\sum_{\alpha_N \geq 0} \frac{1}{p_N^{\alpha_N}} \right) = \prod_{j=1}^{N} \sum_{i \geq 0} \frac{1}{p_j^i} = \prod_{j=1}^{N} \frac{p_j}{p_j - 1},$$

which is a finite quantity. (The reader is encouraged to verify this equation.) Hence, the sum of reciprocals converges.

The problem would be solved if the sum of the reciprocals of $f(i)$, $i \geq 1$ converges. This, however, is not the case. Take, for example, $f(x) = x^2$: it is well known that

$$\sum_{k=0}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

We attempt to modify the sum of the reciprocals of $f(i)$ in such a way that it will diverge, but in such a way that is still compatible with the earlier prime-reciprocals computation.

A possible approach would be to "reduce" the degree of $f(x)$ by instead taking $f(x)^k$, where $k < 1$. This way, it is still possible to proceed with the prime computations, as we are instead taking the sum of the kth powers of the reciprocals. Indeed, we modify the equations to

$$\sum_{\alpha_1, \alpha_2, \cdots, \alpha_N \geq 0} \frac{1}{p_1^{k\alpha_1} \cdots p_N^{k\alpha_N}} = \left(\sum_{\alpha_1 \geq 0} \frac{1}{p_1^{k\alpha_1}} \right) \cdots \left(\sum_{\alpha_N \geq 0} \frac{1}{p_N^{k\alpha_N}} \right) = \prod_{j=1}^{N} \sum_{i \geq 0} \frac{1}{p_j^{ki}} = \prod_{j=1}^{N} \frac{p_j^k}{p_j^k - 1},$$

which is still a finite quantity.

As the sum of the reciprocals of the natural numbers diverges, it may be a good idea to take $k = 1/2\deg(f)$. Then, $1/(f(n))^k$ diverges. (Intuitive; try devising a formal proof.) Thus, if there exists a finite number of primes, consider the sum R of the kth powers of the reciprocals of the terms in (a_n). As $a_n \leq f(n)$ for all n,

$$R > \sum_{n \geq 1} \frac{1}{f(n)^k} = \infty.$$

Based on the sum of kth powers of the reciprocals of primes, however,

$$R \leq \prod_{j=1}^{N} \frac{p_j^k}{p_j^k - 1},$$

which is finite, contradiction. □

Discussion. The premise of a finite set of prime numbers being the only prime factors of a set of numbers implies that almost all natural numbers are ineligible, and the density is roughly logarithmic. This gives a large amount of "leeway" to apply polynomial-magnitude modifications, such as taking the kth power.

Density problems typically allow extremely loose bounds, as long as the magnitude of the largest term is maintained, so when solving such problems, it is recommended to focus on the overall magnitude rather than specifics, and repeatedly argue "let N be large enough" in order for a high-magnitude term to overtake any lower-magnitude term. The three problems below are rather varied, but all of them involve calculating the density of a function or a set to establish some contradiction.

Related problems

A.10. ISL 2012 A6 (M3) – *relating function growth to density of occurrence*

A.25. PFTB Example 7 (M10) – *reciprocals to quantify sparsity, use counting for loose estimates*

A.29. USAMO 2014 P6 (M7) – *estimating the quantity based on primes, show that primes large enough must occur*

3 Refocusing

The four heuristics in this section all involve considering the sequence from another point of view: either for the terms considered or for the sequence as a whole. Sometimes, examining the sequence at face value would not suffice: manipulations must be performed to make the analytic arguments more apparent. Refocusing may be done by focusing on a subsequence, reinterpreting the sequence, considering an alternative quantity, or considering a sequence of subsequences.

3.1 Focusing on a subsequence

Some problems provide information about the sequence at all indices n, and some others ask us to prove that a specified property holds for infinitely many positive integers n. While it may seem more "efficient" and less error-prone at first by considering all $n \in \mathbb{Z}^+$, this may complicate the proof. Only utilizing specific information or restricting attention based a subsequence may significantly simplify arguments.

Example 3.1 (PFTB Example 4)**.** Let a_1, a_2, $\cdots$ a_k be positive real numbers such that at least one of them is not an integer. Prove that there exists infinitely many positive integers n such that n and $\lfloor a_1 n \rfloor + \lfloor a_2 n \rfloor + \cdots + \lfloor a_k n \rfloor$ are relatively prime.

Solution. Just as in previous problems, we take an indirect approach: since the problem statement asks to prove that there exists infinitely many such positive integers n, we suppose otherwise - that there exists only finitely many such n. This implies that there exists $M \in \mathbb{Z}^+$ such that n and $\lfloor a_1 n \rfloor + \cdots + \lfloor a_k n \rfloor$ share a common factor for all $n \geq M$.

We examine what it means for a quantity to share a common factor with n, for various $n \in \mathbb{Z}^+$. If $n = 2016$, then it suffices for the quantity to be even, a multiple of 3, or a multiple of 7. If $n = 2015$, then it suffices for the quantity to be a multiple of 5, 13, or 31. On the other hand, if $n = 2017$, which is a prime number, then the quantity must be divisible by 2017. As considering a single property is far simpler than considering three properties (where the requirement is that at least one has to be true), we restrict the numbers that would give only one possible common factor: the prime numbers.

Thus, we would attempt to contradict the slightly stronger statement: "there exists $M \in \mathbb{Z}^+$ such that n and $\lfloor a_1 p \rfloor + \cdots + \lfloor a_k p \rfloor$ share a common factor for all primes $p \geq M$." Note that p must divide $\lfloor a_1 p \rfloor + \cdots + \lfloor a_k p \rfloor$. Let the primes considered be p_1, p_2,$\cdots$. Write $\lfloor a_1 p_n \rfloor + \cdots + \lfloor a_k p_n \rfloor = x_n p_n$, then

$$x_n = \frac{\lfloor a_1 p_n \rfloor + \cdots + \lfloor a_k p_n \rfloor}{p_n}. \tag{4}$$

Upon careful examination of Equation 4, we observe that without the floors, this is equal to $a_1 + a_2 + \cdots + a_k$. We use a heuristic from the previous section, namely minimizing local discrepancies' impact by considering large: the distance of the quantity to the $a_1 + a_2 + \cdots + a_k$ is less than $\frac{k}{p_n}$, as each floor function makes the numerator at most 1 away from the actual value. As n grows large, this difference becomes infinitesimally small, and thus x_n must converge to $a_1 + a_2 + \cdots + a_k$. As this is a sequence of an integers, for large enough n, $x_n = a_1 + a_2 + \cdots + a_k$.

This means that for large enough n, the floor functions in Equation 4 will have no effect, as its value is the same as if there were no floor functions. A floor function would only have no effect if the term were an integer, so we need $a_i p_n$ to all be integers for all large enough n. But this is impossible: if a_i is irrational then $a_i p_n$ will be irrational and cannot be an integer; if a_i is rational then there exists a prime p_m that does not divide its denominator (when a_i is reduced to lowest terms). □

Discussion. By only considering a specific subsequence, the problem statement's condition was slightly weakened but is still true. After deciding to only consider prime numbers for

n, the problem is greatly simplified, then the remainder is routine analysis based on the concepts of the previous section.

Considering a specific subsequence would not always yield a solution because some problems require use of all information, and it is possible for the natural subsequences to be "too weak" to give any meaningful conclusions.

The subsequence chosen need not to be predefined or even a well-known set: it could be any subsequence that possess a certain property (see ISL 2012 A4 from the related problems, where the solution focuses on a convergent subsequence). The property may be dependent on the sequence's dynamics, such as in ISL 2004 A2.

Related problems

A.6. ISL 2004 A2 (M3) – *focus on a segment where there would be a contradiction*

A.9. ISL 2012 A4 (M5) – *focusing on convergent subsequence (twice)*

A.10. ISL 2012 A6 (M4) – *focus on a segment that is useful (somewhat clear from problem statement)*

A.33. USAMTS 26/3/3 (M3) – *focus on a segment where there would be a contradiction*

3.2 Reinterpreting the sequence

The solution to a problem may not be apparent when the sequence is taken at face value, as problem statements may tend to obscure hints to their solutions. A common technique to approach analytical-flavored problems, where the idea is conceptually simple yet hidden deep into the problem, is examine "what is actually going on" in the sequence, and then devise an alternative interpretation.

Example 3.2 (ISL 2003 A3, part (a))**.** Consider two monotonically decreasing sequences (a_k) and (b_k), where $k \geq 1$, and a_k and b_k are positive real numbers for every k. Now, define the sequences

$$\begin{aligned} c_k &= \min(a_k, b_k) \\ A_k &= a_1 + a_2 + \ldots + a_k \\ B_k &= b_1 + b_2 + \ldots + b_k \\ C_k &= c_1 + c_2 + \ldots + c_k \end{aligned}$$

for all natural numbers k. Construct two monotonically decreasing sequences (a_k) and (b_k) of positive real numbers such that the sequences (A_k) and (B_k) are not bounded, while the sequence (C_k) is bounded?

Solution. The problem statements defines c_k in terms of a_k and b_k, so it may be natural to assume that we must first construct unbounded sequences a_k and b_k, then prove that the resulting c_k is bounded. Although this is possible, it is certainly difficult as we have to make sure that when identifying the values of each of the a_is and b_is, A_k and B_k do not converge, meaning that the terms must not be too small. On the other hand, the minimum of them may not be too large; otherwise, C_k may become unbounded.

This motivates the alternative view, where we first construct (c_i) such that its partial sum sequence converges, and then construct (a_i) and (b_i) accordingly. Suppose that we have a sequence c_i such that its sum converges to some finite value and is monotonically decreasing. We attempt to construct (a_i) and (b_i), one whole number at a time. Since c_k is the minimum of a_k and b_k, it suffices to keep one small, following c_k, while the other maintains some value. Recall the idea from the previous section on adding constants infinitely many times. This maintained value, when sustained, would add whole numbers to one of the sequences. If we show that this may be done infinitely, alternating between (a_i) and (b_i), then we have constructed the desired sequences.

We desire to formalize this idea and establish the length intervals where a specific value is placed for the purpose of adding to the sums A_k and B_k. As in the official solution, we define a strictly increasing sequence of integers (k_m) starting with 1 and satisfying the condition $(k_{m+1} - k_m)c_{k_m} \geq 1$. The continuation of the solution is left as an exercise:

Exercise 3.3. Define the sequences (a_i) and (b_i) based on (k_m) and the ideas presented in the second to the last paragraph. Complete the proof of A_k and B_k being unbounded. □

Discussion. This example problem shows how alternative interpretation may be useful, even though the problem statements strongly hints towards a standard interpretation. While the problem may seem challenging at first, after the re-interpretation and the observation that we may maintain a value on either (a_i) or (b_i) for an arbitrary length, the problem is reduced to routine details.

A weakness of this heuristic, however, is that unless the solution is quickly seen or the solver has accumulated enough experience to be confident that the reinterpretation is more conducive to solving the problem than the original formulation, reinterpretations may complicate the problem.

The first two of the three problems below have a somewhat-combinatorial reinterpretation despite their statements being sequence-based.

Related problems

A.8. ISL 2010 A7 (M3) – *repeatedly use the recursive definition – what happens when the process stops?*

A.29. USAMO 2014 P6 (M3) – *interpret as a grid of primes*

A.32. USAMTS 24/3/4 (M5)

3.3 Considering an alternative quantity

Some problems may be simplified by defining a quantity based on terms of the sequence rather than dealing with sequence terms individually. The problem statement of the example problem below has a clear alternative quantity to be considered and yields a quick solution.

Example 3.4 (USA TST 2008 P5)**.** Two sequences of integers, $a_1, a_2, a_3, \dots$ and $b_1, b_2, b_3, \dots$, satisfy the equation

$$(a_n - a_{n-1})(a_n - a_{n-2}) + (b_n - b_{n-1})(b_n - b_{n-2}) = 0$$

for each integer n greater than 2. Prove that there exists a positive integer k such that $a_k = a_{k+2008}$.

Solution. The given strongly hints to consider the quantities $c_i = a_i - a_{i-1}$ and $d_i = b_i - b_{i-1}$. It is simplified into

$$\begin{aligned} c_i(c_i + c_{i-1}) + d_i(d_i + d_{i-1}) &= 0 \\ c_i^2 + c_i c_{i-1} + d_i^2 + d_i d_{i-1} &= 0 \end{aligned} \tag{5}$$

The terms c_i^2 and $c_i c_{i-1}$ is reminiscent of the expansion $(c_i + c_{i-1})^2 = c_i^2 + 2c_i c_{i-1} + c_{i-1}^2$. We incorporate this in Equation 5. Indeed,

$$2c_i^2 + 2c_i c_{i-1} = (c_i + c_{i-1})^2 + c_i^2 - c_{i-1}^2,$$

and similarly,

$$2d_i^2 + 2d_i d_{i-1} = (d_i + d_{i-1})^2 + d_i^2 - d_{i-1}^2.$$

Hence, taking Equation 5 doubled gives

$$\begin{aligned} (c_i + c_{i-1})^2 + c_i^2 - c_{i-1}^2 + (d_i + d_{i-1})^2 + d_i^2 - d_{i-1}^2 &= 0 \\ (c_i + c_{i-1})^2 + (d_i + d_{i-1})^2 + c_i^2 + d_i^2 &= c_{i-1}^2 + d_{i-1}^2 \\ c_i^2 + d_i^2 &\geq c_{i-1}^2 + d_{i-1}^2 \end{aligned}$$

Equality holds if and only if $c_i + c_{i-1} = 0$ and $d_i + d_{i-1} = 0$.

Note that the $c_n^2 + d_n^2$ only takes nonnegative integer values, but in each step either stays the same or decreases. The number of decreases is finite; therefore, at some point, it stops decreasing and stays constant throughout. Let this point be K. Then $c_n^2 + d_n^2$ is the same for all $n \geq K$. By the equality case, we get that

$$c_i + c_{i-1} = 0 \implies (a_i - a_{i-1}) + (a_{i-1} - a_{i-2}) = 0 \implies a_i = a_{i+2}$$

. From this, the conclusion is clear. We may repeatedly apply the equation $a_i = a_{i+2}$ to get that $a_i = a_{i+2} = a_{i+4} = \cdots = a_{i+2008}$, from which taking $i = k$ solves the problem. □

Discussion. Despite being the second problem in a three-item test, the problem is relatively straightforward after noticing the alternative quantity to be defined, which is clear from the problem statement. A possible difficulty is the pitfall of work backwards, as the conclusion is extremely loose and thus would avert any attempts towards a solution based on it. Solving it forward based on the given, however, would yield a simple and motivated solution.

Not all problems have their alternative quantities spelled out as clearly in the problem statement as in the example. In some problems, a clever observation or understanding of the sequence's dynamics may be required to discover the alternative quantity. Typically, before identifying the alternative quantity, one must experiment with a few values and possible sequences. Among the related problems, a problem that is largely dependent on this heuristic is ISL 2014 A1, whose simplest solution involves rearranging an inequality to uncover the desired quantity.

Related problems

A.12. ISL 2014 A1 (M5) – *expanding out 'averages' to reinterpret our desired condition*

A.13. ISL 2014 A2 (S3) – *consider successive differences and their growth*

A.23. PFTB Example 5 (S3) – *perfect square* $\implies$ *consider square root*

A.24. PFTB Example 6 (S4) – *adjust the 'controlled quantity' based on the other one in the desired contradiction*

A.30. USAMO 2015 P6 (M7) – *write a quantity for the imperfection (deviation from ideal scenario without restrictions) and focus on that*

3.4 Considering a sequence of subsequences

The behavior of a sequence may sometimes be more suitably analyzed in chunks rather than as separate terms. This is recommended when it is possible to determine subsequences with distinctive behavior. The example below demonstrates this heuristic. As this example is fairly involved, some details would be left as an exercise to the reader.

Example 3.5 (ISL 2015 N4)**.** Suppose that $a_0, a_1, \cdots$ and $b_0, b_1, \cdots$ are two sequences of positive integers such that $a_0, b_0 \geq 2$ and

$$a_{n+1} = \gcd(a_n, b_n) + 1, \qquad b_{n+1} = \operatorname{lcm}(a_n, b_n) - 1.$$

Show that the sequence a_n is eventually periodic; in other words, there exist integers $N \geq 0$ and $t > 0$ such that $a_{n+t} = a_n$ for all $n \geq N$.

Solution. The recurrence for (a_n) is interesting: a_{n+1} is one more than the greatest common divisor of a_n and another number. As when a_n is large, taking the greatest common divisor of it and another number may only either keep it the same or reduce it by a significant amount (by at least half of the number). This implies that a_n would have difficulty increasing, so we may hypothesize that it is bounded. A similar analysis, on the other hand, would not hold for (b_n). (Why?)

In the previous paragraph, we noticed two types of behavior of (a_n). The first is if it increases by 1, and this only happens if $a_n|b_n$. On the other hand, if a_n is not a factor of b_n then it would decrease, as long as it is greater than 2. (Why?) We try the initial pair $(a_0, b_0) = (3, 117)$. Following the sequence we get

$$(4,116),(5,115),(6,114),(3,341),(2,1022),(3,1021),(2,3062),\ldots$$

Notice that the sequence occurs in chunks of increasing a_i, drops down, increases again, drops down, etc. Furthermore, we are sure that in any increasing contiguous subsequence, it must eventually drop down when $a_n|b_n$, i.e. $a_n|(a_n+b_n)$, which is kept invariant in a purely increasing contiguous subsequence.

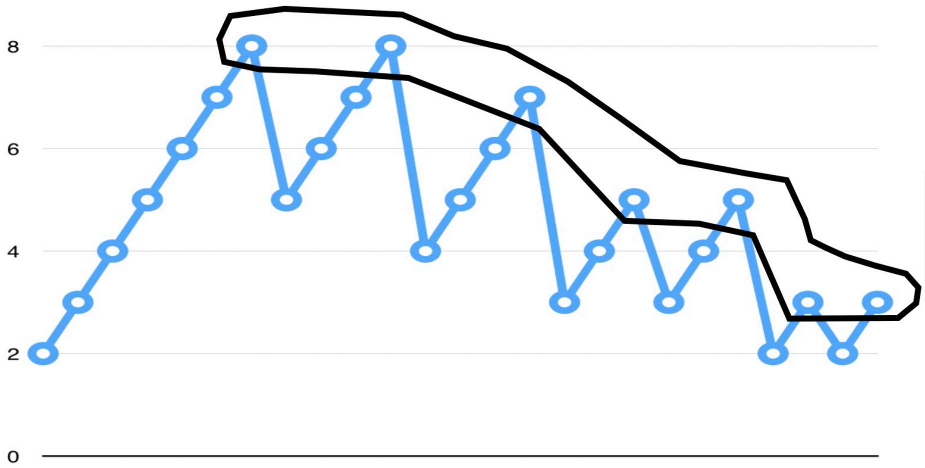

Figure 4: Visualization of the behavior of a_n.

Another observation we make that would imply (a_n) being bounded is that the "peaks" of the sequence are non-increasing. To quantify the notion of "peaks", define

$$W_n = \{m = \mathbb{Z}_{\geq 0} : m \geq a_n \text{ and } m \nmid (a_n + b_n)\}$$

and $w_n = \min W_n$. We split our proof into two cases: $a_n \mid b_n$ and $a_n \nmid b_n$.

If $a_n \mid b_n$ then $a_{n+1} = a_n + 1$. As $a_n \mid (a_n + b_n)$, then by definition of W_n, $a_n \notin W_n$. We also have $a_{n+1} + b_{n+1} = a_n + b_n$, so no element is added to the set W_n, and the condition $m \nmid s_n$ in the definition of W_n is maintained, so $W_{n+1} = W_n$ and $w_{n+1} = w_n$.

For the other case, $a_n \nmid b_n$, we have $a_n \nmid (a_n + b_n)$, so $a_n \in W_n$ and $w_n = a_n$. As we would want to show that $w_{n+1} \leq w_n$, it would suffice to show that $a_n \in W_{n+1}$.

As we analyze the differences between the cases $a_n \nmid b_n$ and $a_n \mid b_n$, we derive two key observations: $a_n \geq a_{n+1}$ and $a_n \nmid (a_{n+1} + b_{n+1})$. The proof of these two observations are intuitive and thus are left as an exercise to the reader. By the definition of W_{n+1}, these two observations imply $a_n \in W_{n+1}$.

Exercise 3.6. Prove that if $a_n \nmid b_n$, then $a_n \geq a_{n+1}$ and $a_n \nmid s_{n+1}$.

From here, we have $a_n \leq w_n \leq w_0$, so the sequence (a_n) is bounded. Hence, it takes only finitely many values. The problem is, however, that (b_n) is likely unbounded. As (a_n) is bounded, however, and we are mainly concerned with (a_n) and its common factors with (b_n), this motivates us to consider the sequence (b_n) taken modulo the least common multiple of all the a_is (the set of numbers is bounded, so the least common multiple is finite). Let this sequence be (r_n). Note that there are only finitely many possible pairs (a_n, r_n) as both a_n and r_n are bounded and are positive integers.

Next, we observe that the pair (a_n, r_n) uniquely determines the pair (a_{n+1}, r_{n+1}). The proof of this is again left as an exercise to the reader. Note that there are only finitely many possible pairs (a_n, r_n), so once it cycles, it would never leave the cycle, and the sequence of pairs (a_n, r_n) is eventually periodic. Hence, the sequence (a_n) is eventually periodic and our proof is complete.

Exercise 3.7. Prove that it is possible to uniquely compute (a_{n+1}, r_{n+1}) given (a_n, r_n). □

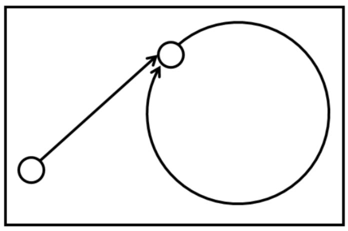

Figure 5: Visualization of cycling

Discussion. Although the above problem is difficult and contains a long solution with many details, the motivation is simple: distinguish the behavior in the scenarios $a_n \mid b_n$ and $a_n \nmid b_n$. After noticing this and dividing the sequence as follows, the details are simply routine calculations.

This heuristic is relatively less common compared to the others. Below is a problem discussed earlier in "reinterpreting the problem", but also contains ideas related to focusing on subsequences.

Related problems

A.5. ISL 2003 A3 (M4) – *focus on chunks of a sequence to quickly prove divergence*

4 Investigating dynamics or processes

Investigating the dynamics or processes involved in a sequence is equivalent to gaining a deep understanding of it. It involves examining the given information about the sequence to deduce properties about it and to have a more intuitive understanding of how the sequence grows as the indices grow large. The slope and rate of change and convergence are key concepts to understanding a sequence's long term behavior. Also essential in understanding the problem are investigating the desired outcome (what we want to prove) as well as how a sequence's first terms affect its long-term behavior.

4.1 Determining the rate of change

This heuristic is applicable to sequences that grow at a linear or approximately linear rate. It involves examining the given information about the sequence and substituting several example sequences to make an educated guess on the approximate slope of the sequence. The example below is a fairly simple problem utilizing this heuristic, and interprets a given quantity as a common difference of a sequence.

Example 4.1 (APMO 2013 P3)**.** For $2k$ real numbers $a_1, a_2, \ldots, a_k,\ b_1, b_2, \ldots, b_k$ define a sequence of numbers X_n by

$$X_n = \sum_{i=1}^{k} [a_i n + b_i] \quad (n = 1, 2, \ldots).$$

If the sequence X_N forms an arithmetic progression, show that $\sum_{i=1}^{k} a_i$ must be an integer. Here $[r]$ denotes the greatest integer less than or equal to r.

Solution. We observe that the problem statement involves the floor function. Similar to a previous problem, Example 3.1, we examine the behavior of the sequence without the floor function. The sequence becomes

$$X_n' = \sum_{i=1}^{k} (a_i n + b_i) = n \sum_{i=1}^{k} a_i + \sum_{i=1}^{k} b_i,$$

which is an arithmetic sequence with common difference $\sum_{i=1}^{k} a_i$. Thus both X_n and X_n' are arithmetic sequences. They are related as follows: for every i, $X_i' < X_i < X_i' + n$. (The reader is encouraged to prove this, using the fact $x - 1 < \lfloor x \rfloor \le x$.) This makes it apparent that the common difference of X_i is close to that of X_i'.

The problem wants us to show that $\sum_{i=1}^{k} a_i$, which is the common difference of X_i', is an integer. As X_n is an integer arithmetic sequence, finding its common difference (which is an integer) would be helpful. In particular, if we are able to show that X_n has the same common difference as X_n' (which is $\sum_{i=1}^{k} a_i$), then $\sum_{i=1}^{k} a_i$ must be an integer, and we are done.

Observe that since both (X_n) and (X'_n) are arithmetic sequence, the difference between them (defined as the difference between the corresponding terms), is also an arithmetic sequence. Call this sequence as (D_n) As $X'_i < X_i < X'_i + n$, however, (D_n) are bounded below by 0 and above by n, then (D_n) must be constant (otherwise, after a sufficient number of increases/decreases by a fixed common difference, the value of D_i would either go below 0 or above n). Thus, (X_n) and (X'_n) have the same common difference. □

Discussion. The focus of the solution to this problem is finding the common difference of (X_n). As we are aware of a closely-related arithmetic sequence, namely (X'_n) that must have a very close common difference, it becomes natural to hypothesize that the common differences of (X_n) and (X'_n) are equal, which soon solves the problem.

This idea is also extremely useful in more advanced problems where the is a sequence whose slope / rate of change is roughly linear (meaning that small, local discrepancies are allowed - recall the idea from the section on "taking a global view"). In some problems, such as ISL 2010 A7, it is unclear from the given that the sequence increases roughly linearly, but it becomes apparent when one works backwards from the desired conclusion.

Related problems

A.7. ISL 2009 A6 (M3) – *work backwards*

A.8. ISL 2010 A7 (M3) – *work backwards*

A.14. ISL 2015 C5 (M3) – *it is zero, so what?*

A.21. PFTB Example 2 (M6) – *this is the convergent sequence*

4.2 Proving convergence

Convergence is essentially a sequence approaching a certain value in the long-term, meaning that after some point, the difference between the sequence's terms to the suggested value would be below a certain threshold. An important theorem in proving convergence in real-numbered sequence is the *Bolzano-Weierstrass Theorem*, which states that a monotonic bounded sequence must be convergent. A visualization of a bounded, monotonic sequence is shown.

Example 4.2 (PEM Handout P5). Prove that for $n \geq 2$, the equation $x^n + x - 1 = 0$ has a unique root in the interval $(0,1]$. If x_n denotes this root, prove that the sequence $(x_n)_n$ is convergent.

Solution. It is clear that x is bounded because x_n is always within $[0,1]$, so it suffices to prove monotonicity (then it is bounded by Bolzano-Weierstrass). For $n = 1$, our root is simply $2x - 1 = 0 \to x = \frac{1}{2}$, and for $n = 2$, our root is $x^2 + x - 1 = 0 \to x = \frac{\sqrt{5}-1}{2} \approx 0.618$.

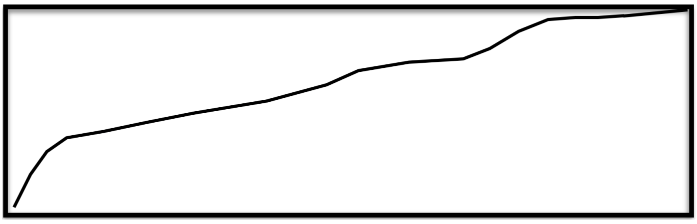

Figure 6: Visualization of a convergent, monotonic sequence

We hypothesize that for all $n \geq 1$, the root of $x^n + x - 1$ is less than or equal to the root of $x^{n+1} + x - 1$. Proving this is a simple contradiction. Suppose that we have $x_n > x_{n+1}$ and

$$x_n^n + x_n - 1 = x_{n+1}^{n+1} + x_{n+1} - 1 = 1.$$

As $x_n > x_{n+1}$ and $x_n^n \geq x_{n+1}^n > x_{n+1}^{n+1}$, we combine them to get $x_n^n + x_n > x_{n+1}^{n+1} + x_{n+1}$, which contradicts $x_n^n + x_n - 1 = x_{n+1}^{n+1} + x_{n+1} - 1$. □

Discussion. This is a standard exercise in using the Bolzano-Weierstrass Theorem, which breaks down convergence into two conditions that are much easier to prove. The proofs, however, may not be trivial and may involve careful examination of the sequence and some bounding, just as seen in the problem.

For integer sequences, the Bolzano-Weierstrass Theorem is not necessary to establish the same conclusion; it suffices to notice that at some point, our sequence must stop increasing or decreasing and become constant (otherwise the whole number differences would accumulate past the bound). Take note that these are not the only ways to prove convergence, as it is possible for non-monotonic sequences to be convergent: a proof that the difference to the value the terms converge to would get infinitesimally small.

Related problems

A.5. USA TST 2008 P5 (M5) – *finding a monovariant that takes on integers and is bounded below*

A.9. ISL 2012 A4 (S4) – *Use Bolzano Weierstrass to prove a crucial lemma on bounded numerator and denominator sequence of a rational number sequence*

A.12. ISL 2014 A1 (M5) – *anti-example: strictly monotonic function on integers*

A.23. PFTB Example 5 (M7) – *find a quantity that converges to zero*

4.3 Interpreting the desired outcome

This heuristic is analogous to working backwards from what the problem asks to prove and using this to identify an easier-to-prove statement about the sequence in consideration.

It may also involve using information about some eventual large index (i.e. the desired outcome) to deduce information about the first terms of the sequence, which is extremely helpful in problems where the solver is asked to find sequences that satisfy a certain condition. Below is an example of such a problem.

Example 4.3 (Benelux 2011 P3)**.** If k is an integer, let $\mathrm{c}(k)$ denote the largest cube that is less than or equal to k. Find all positive integers p for which the following sequence is bounded: $a_0 = p$ and $a_{n+1} = 3a_n - 2\mathrm{c}(a_n)$ for $n \geq 0$.

Solution. One would immediately notice that the condition $a_{n+1} = 3a_n - 2\mathrm{c}(a_n)$ implies that $a_{n+1} \geq a_n$, as $c(a_n) \leq a_n$ with equality case when a_n is a perfect cube. Also, note that if a_n is a perfect cube, then $a_{n+1} = a_n$. a_{n+1}, however, is still a perfect cube, so this implies that a_{n+2} is a perfect cube, and so on, so for all $i \geq 0$, a_{n+i} is a perfect cube. Thus, if one term of the sequence is a perfect cube, then all terms after a particular index will be, and the sequence would be bounded.

It is natural to consider the contrary case, where no terms are a perfect cube. Then the equality case in $a_{n+1} \geq a_n$ would never occur, and the sequence would not stop increasing. As this is a sequence of integers, the sequence would become unbounded. Hence, at least one term of a sequence is a perfect cube. This motivates us to think how one term being a perfect cube could influence previous terms. If p is a perfect cube, then it is easy to see that all succeeding terms would be perfect cubes. We try a non-cube. If $p = 2$, then

$$\begin{aligned} a_0 &= 2 & a_3 &= 3(10) - 2(8) = 14 \\ a_1 &= 3(2) - 2(1) = 4 & a_4 &= 3(14) - 2(8) = 26 \\ a_2 &= 3(4) - 2(1) = 10 & a_5 &= 3(26) - 2(8) = 62. \end{aligned}$$

It appears that the first few a_is are not perfect cubes.

We attempt to prove the stronger claim that if a_i is not a perfect cube, then a_{i+1} would not be a perfect cube. Its contrapositive is that a_{i+1} being a perfect cube would imply that a_i is, and since we know that at least one term of the sequence is a perfect cube, we may induct backwards to show that a_1 is a perfect cube, and the solution is all p that are perfect cubes. (Why?)

Suppose the contrary and assume that a_i is not a perfect cube while a_{i+1} is. Then we may write $k^3 < a_i < (k+1)^3$. Then $a_{i+1} = 3a_i - 2c(a_i) = 3a_i - 2k^3$. We, however, are unable to show that this is less than $(k+1)^3$, failing to prove a bound. We consider terms modulo 3 to get that $a_{i+1} \equiv k^3 \pmod 3$ (why?). But as $a_{i+1} > a_i > k^3$, we need $a_{i+1} \geq (k+3)^3$. Bounding becomes much simpler:

$$\begin{aligned} a_{i+1} &= 3a_i - 2k^3 \\ &< 3(k+1)^3 - 2k^3 \\ &= k^3 + 9k^2 + 9k + 1 \\ &< k^3 + 9k^2 + 27k + 27, \end{aligned}$$

thus $a_{i+1} < (k+3)^3$, giving a contradiction. □

Discussion. Working backwards would provide a natural motivation to the otherwise–unguessable assertion that a_n being a perfect cube would imply that a_{n-1} is a perfect cube, and the proof of this assertion is a simple bounding argument. Essentially, we deduce that at some point, we must have a cube for the sequence to be bounded, and then work backwards from there.

The first problem below involves working backwards to reinterpret the sequence, which may have no clear pattern from its origins at first. Its recursive formula gives multiple possibilities for the term to be decided, of which the maximum is taken. It is helpful to generalize using a backward-forward style, just as done in the .

Related problems

A.8. ISL 2010 A7 (M3) – *interpret "stabilizing" in an easy-to-process form*

A.16. ISL 2015 N6 (S3) – *proving infinitely many satisfies some property; gives an easier-to-prove bound*

A.36. USA TSTST 2011 P8 (S2) – *guess the eventual value; work backwards from there*

4.4 Investigating origins

Many sequences are defined recursively, meaning that terms with sufficiently large indices are computed solely based on previous terms. A well-known example is the Fibonacci sequence, which has a closed-form Binet's formula. Some recursive sequences, however, do not have each term's relation to the starting numbers stated clearly. The example below utilizes this heuristic by relating each term in a subsequence to a chosen set of starting terms. As the proof is rather long and rigorous, details will be left as an exercise to reader, so that the discussion may be focused to the heuristic at hand and the analytic thinking involved.

Example 4.4 (ISL 2012 A6)**.** Let $f : \mathbb{N} \to \mathbb{N}$ be a function, and let f^m be f applied m times. Suppose that for every $n \in \mathbb{N}$ there exists a $k \in \mathbb{N}$ such that $f^{2k}(n) = n+k$, and let k_n be the smallest such k. Prove that the sequence $k_1, k_2, \ldots$ is unbounded.

Solution. The problem statement gives a very loose condition on how a term in the sequence may be related to later terms, so there is much freedom in constructing our sequence. We are unsure which terms are directly related through the condition to a certain term as it depends on the result of several compoundings of f. Looking instead at the problem statement, a likely contradiction is that there is *too little space* for the values $f(1)$, $f^2(1)$, ..., f_n^{2k}, if $f_n^{2k} = n+k$, meaning that the terms in this subsequence must not be increasing, even in the analytical, non-rigid sense. This motivates us to focus on the set $S = \{1, f(1), f^2(a), \ldots\}$, and hope for a contradiction if we assume the hypothesis to be false.

Another benefit of choosing this set is that for every $x \in S$, we are certain that there is another term $y \in S$ that is directly related through the problem condition, and it is possible to access this by simply moving a given number of steps forward. As the only information given about the sequence is the $f^{2k}(n)$ condition, we are inclined to establish a link between n and $f^{2k_n}(n)$. An array of chains visualization seems to be a good idea to visualize the relationship between the elements of S. For this to be possible, however, we need f to be injective on S and for $g(n) = f^{2k_n}(n) = n + k_n$ to also be injective on S. (Why are these needed?) The proofs of these are left as an exercise to the reader.

Exercise 4.5. Prove that f and g are both injective on S. (Hint: use contradiction.)

Given the two conditions, it is possible to construct chains as follows: Take $f(1)$, $g(f(1))$, $g^2(f(1))$, $\ldots$. Repeatedly take the first term in S that does not belong yet in a chain, and build its chain by applying g repeatedly (infinitely many times). Note that as g is injective over S, these chains must be disjoint. Label the set of elements of S used as the starting number of the chains as T. It is clear that this covers all elements of S.

From here, by tracing the value of f back using the equation $f^{2k_n}(n) = n + k$ repeatedly, we are able to show that if $f^{n_t}(1)$ is the start of the chain where $f(n)$ is in, then $f^n(1) = t + \frac{n-n_t}{2}$. (Why?)

Exercise 4.6. Rigorously prove that $f^n(1) = t + \frac{n-n_t}{2}$ by using $n = n_t + 2a_1 + \cdots + 2a_j$, where $f^n(1) = f^{2a_j}(f^{2a_{j-1}}(\cdots f^{2a_1}(f^{n_t}(1))))$.

It appears that we are close, because note that each chain has a portion of the set S that is completely blocked from other chains. As our conclusion wants us to show that the gaps between two successive terms in a chain is unbounded, we prove that there exists an infinitely number of chains, i.e. that T is infinite.

We suppose otherwise. Then if we let M be the last starting point, by $f^n(1) = t + \frac{n-n_t}{2}$, we are sure that $f^i(1) \le t + \frac{n-n_t}{2} \le M + \frac{N}{2}$ for sufficiently large N and $1 \le i \le n$. As we have earlier established that f is injective over S, this needs $N + 1 \le M + \frac{N}{2}$, which yields a contradiction if N is sufficiently large.

Now that we have proven that T is infinite, we only have to show that for any $k \in \mathbb{N}$, there exists a gap between two successive terms in a chain that is of at least length k. For this, we could consider only the first $k+1$ chains. Intuitively, it is apparent that we have k chains to fill in the gaps between the one term in a chain and the next. To rigorize this, consider M again. All of the chains has a term $f^i(1)$, where $i \le M$. By the pigeonhole principle, however, at least one of them does not contain any term among $f^{M+1}(1)$, $f^{M+2}(1)$, $\ldots$ $f^{M+k}(1)$, so we are sure to have an index gap of at least k, and (k_i) is unbounded.

Exercise 4.7. Many details in this example are intuitive rather than rigorous. As this is the last example problem for this article, it is recommended to attempt writing a full, rigorous proof based on the above ideas. □

Discussion. The central idea in this problem is to consider each term relative to some origin. A far more advanced derivative of this problem is ISL 2015 N6, which involves similar ideas. In solving this, however, we are not allowed to focus solely on a subsequence as we have to show a property for the entire sequence.

Related problems

A.8. ISL 2010 A7 (M1)

A.14. ISL 2015 C5 (M4)

A.16. ISL 2015 N6 (M3) – *interpreting functions as chains of input-output-input-output*

A.36. USA TSTST 2011 P8 (M5) – *constructing a term easily based on previous terms*

5 Conclusion and Tips

Upon performing a thematic categorization of the heuristics involved in solutions to thirty-six analysis-flavored olympiad problems, the author discovered that the three main themes (categories of heuristics) are *Taking a Global View*, *Refocusing*, and *Investigating Processes and Dynamics*. Each theme contains three to four heuristics, which are specific techniques or approaches. While these may be viewed as merely techniques, they are in fact the central ideas that appear in the problems. As seen in the previous examples, devising the solution involves understanding the problem and how the given sequence behaves at the large-scale; therefore, knowledge of the techniques and heuristics are insufficient.

The author has also noticed that the more difficult ISL problems, determined by the order it appears at the ISL, require simultaneous or directed thought through multiple themes and may have multiple key steps (e.g. ISL 2010 A7, ISL 2015 N6), unlike the easier problems whose key step rests on one theme or heuristic. This is consistent with the observation in the article *On Reading Solutions*, that the more advanced olympiad problems involve a larger number of main ideas compared to introductory problems (Chen, 2017).

5.1 Recommendations

The author noted several possible improvements to expand upon the ideas in this article and improve its instructional value, which were not possible due to time constraints of the course.

1. Examine the low-level techniques involved in verifying *routine* details of the solution, as this article is focused on high-level problem-solving strategies.
2. Include worked out solutions to advanced problems (at the level of the hardest IMO/ISL problems), which likely incorporate ideas from different themes.
3. Expand on the examples given in the heuristics discussion by explaining how a heuristic may be used as a minor idea in more advanced problems.
4. Classify problems for other prominent mathematical olympiads not included in consideration for the problems discussed in this article.

References

[1] Andreescu, T., & Dospinescu, G. (2010). Problems from the book. Plano (Tex.): XYZ Press.

[2] Asuncion, J., & Atienza, K. (2017). NOI.PH 2017 Training Week 8. Unpublished manuscript, IOI 2017 Training, National Olympiad in Informatics - Philippines, Manila.

[3] Chan Shio, C. (2016, September 03). Problems on Convergence of Sequences. Lecture presented at PEM Advanced Class, 03 September 2016 in Ateneo de Manila University, Quezon City.

[4] Chen, E. (2017, March 9). On Analysis Olympiad Problems [E-mail interview].

[5] Chen, E. (2017, March 6). On Reading Solutions [Web log post]. Retrieved May 5, 2017, from https://usamo.wordpress.com/2017/03/06/on-reading-solutions/

[6] Chen, E. (2017, April 8). Some Thoughts on Olympiad Material Design [Web log post]. Retrieved May 5, 2017, from https://usamo.wordpress.com/2017/03/06/on-reading-solutions/

[7] Wang, V. (2017, March 27). On Analysis-flavored Olympiad Problems [E-mail interview].

[8] Wang, Victor (2014). [MOP Experiment]. Unpublished raw data.

[9] Renze, J., & Weisstein, E. (n.d.). Analysis. Retrieved May 02, 2017, from http://mathworld.wolfram.com/Analysis.html