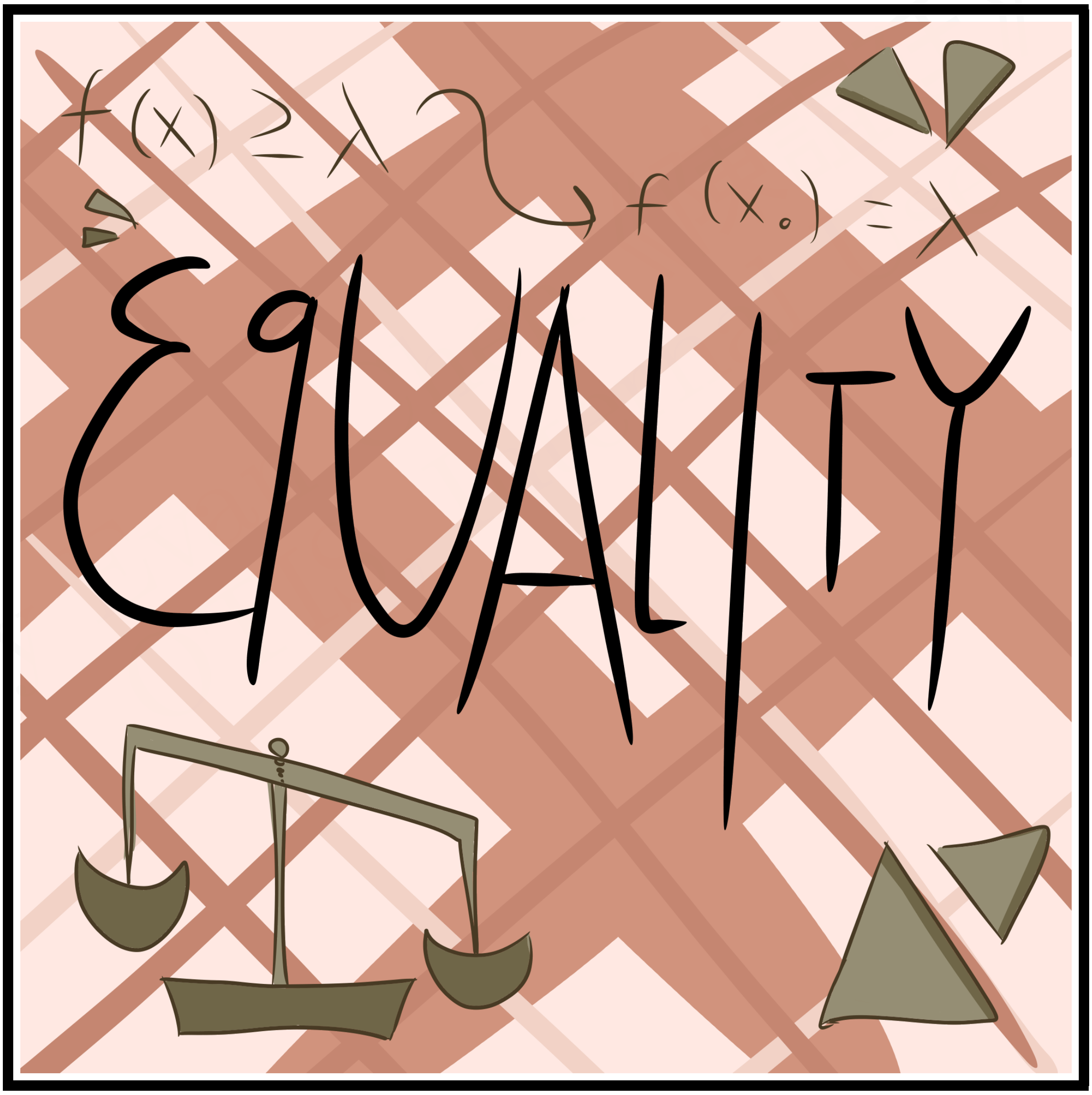
Look at Equality Cases

Evan Chen《陳誼廷》

28 December 2023
BCW-EQUALITY



§1 Lecture Notes

Welcome to Equality, the flagship soft-skill unit (and the first, but not the last, you'll find).

The term "soft skill" is a thing I made up to refer to the general idea of actions that will help you solve a problem (like examining small/special/interesting cases) but do not end up as part of your official solution. If you want, you can read more about what I mean by this at the following URL:

https://blog.evanchen.cc/2019/05/03/hard-and-soft-techniques/

§1.1 Discussion

This lecture focuses on the specific soft skill of looking at the equality cases of an estimate. Unlike things you saw in previous OTIS units, this is not really a technique or theorem, but more philosophical.

The basic premise is that if you know the equality cases of an optimization problem, you can use this information to figure out what to do. In particular, many of the example problems you'll see will have solutions that look they come out of nowhere, until you realize they were engineered based on a good understanding of those equality cases, even though the solution won't mention it.

Concrete advice:

- If some estimate doesn't preserve equality, **it won't work**. For example, the wrong way to start proving Schur's inequality
$$a^3 + b^3 + c^3 + 3abc \ge \sum_{\text{sym}} a^2 b$$
is to try starting with the estimate $a^3 + b^3 \ge a^2 b + ab^2$. This is because Schur has an equality case not preserved by this estimate. (What is it?)

- If something holds at equality, it's worth trying no matter how stupid it looks.

- If you have lots of weird equality cases, stop doing the inequality for at least a little while and try to figure out where the equality cases are coming from.

§1.2 Sharpness principle

There is also a sort of pseudo-theorem I want to mention.

You should by now know that any "find minimum" problem is always a *two-part* problem, where, say

(i) you show that $f(x) \ge \lambda$ for all x (proving a *bound*), and then

(ii) you show that $f(x_0) = \lambda$ for some x_0, showing *attainability*.

Here λ is a constant. Step (ii) upgrades the bound from step (i) into a true minimum, solving the problem.

The sharpness principle is the following philosophy:

> **Sharpness principle**: Before trying to prove the bound, guess what the x_0 in (ii) should be. Then throughout step (i), ensure every estimate used has exact equality at x_0.

If you can do this, then step (ii) will automatically follow. That is, if you get a bound at all, it will automatically be a true minimum. In still other words:

You are promised the constant you want, if you get a constant at all.

This won't make sense without an example; see SL 2010 A2.

§1.3 Warm-up

Example 1.1

Let A and B be finite nonempty sets of real numbers. Let $A + B = \{a + b \mid a \in A, b \in B\}$.

(i) Prove that $|A + B| \geq |A| + |B| - 1$.

(ii) Determine the cases of equality.

SUMSETS

Walkthrough. Let $n = |A|$ and $m = |B|$. We begin by showing (i).

(a) Find examples of equality cases which achieve $n + m - 1$.

(b) Assume $a_1 < \cdots < a_n$ and $b_1 < \cdots < b_m$. Explicitly list $n + m - 1$ elements in $A + B$.

This solution for (i) might feel infantile, but roughly the reason it works is that the equality case (in which $|A+B|$ is linear) is way different from the "generic" case in which we expect $|A + B| = \Omega(nm)$.

Part (ii) is more subtle. Let's write out all mn sums:

$$\begin{bmatrix} a_1 + b_1 & a_1 + b_2 & a_1 + b_3 & \dots & a_1 + b_m \\ a_2 + b_1 & a_2 + b_2 & a_2 + b_3 & \dots & a_2 + b_m \\ a_3 + b_1 & a_3 + b_2 & a_3 + b_3 & \dots & a_3 + b_m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_n + b_1 & a_n + b_2 & a_n + b_3 & \dots & a_n + b_m \end{bmatrix}$$

The idea behind (b) was to just take a path starting from $a_1 + b_1$ to $a_n + b_m$. It increased.

(c) Show that if $|A + B| = n + m - 1$ then in fact the diagonals are constant.

(d) Prove that if $\min(m, n) > 1$ then A and B are arithmetic progressions.

(e) What if $\min(m, n) = 1$?

(f) Write down the final answer to (ii).

§1.4 Algebraic examples

Example 1.2 (MOP 2003)

Prove that for $a, b, c > 0$ we have

$$\sum_{\text{sym}} (a^4 b^2 + 2a^3 b^2 c) \geq \sum_{\text{sym}} (a^4 bc + a^3 b^3 + a^2 b^2 c^2).$$

03MOP

Walkthrough. Write the inequality as LHS − RHS ≥ 0. To help visualize the problem, we will use the *Chinese dumbass notation* in which we write the coefficient of each monomial in a triangle. The triangle is formatted as follows, where $[p]$ denotes the coefficient of the monomial p.

$$\begin{array}{ccccccccccccc}
&&&&&&[a^6]&&&&&& \\
&&&&&[a^5b]&&[a^5c]&&&&& \\
&&&&[a^4b^2]&&[a^4bc]&&[a^4c^2]&&&& \\
&&&[a^3b^3]&&[a^3b^2c]&&[a^3bc^2]&&[a^3c^3]&&& \\
&&[a^2b^4]&&[a^2b^3c]&&[a^2b^2c^2]&&[a^2bc^3]&&[a^2c^4]&& \\
&[ab^5]&&[ab^4c]&&[ab^3c^2]&&[ab^2c^3]&&[abc^4]&&[ac^5]& \\
[b^6]&&[b^5c]&&[b^4c^2]&&[b^3c^3]&&[b^2c^4]&&[bc^5]&&[c^6]
\end{array}.$$

Written this way, the inequality becomes

$$\begin{array}{ccccccccccccc}
&&&&&&0&&&&&& \\
&&&&&0&&0&&&&& \\
&&&&1&&-2&&1&&&& \\
&&&-2&&2&&2&&-2&&& \\
&&1&&2&&-6&&2&&1&& \\
&0&&-2&&2&&2&&-2&&0& \\
0&&0&&1&&-2&&1&&0&&0
\end{array} \ge 0.$$

For example, the third row says that the coefficients of a^4b^2, a^4bc, a^4c^2 are respectively 1, -2, 1.

(a) Find any example of an equality case other than $a = b = c$. (It's permissible for a, b, c to be zero, since if the inequality holds for positive reals it should certainly hold for nonnegative reals too.)

(b) Study the rows in the triangle above; what are the sums of the rows? Find a larger curve of equality cases. You should find at least a two-dimensional curve.

(c) Let $f(a, b, c) = \text{LHS} - \text{RHS}$. Based on your answer to (b), what polynomials must divide f?

Important intuition to note: let $p(x)$ be a one-variable polynomial such that $p(x) \ge 0$ for all $x \in \mathbb{R}$, and p has a root at r. Then r must be a double root of p (the graph of f is tangent to x-axis at $x = r$). By using this lemma, you should be able to deduce square factors which must divide f in order for the problem to possibly be true.

(d) Find the factorization of f.

This is an example of a problem where merely discovering the entire curve of equality cases is somehow enough to extract a complete solution.

Example 1.3 (Shortlist 2010 A2)

Let a, b, c, d be real numbers satisfying $a + b + c + d = 6$ and $a^2 + b^2 + c^2 + d^2 = 12$. Prove that

$$36 \le 4(a^3 + b^3 + c^3 + d^3) - (a^4 + b^4 + c^4 + d^4) \le 48.$$

10SLA2 **Walkthrough.** This is such a poster example that I had to include it despite it being an algebra problem: it's perfect in so many ways.

The first part is easy:

(a) Find the quadruplets (a, b, c, d) achieving 36 and 48. (They are integers.)

In some sense the solution that follows is impossible to motivate without this step.

Now, let's prove one direction.

(b) Come up with an inequality of the form

$$x^4 - 4x^3 + kx^2 + \ell x + m \geq 0 \qquad \forall x \in \mathbb{R}$$

which is sharp (equality holds) when $x \in \{0, 2\}$. (There is only one possible answer.)

(c) *Without writing anything down or doing any calculation*, figure out what the result *must* be when we sum the inequality in (b) cyclically. (This is an application of Sharpness Principle.)

(d) Now actually carry out the calculation and verify that it matches the result of (c).

It's worth noting (if you haven't seen already) that we didn't even use the condition $a + b + c + d = 6$ yet. We'll need to for the other bound, though.

(e) Show that the approach of (c) won't work for the other bound: one can't conjure an inequality as in (b) which is true for all $x \in \mathbb{R}$, sharp at $x \in \{1, 3\}$.

However, we make the following observation: we know that

$$(b + c + d)^2 \leq 3(b^2 + c^2 + d^2)$$

and thus we should be able to get a one-variable inequality in a; this will give some constraints on possible values of a.

(f) It turns out that a will lie in some interval $[u, v]$. *Without writing anything down or doing any calculation*, figure out what u and v *must* be.

(g) Now actually carry out the calculation and verify that it matches the result of (f).

We now have an extra condition that will pay off.

(h) Come up with an inequality of the form

$$x^4 - 4x^3 + kx^2 + \ell x + m \leq 0 \qquad \forall x \in [u, v]$$

which is sharp when $x \in \{1, 3\}$. (There is only one possible answer).

(i) *Without writing anything down or doing any calculation*, figure out what the result *must* be when we sum the inequality in (h) cyclically.

(j) Now actually carry out the calculation and verify the result of (i).

Thus we see the Sharpness Principle is used three times: once in part (c), once in part (f), and once in part (i).

Example 1.4 (Putnam 2014 B2)

Suppose that f is an integrable function on the interval $[1, 3]$ such that $-1 \leq f(x) \leq 1$ for all x and $\int_1^3 f(x)\, dx = 0$. Determine the largest possible value of

$$\int_1^3 \frac{f(x)}{x}\, dx.$$

14PTNMB2 **Walkthrough.** Don't be scared of the integral. You won't need much calculus at all; only two facts will really be used.

- $\int 1/x \, dx = \log x$
- $\int_a^b u(x) \, dx \le \int_a^b v(x) \, dx$ if $u \le v$.

First let's figure out the answer.

(a) Guess the function g which obtains the maximum. (The function is not continuous).

(b) Verify that $\int_1^3 g(x)/x \, dx = \log(4/3)$.

Reasonable, right? We have one direction; okay, now let's solve the problem. The point is that we will prove the inequality

$$\int_1^3 \frac{g(x) - f(x)}{x} \, dx \ge 0.$$

In particular, you won't use your answer to (b) at all. The point is to find the function with the desired *properties*.

(c) Split the integral into $[1, 2]$ and $[2, 3]$ and expand out the definition of g.

(d) Bound $\frac{g(x)-f(x)}{x}$ for $1 \le x < 2$ by an expression in terms of $f(x)$ which doesn't have a denominator.

(e) Do the same for $2 < x \le 3$.

(f) Add (d) and (e) and use $\int_1^3 f(x) \, dx = 0$ to finish.

§1.5 Combinatorial examples

The first one actually might be familiar; it comes from one of the versions of the Global unit.

Example 1.5 (ELMO 2013/1)
Let $a_1, a_2, \dots, a_9$ be nine real numbers, not necessarily distinct, with average m. Let A denote the number of triples $1 \le i < j < k \le 9$ for which $a_i + a_j + a_k \ge 3m$. What is the minimum possible value of A?

13ELMO1 **Walkthrough.** We say a triple $t = (a_i, a_j, a_k)$ is *large* if $a_i + a_j + a_k \ge 3m$.

(a) Show that among any three disjoint triples, at least one triple is large. Give a heuristic argument why we expect $A \ge 28$ as a result.

(b) Give a construction for $A = 28$. (Try making one element large.)

(c) We now proceed to the "global" idea of looking at every possible partition in (a) at once. Show that there are

$$C = \frac{1}{3!}\binom{9}{3, 3, 3} = 280$$

ways to partition the 9 elements into three disjoint triples.

(d) How many of the C partitions does each triple t appear in?

(e) Use your answer to (d) to prove $A \geq 28$, thereby solving the problem.

(f) Optionally, for an alternate solution, explicitly construct a partition of the $\binom{9}{3} = 84$ triples into 28 disjoint triples. This would give another proof that $A \geq 28$.

When doing this calculation for the first time, you might be surprised that the division of seemingly random constants ends up with 28 in the end. It's important to recognize that the argument in (e) is "guaranteed" to work in a sense.

To elaborate: we constructed in (b) an example of an equality case, and every estimate we used was sharp. At the end of (e) we get some number again. The existence of the equality case means that this number *must* match the corresponding constant in (a), namely 28. This point is one of the key ideas in the Equality unit; the so-called "Sharpness Principle".

Finally, here is an absurd example which is going to really exploit the idea.

Example 1.6 (EGMO 2013/6)

Snow White and the Seven Dwarves are living in their house in the forest. On each of 16 consecutive days, some of the dwarves worked in the diamond mine while the remaining dwarves collected berries in the forest. No dwarf performed both types of work on the same day. On any two different (not necessarily consecutive) days, at least three dwarves each performed both types of work. Further, on the first day, all seven dwarves worked in the diamond mine. Prove that, on one of these 16 days, all seven dwarves were collecting berries.

13EGMO6

Walkthrough. We'll view the problem in terms of 16 binary strings of length 7, with 0000000 corresponding to all seven working in the diamond mine. Thus any two of the binary strings must differ in at least 3 positions. Let $v_1, \dots, v_{16} = 0000000$ denote these binary strings and let n_i be the number of 1's in the ith string. (Hence $n_{16} = 0$, and $n_i \geq 3$ for $i = 1, \dots, 15$.)

(a) Show that *if* the problem statement holds, then the multiset $\{n_i\}$ equals

$$\{0, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 7\}.$$

(There are seven 3's and seven $4'$'s.) Essentially, the problem statement asks to show that if v is in the set then so is its bitwise complement.

Thus, our main goal is to show that indeed $n_i \in \{3, 4, 7\}$ for all $i \neq 16$, i.e. to prove the conjecture from (a) which we know has to be true. To do this, we will use the following very artificial inequality:

$$0 \geq (n_i - 3)(n_i - 4)(n_i - 7) = n_i^3 - 14n_i^2 + 61n_i - 84 \qquad (\heartsuit)$$

with equality when $n_i \in \{3, 4, 7\}$ only. Let's get the ingredients we need to make this work.

(b) Show that exactly two vectors start with each of the eight prefixes 000, 001, ... (To do this, assume for contradiction three vectors all start with 000, say, and derive a contradiction.)

(c) Using either (a) or (b), compute $\sum_{i=1}^{16} n_i$.

(d) Generalize this by using (b) to show that $\sum_{i=1}^{16} \binom{n_i}{2} = 84$ and $\sum_{i=1}^{16} \binom{n_i}{3} = 70$.

(e) Using the answer to (d), compute $\sum_{i=1}^{16} n_i^2 = 224$ and $\sum_{i=1}^{16} n_i^3 = 980$.

(f) Take the sum of ($\heartsuit$) across $1 \le i \le 15$. One should encounter the expression

$$1 \cdot 980 - 14 \cdot 224 + 61 \cdot 56 - 15 \cdot 84 = 0. \qquad (\spadesuit)$$

Use this along with (b) and (d) to conclude that $n_i \in \{3, 4, 7\}$ for all i.

(g) Show that $n_i = 7$ for some i, solving the problem.

The heart of this solution is the inequality ($\heartsuit$). We contrive an inequality which we know will be sharp in the equality case we desire, and then combine it with the global information from (b) to force the solution to work. Thus ($\spadesuit$) is not a coincidence; we know *a priori* that the numbers will work out by Sharpness Principle, and this extreme example is hand-picked to show just how far we can take this idea.

§1.6 (Stray digression) Food for thought from a YouTube video on the Graham-Pollak theorem

Royce Yao forwarded the following comment to me from `https://youtu.be/6ByCnuCbPsg`. That video talks about a certain graph-theory theorem which is simple to state, but for which all known proofs seem to use linear algebra, and no combinatorial proof is known right now.

The theorem is known as Graham-Pollak: it states that K_n can't be partitioned into fewer than $n-1$ complete bipartite graphs. You can watch the video for an explanation if these graph theory terms are new to you.

What's the connection to the Equality unit? It's the top comment on the video from Yan Sheng Ang, reproduced below:

> *Here's a guideline for solving combinatorics problems that served me well in my olympiad years:*
>
> - *If the problem has only one (or very few) equality cases/optimal solutions, the proof will likely be combinatorial in nature.*
> - *If the problem has many equality cases/optimal solutions, the proof will likely be algebraic in nature.*
>
> *The Graham–Pollak theorem is definitely of the latter type. There are so many different ways of writing K_n as the edge-disjoint union of $n-1$ many complete bipartite graphs, and these ways have hardly any common structure, so it's very hard for a combinatorial argument to be tight at all of these ways.*
>
> *I think we can see that in the many proof attempts in this comment section: they try to argue that "all optimal solutions must look like [some combinatorial characterisation], so there must be at least $n-1$ parts", but invariably the claimed characterisation excludes some optimal examples.*

§1.7 Difficult example

The following example is fairly technically involved, so it might be best to come back to this after a few practice problems. (This is also a required practice problem in DCW-equality.)

Example 1.7 (Shortlist 2018 C5)

Let k be a positive integer. Evan is organizing a round-robin StarCraft tournament with $2k$ players. The tournament is played over $\binom{2k}{2}$ consecutive days, with one match on each day. Each player arrives at the hotel on the day of their first match, and leaves on the day of their last match. For each day, Evan must pay 1 coin per player in the hotel (including arrival and departure days). Determine the least possible number of coins Evan must pay.

18SLC5

Walkthrough. Let us denote the arrival and departure dates as

$$1 = a_1 \le a_2 \le \dots \le a_{2k}$$
$$\binom{2k}{2} = b_1 \ge b_2 \ge \dots \ge b_{2k}$$

in *sorted* order. Note this does not necessarily corresponding to players!

(a) Show that, despite the fact that b_n and a_n might be referring to different people, we still have
$$S = \sum_n (b_n - a_n + 1).$$

The key to the problem is figuring out what the optimal situation is. Here is one optimal construction for $k = 4$, with the eight players labeled A, B, C, D, W, X, Y, Z. The arrival and departure days are highlighted.

			Day 7	**W**	A			
			Day 8	W	B			
			Day 9	W	C			
			Day 10	W	D			
			Day 11	**X**	A			
Day 1	**A**	**B**	Day 12	X	B	Day 23	W	X
Day 2	**C**	A	Day 13	X	C	Day 24	W	Y
Day 3	C	B	Day 14	**Y**	A	Day 25	**W**	Z
Day 4	**D**	A	Day 15	Y	B	Day 26	X	Y
Day 5	D	B	Day 16	**Z**	**A**	Day 27	**X**	Z
Day 6	D	C	Day 17	Z	**B**	Day 28	**Y**	**Z**
			Day 18	Y	C			
			Day 19	Z	**C**			
			Day 20	X	D			
			Day 21	Y	D			
			Day 22	Z	**D**			

(b) Come up with similar examples for $k = 1, 2, 3$ which achieve the optimal values 2, 17, 57. Then generalize to all k.

(c) Is the construction we gave for $k = 4$ unique? If not, write down a few more examples.

Once you have this, you can start playing around to try and see why it's not possible to improve. In fact it's not too bad to do so.

(d) For $k \ge 3$, I claim that $a_5 \le 7$, i.e. the fifth arrival can happen no later than day 7. Why is that?

(e) Generalize (d): for each $n = 1, \dots, 2k$, find the maximum possible value of a_n (i.e. the latest we could postpone the nth arrival).

(f) For each $n = 1, \dots, 2k$, find the minimum possible value of b_n (i.e. the earliest we could kick a player out of the hotel).

(g) Add the answers to (e) and (f) to find a polynomial F (depending on k) such that

$$b_n - a_n + 1 \ge F(n)$$

must hold for all n.

(h) In our equality case of (b), for which n is that bound achieved?

So for a lot of values of n, the bound $F(n)$ has been proved and is in fact sharp. But not all n; we need to deal with the rest.

(i) For all the n not covered by (h), try to show that the bound of $F(n)$ can be improved to a function $G(n)$ which *is* sharp at the equality case.

The idea is that in your proofs of (d) and (e), some games were probably counted in both, and you can try to subtract them out.

(j) Sum up the modified bound for these other n to get the answer $F(1) + F(2) + \cdots + G(n-1) + G(n)$ (the turning point from F to G depends on your answer to (h)). It should be a cubic polynomial in k.

(k) Verify (admittedly painfully) that the bound you got has the same score S as the equality case. You know this will work in advance, by the Sharpness Principle; it's just a matter of actually grinding the algebra out.

§2 Practice Problems

Instructions: Solve [35♣]. If you have time, solve [45♣]. Problems with red weights are mandatory.

> Look, I'm going to make this simple for you. You've got two choices: YES or YES?
>
> Yes or Yes, by TWICE

Z3E0ED6A

[3♣] **Problem 1** (Bhavya Tiwari). Let $x_1 \le x_2 \le \dots \le x_{100}$ be real numbers with sum 0. Prove that

$$x_1x_{100} + x_2x_{99} + \dots + x_{50}x_{51} \le -50 \cdot x_{50}^2.$$

10JMO2

[2♣] **Problem 2** (JMO 2010/2). Let $n > 1$ be an integer. Find, with proof, all sequences $x_1, x_2, \dots, x_{n-1}$ of positive integers with the following three properties:

(a) $x_1 < x_2 < \dots < x_{n-1}$;

(b) $x_i + x_{n-i} = 2n$ for all $i = 1, 2, \dots, n-1$;

(c) given any two indices i and j (not necessarily distinct) for which $x_i + x_j < 2n$, there is an index k such that $x_i + x_j = x_k$.

16EGMO1

[3♣] **Problem 3** (EGMO 2016/1). Let n be an odd positive integer, and let $x_1, x_2, \dots, x_n$ be nonnegative real numbers. Show that

$$\min(x_i^2 + x_{i+1}^2) \le \max(2x_jx_{j+1})$$

where $1 \le i, j \le n$ and $x_{n+1} = x_1$.

06AMO2

[3♣] **Problem 4** (USAMO 2006/2). Let $k > 0$ be a fixed integer. Compute the minimum integer N (in terms of k) for which there exists a set of $2k+1$ distinct positive integers that has sum greater than N, but for which every subset of size k has sum at most $N/2$.

KAZ

[2♣] **Problem 5** (Kazakhstan). You are given 200 coins of value 1, 200 coins of value 2, and 200 coins of value 5. Determine the maximum possible number of disjoint groups you can create such that each group has total value 9.

19PTNMA3

[2♣] **Problem 6** (Putnam 2019 A3). Given real numbers $b_0, b_1, \dots, b_{2019}$ with $b_{2019} \neq 0$, let $z_1, z_2, \dots, z_{2019}$ be the roots in the complex plane of the polynomial

$$P(z) = \sum_{k=0}^{2019} b_k z^k.$$

Let $\mu = (|z_1| + \dots + |z_{2019}|)/2019$ be the average of the distances from $z_1, z_2, \dots, z_{2019}$ to the origin. Determine the largest constant M such that $\mu \ge M$ for all choices of $b_0, b_1, \dots, b_{2019}$ that satisfy

$$1 \le b_0 < b_1 < b_2 < \dots < b_{2019} \le 2019.$$

14TWNQ3J5

[5♣] **Required Problem 7** (Taiwan TST Quiz). Positive integers $x_1, x_2, \dots, x_n$ ($n \ge 4$) are arranged in a circle such that each x_i divides the sum of the neighbors; that is,

$$\frac{x_{i-1} + x_{i+1}}{x_i} = k_i$$

is an integer for each i, where $x_0 = x_n$, $x_{n+1} = x_1$. Prove that

$$2 \le \frac{k_1 + \dots + k_n}{n} < 3.$$

16SLC5

[9♣] **Problem 8** (Shortlist 2016 C5)**.** Let $n \geq 3$ be a positive integer. Determine the maximal number of diagonals of a regular n-gon that can be drawn such that: any two drawn diagonals which intersect in the interior of the n-gon are perpendicular.

20SRBJST3

[2♣] **Problem 9** (Serbia JBMO TST 2020, added by Lum Jerliu)**.** Given are real numbers $a_1, a_2, \dots, a_{101}$ from the interval $[-2, 10]$ such that their sum is 0. Prove that the sum of their squares is strictly less than 2020.

Z7D3D3AD

[2♣] **Problem 10** (Added by Bhavya Tiwari)**.** Let $e > 0$ be a real number such that $e^x \geq x + 1$ holds for all real numbers x. Show directly (i.e. without using calculus) that $t^{1/t} \leq e^{1/e}$ for all real numbers $t > 0$.

09USATST7

[3♣] **Required Problem 11** (USA TST 2009/7)**.** Find all real numbers x, y, z which satisfy

$$\begin{aligned} x^3 &= 3x - 12y + 50, \\ y^3 &= 12y + 3z - 2, \\ z^3 &= 27z + 27x. \end{aligned}$$

11CSUR6

[5♣] **Problem 12** (Cono Sur 2011/6)**.** Let Q be a $(2n+1) \times (2n+1)$ board. Some of its cells are colored black in such a way that every 2×2 board of Q has at most 2 black cells. Find the maximum amount of black cells that the board may have.

16HRVTST22

[5♣] **Required Problem 13** (Croatia TST 2016)**.** Let $n > 1$ be an integer. We have an $n \times n$ board with two diagonally opposite corner cells colored in black; the rest is colored white. A legal move consists of picking row or column and inverting the colour of cells which are in that row/column. What is the least additional number of cells that we need to colour black so that we can eventually turn all cells black?

ZD4AC005

[9♣] **Problem 14.** Find the smallest possible value of

$$S = \sum_{k=1}^{10} a_k^2 + \left(\sum_{k=1}^{10} a_k \right)^2$$

over all integers $a_1, \dots, a_{10}$ satisfying $a_1 + 2a_2 + 3a_3 + \dots + 10a_{10} = 111$.

11RUS93

[5♣] **Problem 15** (Russia 2011/9.3)**.** A convex 2011-gon is drawn on the board. Peter keeps drawing its diagonals in such a way that each newly drawn diagonal intersected no more than one of the already drawn diagonals (two diagonals which share a common vertex are not considered to intersect). What is the greatest number of diagonals that Peter can draw?

99CAN4

[2♣] **Problem 16** (Canada 1999/4, added by Haozhe Yang)**.** Suppose $a_1, a_2, \dots, a_8$ are eight distinct integers from $\{1, 2, \dots, 16, 17\}$.

(a) Show that there is an integer $k > 0$ such that the equation $a_i - a_j = k$ has at least three different solutions.

(b) Find a specific counterexample when 8 is replaced by 7.

21SLC2

[9♣] **Problem 17** (Shortlist 2021 C2)**.** Let $n \geq 3$ be a fixed integer. There are $m \geq n+1$ beads on a circular necklace. You wish to paint the beads using n colors, such that among any $n + 1$ consecutive beads every color appears at least once. Compute the number of values of m for which this task is *not* possible.

13SLA4

[9♣] **Required Problem 18** (Shortlist 2013 A4)**.** Let n be a positive integer, and consider a sequence $a_1, a_2, \dots, a_n$ of positive integers. Extend it periodically to an infinite sequence $a_1, a_2, \dots$ by defining $a_{n+i} = a_i$ for all $i \ge 1$. If $a_1 \le a_2 \le \dots \le a_n \le a_1 + n$ and

$$a_{a_i} \le n + i - 1 \quad \text{for } i = 1, 2, \dots, n,$$

prove that $a_1 + \dots + a_n \le n^2$.

[1♣] **Mini Survey.** Fill out feedback on the OTIS-WEB portal when submitting this problem set. Any thoughts on problems (e.g. especially nice, instructive, easy, etc.) or overall comments on the unit are welcome.

In addition, if you have any suggestions for problems to add, or want to write hints for one problem you really liked, please do so in the ARCH system!

The maximum number of [♣] for this unit is [81♣], including the mini-survey.

§3 Solutions to the walkthroughs

§3.1 Solution 1.1

First, we prove the inequality in (i). Suppose $a_1 < a_2 < \dots < a_n$, and $b_1 < \dots < b_m$. Then

$$\begin{aligned} a_1 + b_1 &< a_1 + b_2 < a_1 + b_3 < \dots \\ &< a_1 + b_m \\ &< a_2 + b_m < a_3 + b_m < \dots < a_n + b_m \end{aligned}$$

is a strictly increasing sequence of $n + m - 1$ elements.

For (ii), we claim that equality holds if and only if either one of the sets is a singleton, or both sets are arithmetic progressions with the same common difference (which both work). We assume $m, n > 1$ since otherwise there is nothing to prove.

Maintaining the same order relation, arrange all the sums in the following table:

$$\begin{bmatrix} a_1 + b_1 & a_1 + b_2 & a_1 + b_3 & \dots & a_1 + b_m \\ a_2 + b_1 & a_2 + b_2 & a_2 + b_3 & \dots & a_2 + b_m \\ a_3 + b_1 & a_3 + b_2 & a_3 + b_3 & \dots & a_3 + b_m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_n + b_1 & a_n + b_2 & a_n + b_3 & \dots & a_n + b_m \end{bmatrix}$$

Let D_k for $k \in \{2, \dots, m+n\}$ be the diagonal with entries $a_i + b_j$ with $i + j = k$; this is a diagonal of the table going from bottom left to upper right. In general, no two diagonals have common entries for size reasons. The previous proof proceeded by picking an element from every diagonal, and in fact we can get multiple different proofs of (i) by picking different paths south-east paths from $a_1 + b_1$ to $a_n + b_m$, e.g.

$$\underbrace{a_1 + b_1}_{\in D_2} < \underbrace{a_1 + b_2}_{\in D_3} < \underbrace{a_2 + b_2}_{\in D_4} < \underbrace{a_2 + b_3}_{\in D_5}$$

As there are $m + n - 1$ diagonals, if $|A + B| = m + n - 1$ then in fact each diagonal is *constant*, i.e. the value of $a_i + b_j$ depends only on $i + j$. If $m, n > 1$ this means in particular that $a_1 + b_j = a_2 + b_{j-1}$ for each $j \ge 2$, which means B is an arithmetic progression with common difference $a_2 - a_1$. Similarly, A is an arithmetic progression with common difference $b_2 - b_1$, and we're done.

§3.2 Solution 1.2, MOP 2003

In triangle notation,

$$\begin{array}{ccccccccccccc} &&&&&& 0 &&&&&& \\ &&&&& 0 && 0 &&&&& \\ &&&& 1 && -2 && 1 &&&& \\ &&& -2 && 2 && 2 && -2 &&& \\ && 1 && 2 && -6 && 2 && 1 && \\ & 0 && -2 && 2 && 2 && -2 && 0 & \\ 0 && 0 && 1 && -2 && 1 && 0 && 0 \end{array} \ge 0.$$

It's equivalent to $(a-b)^2(b-c)^2(c-a)^2 \ge 0$.

§3.3 Solution 1.3, Shortlist 2010 A2

Although it's not necessary at all we'll remark that equality holds in the upper bound when $(a,b,c,d) = (0,2,2,2)$ and permutations, and equality holds in the lower bound $(a,b,c,d) = (3,1,1,1)$ and permutations. This will motivate the solution to follow.

Let $S = \sum_{\text{cyc}}(4a^3 - a^4)$. The upper bound on S is proved by noting that

$$\begin{aligned} 0 \le \sum_{\text{cyc}} a^2(a-2)^2 &= \sum_{\text{cyc}} \left(a^4 - 4a^3 + 4a^2\right) \\ &= -S + 4(a^2+b^2+c^2+d^2) = 48 - S. \end{aligned}$$

For the lower bound, we need an additional estimate.

Claim — We have $a,b,c,d \in [0,3]$.

Proof. As $3(b^2+c^2+d^2) \ge (b+c+d)^2$, we get $3(12-a^2) \ge (6-a)^2 \implies 0 \le a \le 3$. □

With this, we may now write

$$\begin{aligned} 0 \ge \sum_{\text{cyc}} (a+1)(a-1)^2(a-3) &= \sum_{\text{cyc}} \left(a^4 - 4a^3 + 2a^2 + 4a - 3\right) \\ &= -S + 2(a^2+b^2+c^2+d^2) + 4(a+b+c+d) - 12 \\ &= -S + 2 \cdot 12 + 4 \cdot 6 - 12 = 36 - S \end{aligned}$$

as desired.

Remark (Motivational remark)**.** The "magical" nature of the inequalities here is due to some behind-the-scenes precision about how they were selected. Essentially, once the equality cases are discovered, the essential idea is to try and design inequalities of the form

$$a^4 - 4a^3 + ka^2 + \ell a + m \quad \ge \text{ or } \le \quad 0$$

or the reverse, to sum cyclically.

It suffices for these inequalities to sharp at $a \in \{0,2\}$ in the upper bound, and $a \in \{1,3\}$ in the lower bound.

For the upper bound this works out fine verbatim. In fact, one does not even need $a+b+c+d = 6$ for the upper bound.

For the lower bound, it is necessary to first put bounds on a, since the natural $(a-1)^2(a-3)^2 \ge 0$ would yield an $-8a^3$ coefficient which is not what we want. However, we recognize that because we have a Cauchy-Schwarz inequality $(b+c+d)^2 \le 3(b^2+c^2+d^2)$, it should be possible to obtain some bounds for a. But those bounds are sharpest when $b=c=d$, and so we even know in advance that we're going to get $0 \le a \le 3$ (because the two equality cases we identified at the very beginning satisfied $b=c=d$). By recognizing that we have $a \le 3$, we no longer need the square factor on $a-3$ and can choose the last factor $a+1$ to force the $-4a^3$ term.

This is thus a superb example of how knowing the equality case can lead to a solution which would be impossible to motivate otherwise. It is used three times: once to conjure $a^2(a-2)^2 \ge 0$, once for the observation that we ought to have $0 \le a \le 3$, and once more to conjure $(a+1)(a-1)^2(a-3) \le 0$. Each time, we *know in advance* that the bound will work by Sharpness Principle.

§3.4 Solution 1.4, Putnam 2014 B2

Let g be the function which is 1 on $[1,2]$ and -1 on $[2,3]$. Note that

$$\begin{aligned}\int_1^3 \frac{g(x)-f(x)}{x}\,dx &= \int_1^2 \frac{1-f(x)}{x}\,dx + \int_2^3 \frac{-1-f(x)}{x}\,dx \\ &\ge \int_1^2 \frac{1-f(x)}{2}\,dx + \int_2^3 \frac{-1-f(x)}{2}\,dx \\ &= 0.\end{aligned}$$

Hence, answer is

$$\int_1^3 \frac{g(x)}{x}\,dx = \log 2 - \log(3/2) = \log(4/3).$$

§3.5 Solution 1.5, ELMO 2013/1

The answer is 28, achieved when $a_1 = \dots = a_8 = 0$ and $a_9 = 1$.

Here are three ways to see that $A \ge 28$ is necessary.

- **Linearity of expectation**: Let p be the probability a random triple is $\ge 3m$. Consider a random permutation π and the expected value of the number of triples $\{a_{\pi(3i+1)}, a_{\pi(3i+2)}, a_{\pi(3i+3)}\}$ with sum $\ge 3m$ for $i = 0,1,2$. On one hand it is ≥ 1, and the other hand it is $3p$, so $p \ge 1/3$ as needed.

- **Double counting**: Consider all

  $$C = \frac{1}{6}\binom{9}{3,3,3} = 280$$

  choices of three disjoint triples. So within each of C groups, there is at least one triple $\ge 3m$.

  Consider a triple $t = (a_i, a_k, a_k)$ with sum $\ge 3m$ (there are A of these). It appears $\frac{1}{2}\binom{6}{3} = 10$ times. Therefore:

  $$\frac{1}{2}\binom{6}{3} \cdot A \ge C = \frac{1}{3!}\binom{9}{3,3,3} = 280.$$

  This implies $A \ge 28$.

  (This is really the same as the previous solution.)

- **Explicit partition**: We explicitly break the $\binom{9}{3} = 84$ triples into 28 groups. Imagine the following array on a torus:

  $$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

  We give the following partition.

  - The all-rows partition 123|456|789 and the all-columns partition 147|258|369.
  - The two "diagonal" partitions 159|483|267 and 168|249|357.
  - Six "shift downwards cyclically" partitions: 126|459|783, 129|453|786, 135|486|792, 183|426|759, 423|756|189, 723|156|489.

- Six "shift right cyclically" partitions: 148|259|367, 149|257|368, 157|286|349, 167|248|359, 247|358|169, 347|158|269.
- The four "northeast L-shape" partitions: 125|893|647, 458|236|971, 782|569|314.
- The four "northwest L-shape" partitions: 214|397|658, 547|631|984, 871|964|325.
- The four "southwest L-shape" partitions: 145|697|823, 478|931|256, 712|364|589.
- The four "southeast L-shape" partitions: 245|689|713, 578|923|146, 812|356|479.

At least one of the triples in each of the 28 partitions is large, as needed.

§3.6 Solution 1.6, EGMO 2013/6

Let Q_n denote the vector space $\{0,1\}^n$. For a vector v, let $v[i]$ denote the ith component. We may identify each day with a vector v_k, where $v_k[i] = 0$ if dwarf i worked in the diamond mine, and $v_k[i] = 1$ otherwise. Let $V = \{v_1, v_2, \dots, v_{16}\}$ be the subset of Q_7 in the problem, and assume $v_{16} = 0000000$.

We first prove the following:

Lemma

Exactly two vectors start with each of 000, 001, ..., 111. Similar statements hold for any choice of three indices.

Proof. If three vectors start with 000 (say) then we run into problems. □

Now we know that v_{16} is the all-null vector. Ignoring that vector (and hence considering just the first 15 vectors), let n_i be the number of 1's in v_i, where $i = 1, 2, \dots, 15$.

Claim — We have

$$\sum_{i=1}^{15} \binom{n_i}{1} = \frac{16}{2}\binom{7}{1} = 56$$
$$\sum_{i=1}^{15} \binom{n_i}{2} = \frac{16}{2^2}\binom{7}{2} = 84$$
$$\sum_{i=1}^{15} \binom{n_i}{3} = \frac{16}{2^3}\binom{7}{3} = 70.$$

Proof. This follows using the lemma and double counting. For example, the third equation is counting the number of quadruples (i, j_1, j_2, j_3) such that the ith string has a 1 in the j_1th, j_2th, j_3th position, where $j_1 < j_2 < j_3$. The left-hand side counts the number when summed by i; the right-hand side counts the number according to the $\binom{7}{3}$ choices of (j_1, j_2, j_3). □

We may then rewrite the lemma as

$$\sum_{i=1}^{15} n_i = 56$$
$$\sum_{i=1}^{15} n_i^2 = 2 \cdot 84 + 56 = 224$$

$$\sum_{i=1}^{15} n_i^3 = 6 \cdot 70 + 3 \cdot 224 - 2 \cdot 56 = 980.$$

Now remark that $(n_i - 3)(n_i - 4)(n_i - 7) \le 0$ for each integer $1 \le n_i \le 7$. We compute

$$\begin{aligned} 0 &\ge \sum_{i=1}^{15} (n_i - 3)(n_i - 4)(n_i - 7) \\ &= \sum_{i=1}^{15} n_i^3 - 14n_i^2 + 61n_i - 84 \\ &= 1 \cdot 980 - 14 \cdot 224 + 61 \cdot 56 - 15 \cdot 84 \\ &= 980 - 3136 + 3416 - 1260 \\ &= 0. \end{aligned}$$

This implies that $n_i \in \{3, 4, 7\}$ for each i. If $n_i \neq 7$ for any i then we have

$$224 = \sum n_i^2 \equiv 15 \cdot 2 \not\equiv 0 \pmod 7$$

which is impossible. This means that $n_i = 7$ for some i, and therefore, $1111111 \in V$, as desired.

Remark. Up to permutation there turns out to be a unique set of 16 vectors. Xinyang Chen gives the following compact description of the unique solution:

- 0000000 and 1111111, of course
- 1101000 and its six *cyclic shifts* 0110100, 0011010, 0001101, 1000110, 0100011, 1010001.
- The complements of the above seven strings (i.e. 0010111, 1001011, etc.).

§3.7 Solution 1.7, Shortlist 2018 C5

The required minimum is $k(4k^2 + k - 1)/2$. (The first few answers are 2, 17, 57, 134.)

¶ Construction. We now give an example of a tournament achieving the minimum, although we postpone the calculation until the end of the solution. First we illustrate an example when $k = 4$, where the eight players are named A, B, C, D, W, X, Y, Z. The arrival dates of players are marked in green, while the departure dates are marked in red.

Day 1	**A**	**B**
Day 2	**C**	A
Day 3	C	B
Day 4	**D**	A
Day 5	D	B
Day 6	D	C
Day 7	**W**	A
Day 8	W	B
Day 9	W	C
Day 10	W	D
Day 11	**X**	A
Day 12	X	B
Day 13	X	C
Day 14	**Y**	A
Day 15	Y	B
Day 16	**Z**	**A**
Day 17	Z	**B**
Day 18	Y	C
Day 19	Z	**C**
Day 20	X	D
Day 21	Y	D
Day 22	Z	**D**
Day 23	W	X
Day 24	W	Y
Day 25	**W**	Z
Day 26	X	Y
Day 27	**X**	Z
Day 28	**Y**	**Z**

It should be clear how this adapts to general k. To spell it out, we denote the players $P_1, \dots, P_k$ and $Q_1, \dots, Q_k$ in general. In the first $\binom{k}{2}$ days, the players $P_1, \dots, P_k$ arrive in order, and when each player arrives they play everyone else. The last $\binom{k}{2}$ are done in a symmetric way. In the middle k^2 days we have two phases. First $Q_1, \dots, Q_k$ arrive in order; when Q_i arrives they immediately play $P_1, \dots, P_{k+1-i}$. Then $P_1, \dots, P_k$ depart in order; before P_i departs they play the rest of their matches.

¶ **Proof of the bound.** Let us denote the arrival and departure dates as

$$\begin{aligned} 1 = a_1 \le a_2 \le \dots \le a_{2k} \\ \binom{2k}{2} = b_1 \ge b_2 \ge \dots \ge b_{2k} \end{aligned}$$

in sorted order (not necessarily corresponding to players!). If a player arrives on date a_i and departs on date b_j the cost is $b_j - a_i + 1$, so nonetheless the organizers pay a total of

$$S = \left(\sum_j b_j\right) - \left(\sum_i a_i\right) + (2k) = \sum_n (b_n - a_n + 1).$$

Claim — We have $a_n - 1 \le \binom{n-1}{2}$ and $\binom{2k}{2} - b_n \le \binom{n-1}{2}$ for each $n = 1, 2, \dots, k$.

Proof. This is obvious: before the nth arrival, there were at most $n-1$ players, so at most $\binom{n-1}{2}$ days passed. The other bound is analogous. □

The previous claim of course works for $n \ge k+1$ as well. But we can actually get a better bound in this case (well, the same bound for $n = k+1$, but better when $n \ge k+2$).

Claim — For $n = k+1, k+2, \dots, 2k$ we have

$$(a_n - 1) + \left(\binom{2k}{2} - b_n\right) \le 2\binom{n-1}{2} - \binom{2(n-1)-2k}{2}$$

Proof. The idea of the proof is the same as before. Before the nth arrival, at most $\binom{n-1}{2}$ games were played; after the nth departure, at most $\binom{n-1}{2}$ games were played.

However, when $n > k$ the sum of two binomials double-count certain games. Specifically, there are $n-1$ players corresponding to $a_1, \dots, a_{n-1}$ and $n-1$ players corresponding to $b_1, \dots, b_{n-1}$. So there are at least $2(n-1) - 2k$ players overlapped; the estimate above counts their games twice, but they really should be counted at most once. This gives the correction term above. □

Remark (Yannick Yao)**.** The right-hand side in the second claim evaluates to $\binom{2k}{2} - (2k - n + 1)^2$ which may be thought of more naturally as all the games except for the games in a "complete bipartite graph" between two sets of $2k - (n-1)$ players.

From this, we get the estimate

$$S = \sum_{n=1}^{2k} (b_n - a_n + 1)$$

$$
\begin{aligned}
&\geq \sum_{n=1}^{2k} \left(\binom{2k}{2} - 2\binom{n-1}{2} \right) + \sum_{n=k+2}^{2k} \binom{2(n-1)-2k}{2} \\
&= 2k\binom{2k}{2} - 2\sum_{n=1}^{2k} \binom{n-1}{2} + \sum_{n=k+2}^{2k} \binom{2(n-k-1)}{2} \\
&= \frac{1}{2}k(4k^2+k-1)
\end{aligned}
$$

after some calculation. (Finally, since the example we exhibited at the beginning satisfies all the inequalities in the earlier claims, it achieves the same value.)

> **Remark.** The equality case is not unique, as it's easily possible to for example switch days 20 and 21 in the construction for $k = 4$, et cetera. That's why it is so much more fruitful to define the a's and b's in order rather than relative to the players; one could have, say, defined a_i and b_i to be the arrival and departure dates of the ith player, but then these variables change wildly in the equality case.